# Reeves' Muntjac: a Molecular Genetic Study of an Invading Species

by

Timothy Roy Williams

A thesis submitted to the University of Kent for the Degree of Doctor of Philosophy in the Faculty of Natural Science.

Durrell Institute of Conservation and Ecology 1994

No part of this thesis has been submitted in support of an application for any degree or qualification of the University of Kent or any other University or institute of learning

Timothy Roy Williams

# ACKNOWLEDGEMENTS

This study was funded by a Science and Engineering Research Council Studentship and the Durrell Institute of Conservation and Ecology. Laboratory facilities were provided at the University of Kent by Dr M.F. Tuite and the Biological Laboratory, and at the Institute of Zoology (London) by Dr. R.K. Wayne and the Conservation Genetics Group. The collection of muntjac samples was organised and funded by Professor Steven Harris of Bristol University, with whom the Institute of Zoology are in collaboration.

I am grateful to the above organisations and individuals for providing a unique opportunity to allow me to carry out this work in stimulating, sometimes demanding, but always encouraging environments. In a work of this nature many people make valuable contributions, often to a greater extent than they realise at the time. I am indebted to Dr. Mick Tuite and the many people in his lab for patient guidance during my formative period when molecular genetics was, to me, an unknown planet. Particularly Dave Colhurst for patience and an introduction to the basics, and Ian Stansfield for occasional theoretical discussions.

Because of the nature of this work almost everyone at the Institute of Zoology has made a valuable contribution. I am particularly grateful to Susan Haines for initial practical guidance with DNA fingerprinting and to Dave Chessman for the same in DNA amplification by PCR. Dr. Heather Hall introduced me to the world of mtDNA and, in collusion with Dr. Helen Stanley, kept me supplied with innumerable primers, advice and consolation.

There are many people not mentioned here but who will be remembered simply because they gave a smile or friendly gesture at a fraught time, and thus had a great encouraging effect. In spite of every effort by colleagues, the errors and omissions remain exclusively my own.

# ABSTRACT

The possibility of screening large numbers of individuals directly for genetic variability at the DNA level has only become feasible within the last decade. Prior to this, gel electrophoresis was employed to resolve small differences in the amino acid sequences of proteins. This thesis describes the application of three molecular genetic techniques to elucidate the relationships within sub-populations as well as the distribution and relationships between groups of sub-populations of an invading species of deer; Reeves muntjac (*Muntiacus reevesi*). Samples were taken from eighteen U.K. sub-populations of muntjac, representing increasing geographic distances centred on Woburn Abbey (the putative centre of origin) and one group from Taiwan were investigated.

An appropriate DNA fingerprinting protocol was established, based on published protocols and two commercially available minisatelite probes. Detailed analysis of the banding patterns produced by these probes allowed an estimate of the level of inbreeding within sub-populations to be established. Tentative relationships were also able to be established between the sub-populations sampled. Although there was some overlap between probes, they generally detected independent loci. The mean band sharing coefficients were found to range from 0·19 to 0·33 (probe 33·15) and 0·15 to 0·34 (probe 33·6).

The polymerase chain reaction was used to amplify the control region of muntjac mitochondrial DNA prior to digestion with restriction enzymes. The restriction fragments produced were found to be sufficiently informative to identify eight maternal lineages and demonstrate that genetically closely related groups are in fact geographically widely separated, a finding counter to the hypothesis of a regular mode of dispersal. Genetic partitioning between sub-populations was assessed by three agglomerative clustering methods which demonstrated that there is very little geographic partitioning between sub-populations, a result which may be expected given the short time since the introduction of muntjac to the U.K.

Genetic variability between sub-populations was also assessed by an investigation of the geographic distribution of seven microsatellite loci. Most microsatellite loci were found to be highly polymorphic and there was substantial variation in the number of alleles detected per sub-population. Heterozygosity values were found to be high, ranging from 0·48 to 0·74. Fourteen 'rare' alleles were uncovered distributed between ten of the eighteen sub-populations. Wright's $F_{IS}$ was calculated as a metric of the level of inbreeding within sub-populations. The mean values of this estimator ranged from -0.181, indicating heterozygosity, to 0.222, indicating homozygosity. Nei's $G_{ST}$, an analogue of Wright's $F_{ST}$ was calculated to give an indication of the level of genetic sub-structuring between sub populations. These calculations indicated that differentiation by distance is not significant over the geographic range of the sample area, again indicating that the muntjac have not dispersed in a 'natural' way.

The use of three different molecular techniques has allowed comparisons to be made between their relative merits in terms of the level of genetic information they provide and in their ability to define population structure.

TABLE OF CONTENTS

## List of Tables and Figures

# List Of Abbreviations used in this Thesis

| | |
|---|---|
| ACD | Citric acid, sodium citrate, glucose |
| ADH | Alcohol dehydrogenase |
| ANOVAR | Analysis of variance |
| APD | Average percentage difference |
| bp | base pairs |
| Bdp | Bowden Park (sample site) |
| Bkp | Brocket Park (sample site) |
| BSA | Bovine serum albumin |
| BSC | Band sharing coefficient |
| Btf | Buntingford (sample site) |
| CEPH | Centre d'Etude Polymorphisme Humain |
| CFTR | Cystic fibrosis transmembrane conductance regulator gene |
| dCTP | deoxyCytosine Triphosphate |
| df | degrees of freedom |
| D-loop | Displacement loop |
| DM | Myotonic Dystrophy |
| DMSO | Dimethyl Sulphoxide |
| EDTA | Ethylenediaminetetraacetic acid |
| Elg | Eling Common (sample site) |
| EPC | Extra Pair Copulation |
| EPF | Extra Pair Fertilisation |
| Eus | Euston (sample site) |
| Hex | Hexton (sample site) |
| Hpk | Harrys' Park (sample site) |
| H-Strand | Heavy strand |
| ISBP | Intraspecific Brood Parasitism |
| JMR | Juvenile mortality rate |
| kb | kilo bases |
| Knb | Knebworth (sample site) |
| Kf | King's Forest (sample site) |
| LDH | Lactate dehydrogenase |
| L-Strand | Light strand |
| Min | Minsbury Hill (sample site) |
| MS | mean square |
| mtDNA | Mitochondrial DNA |
| Myr | Million years |
| ND (gene) | Nicotinamide adenine dinucleotide dehydrogenase (gene) |
| NHN | Northern Hairy-Nosed Wombat |
| Nth | Northill (sample site) |
| OD | Optical Density |
| OLB | Oligo labelling buffer |
| ORF | Open Reading Frame |
| Owd | Oldwarden (sample site) |
| PCR | Polymerase Chain Reaction |
| PGI | Phosphoglucose Isomerase |
| RFLP | Restriction Fragment Length Polymorphism |
| rRNA | Ribosomal RNA |
| Sal | Salcey Forest (sample site) |
| Sbl | Southblock (sample site) |
| SD | standard deviation |
| SLS | Sodium Lauryl Sulphate |

| | |
|---|---|
| SHN | Southern Hairy-Nosed Wombat |
| SS | sum of squares |
| SSC (20x) | 3M NaCl, 0.3M Na-citrate |
| TBE | 90mM Tris, 90mM boric acid, 1mM EDTA p*H* 7.8 |
| TE | 10mM Tris[hydroxymethyl]-aminomethane hydrochloride, 1mM EDTA p*H* 7.8 |
| Tris | Tris[hydroxymethyl]-aminomethane hydrochloride |
| tRNA | Transfer RNA |
| Tus | Tusmore (sample site) |
| Twn | Taiwan (sample site) |
| UPGMA | Unweighted Pair Group Method with Arithmetic means |
| V/cm | volts per centimetre |
| Wav | Waverly (sample site) |
| Wob | Woburn (sample site) |
| Wst | Wormsley Estate (sample site) |
| λDNA | Lambda DNA commercially isolated from bacteriophage lambda |

Chapter 1

# INTRODUCTION AND BACKGROUND TO THE PROJECT

## 1·1 Introduction

Population bottlenecks and the consequential effects of a diminution of genetic variation on the surviving small number of individuals are well known and have been the subject of many studies. Other than being of intrinsic value; reaching an understanding of the causes and consequences of diminished genetic variation is of primary importance in the genetic management of species that are in danger of becoming extinct. There have been a few studies on the consequences of bottlenecks in natural populations (Bonnell & Selander 1974, Ryman *et al* 1977, O'Brien *et al* 1983, Wildt *et al* 1987, Gilbert *et al* 1990, Packer *et al* 1991b, Roelke *et al* 1993) but many species have been through a bottleneck followed by a rapid population expansion. There are numerous records of the successful introduction and naturalisation of many species both in the British Isles and the rest of the world (Lever 1977, 1978).

An understanding of the genetic consequences of population expansion after a bottleneck event, as well as a knowledge of colonisation profiles and the mode of range expansions are of considerable importance in the formulation of conservation strategies for rare and endangered species.

The subject of this study is Reeves' muntjac (*Muntiacus reevesi*). Here I shall investigate the dispersal and origins of some of the U. K. sub-populations as well as the level of genetic similarity within and between populations using three molecular genetic techniques, (1) DNA fingerprinting, (2) restriction fragment analysis of a region of mitochondrial DNA, and (3) microsatellite DNA analysis.

Reeves' muntjac is a diminutive deer that originates in China and Taiwan (Whitehead 1972), comparatively little is known about their native habitats other than what has been summarised by Sheng (1992). Twenty four were introduced into the U. K. from within their native range at the beginning of this century (Chapman 1991a) and since then, as a consequence of either escaping or being released, they have dramatically expanded their range to cover most of southern England (Anderson & Cham 1987). They are one of the few species of mammal that have been through a population bottleneck followed by a rapid population expansion, and which hold the potential for a thorough investigation. We are fortunate in knowing the number of founder individuals and in having the source population still essentially intact.

The present-day status of muntjac deer presents an ideal situation in which the genetic consequences of a bottleneck event followed by rapid population expansion, the rate of genetic loss associated with rapid colonisation and the effect of the size of founder populations on genetic variability can be studied. The rate and geographical spread of populations and the development of patterns of genetic variation within and between those populations can also be studied. Few previous studies (Bonnel & Selander 1974, O'Brien *et al* 1983, Wildt *et al* 1987, Packer *et al* 1991b [see section 1·10]) have had the opportunity to investigate such a successful recovery after a bottleneck event.

Reports of the origins of Reeve's muntjac in the U. K. are sparse and sometimes contradictory. According to Chaplin (1977a), Reeve's muntjac was introduced at Woburn about 1930, but Harding (1986) has reported that the Woburn record books show that 24 Reeve's muntjac were introduced between 1892 and 1905. Hastings (1949) maintains that the collection of animals at Woburn began in 'about' 1895.
The expansion of Reeve's muntjac has been recorded, although patchily, from 1922 onwards for the regions close to Woburn and the surrounding counties (Chapman 1991a). Their spread has been recorded by Whitehead (1964), Carne (1968), Clarke (1974), Lever (1977) and Dansie (1983, 1989). Corbet (1971) has constructed a distribution map taking account of all the records up to 1970, and Anderson and Cham (1987), using the same data, have produced a series of maps depicting the expansion of muntjac. By 1960 there were records from only 12 x 10km squares, but by 1985 this figure had increased to 280. Amos (1992) has tabulated the number of counties in which muntjac have been recorded per decade as: 1920s/2, 1930s/4, 1940s/10, 1950s/19, 1960s+1970s/31. These increases may of course reflect an increase in the efficiency of record gathering rather than an increase in range but reports from the Forestry Commission for 1973, 1978 and 1983 represent a standard recording regime and confirm an increase in the number of regions in which muntjac were observed (Forestry Commission unpublished data, Tee *et al* 1985).

There are three hypotheses associated with this work:-

- That muntjac radiated from the original founding population in a regular manner, i.e. dispersed radially from Woburn Abbey, given the limitations of suitable habitat. In this situation I would expect genetic variation to diminish with geographic distance. Radiations in different directions may be found to exhibit their own genetic 'lineages' and consequently sub-populations at the extremes of the radiations may be genetically distinct. This hypothesis assumes that the population density is such that new generations are constrained to migrate ever outwards.
- That muntjac dispersal from the original population has been directional and that discrete sub-populations have been founded. Emigrants from these sub-populations then dispersed to new habitats where, randomly, they may have come into contact with individuals from other sub-populations. This hypothesis assumes that the population density is low enough to allow emigration in any direction from the previously established sub-populations. The resultant pattern of sub-populations would be in the form of a lattice rather than the concentric circles envisaged in the former hypothesis. Under this hypothesis, genetic variation would be more homogenous and any variation between sub-populations may be expected regionally. For example the sub-populations may fall into a north-south or northeast-southeast divide.
- An open hypothesis, that the muntjac have dispersed in some other way that may be elucidated by this study.

### 1·2 A Brief Taxonomy of Muntjac Deer

The extant members of the deer family (Cervidae, order Artiodactyla) comprise about 40 species world wide and are divided into five subfamilies: (1) The Hydropotinae, Chinese water deer. (2) The Muntiacinae, muntjacs of India and Asia. (3) The Cervinae, cervids of Eurasia. (4) The Odocoilinae, a grouping which contains the moose (*Alces*), roe deer (*Capreolus*) and reindeer (Rangifer). (5) The Moschinae, a contentious subfamily that is included only by those who consider the musk deer (*Moschus*) of the Himalayas, China and Southeast Asia as being within the Cervid family.

The present-day Muntiacinae are all asiatic in origin and are regarded as a relatively primitive group, mainly on the basis of their diminutive size and the presence of enlarged, recurved canine tusks in the males. Early members of the Cervidae, now extinct, either had large canines and no antlers, as in *Blastomeryx* of the Miocene of North America, or they possessed a nondeciduous cranial bony outgrowth that was covered with skin, as in *Paleomeryx* of the Miocene of Europe (Young, 1977). Characteristically in muntjacs the bony pedicels from which the simple antlers develop in the males originate on the top of the skull, as in other deer, but extend down the face as two bony prominences, visible externally as distinctive ridges. Thus giving this species the common name 'Rib-faced deer'. Rarely, these antlers may develop a small brow tine. There are six species within the subfamily Muntiacinae: *Muntiacus muntjak* Zimmermann, 1780 (type species) the Indian (or Javan) muntjac of India, Sri Lanka, Tibet, southwest China, Burma, Thailand, Vietnam, Malaysia and Indonesia (15 subspecies). *Muntiacus Reevesi* Ogilby, 1839, variously known as the barking deer or rib-faced deer, Chinese muntjac or Reeves' muntjac of east China and Formosa (two subspecies), introduced into the U. K. *Muntiacus crinifrons* Sclater, 1855, the hairy-fronted or black muntjac of eastern China. *Muntiacus feae* Thomas & Doria 1889, Fea's muntjac of Thailand and Tenasserim. *Muntiacus rooseveltorum* Osgood 1932, Roosevelts' muntjac of north Vietnam. *Elaphodus ephalophus* Milne-Edwards, 1871, the tufted deer of Burma and central and southern China (three subspecies).

### 1·3 The Origins of Reeves' muntjac in the U. K.

*M. reevesi* was introduced into this country at Woburn Abbey, Bedfordshire, by the 11th Duke of Bedford at the beginning of this century (Hastings, 1949), and later in 1930 to Whipsnade Zoological Gardens. However, prior to this in the late 1890s the Indian muntjac (*M. muntjak*) was introduced to Woburn, from where in subsequent years a number escaped. They reproduced successfully both in the captive and feral state (Anderson & Cham 1987) but, due to their innate aggressive nature, it was decided to replace them with the more docile Reeves' muntjac. Whilst this was obviously easily done within the confines of the park, it was a more difficult task with the feral population. Later escapes and releases into the wild of Reeves' muntjac led to the supposition that the two species may co-exist (Lever, 1977) and possibly have interbred, or exist as hybrids (Bray, 1981). Soper (1969) states that '*M. reevesi* and its hybrid with the Indian muntjac have become widely distributed since both species were liberated into the outlying woods around Woburn'. The hybrid *M. reevesi* x *M. muntjak* being midway in size between the two species (Whitehead, 1964). One individual was shot in 1946 and identified as being a hybrid (Whitehead, 1950). Morphological studies of the skull and jaw of authentic adult Reeves' muntjac from China and Taiwan, authentic Indian muntjac, and English feral muntjac (Chapman & Chapman, 1982) have shown that the feral individuals are not the Indian species and are most probably Reeves'. The probability of there being hybrids could not, at that stage, be categorically excluded. Later work by Chapman *et al* (1983) has provided stronger evidence for the absence of both the Indian species and the hybrid. The Indian muntjac (*M. m. muntjak*) has a diploid chromosome number of nine in the male and eight in the female (Wurster & Atkin, 1972), whilst Reeves' muntjac (*M. reevesi*) has a diploid chromosome number of 46 in both male and female (Wurster & Benirschke, 1967). The hybrid would then be expected to be either 2n = 27 (female) or 2n = 28 (male). The hybrid *M. m. vaginalis* (2n = 7/female; 2n = 6/male) x *M. reevesi* (2n = 46) has a diploid chromosome number of 27 and 26, male and female respectively (Shi, Ye, & Duan, 1980). All the deer studied in this work of Chapman *et al* had a chromosome number of 46 and thus they failed to find evidence for the presence of either the Indian species or the hybrid. Furthermore, the disparity in chromosome number between these two species would suggest that the hybrid is likely to be infertile, as was the case of a male hybrid resulting from an Indian muntjac (female) x Reeves' muntjac (male) (Shi & Pathak, 1981). Any hybrids would be expected to be the progeny of viable sympatric populations of both species.

Sightings of feral muntjac date back to 1922 when one was seen at Wrest Park, 11km east of Woburn and two years later another was reported from Ashridge Park, 19km south of Woburn (Anderson & Cham 1987). These were probably the Indian species as no more sightings are reported from the 1920s, which coincides with the eradication of the Indian species and the introduction of Reeves' muntjac. Table 1·1, a record of the sightings from 1922 to 1974, has been constructed from data on sightings which have been taken from Lever (1977) and Anderson and Cham (1987).

| Date | Location | Location re: Woburn |
|---|---|---|
| 1922 | Wrest Park (Bedfordshire) | 11km East |
| 1924 | Ashridge Park (Hertfordshire) | 19km South |
| 1933/34 | Brackley (Buckinghamshire)<br>Salcey Forest (Buckinghamshire) | 38km West<br>25km N.West |
| 1934 | Ryton Wood (Warwickshire) | 60km West |
| 1934/39 | Rockingham (Northamptonshire) | 57km North |
| 1937 | Kelmarsh (? escape Croyland Abbey) | 52km N.N.West |
| 1938 | Broxbourne (Hertfordshire)<br>Ampthill Forest 'around' Luton (1938 Deer Census) | 50km S.East<br>10km N.East |

| | | |
|---|---|---|
| 1940 | Parham East Suffolk. | 140km E.N.East |
| 1941/?41 | Walton Wood - Kineton (Warwickshire)<br>Needwood Forest (Sudbury) | 70km N.West<br>95km East |
| 1941/?42 | Easthorpe (Colchester) | 100km East |
| 1947 | Banner Hill Farm(Warwickshire)<br>Oakley Wood (Warwickshire) | 80km N.West<br>68km N.West |
| 1948 | Hazelborough. (1948 Deer Census) | 34km N.West |
| 1949 | Bolderwood Forest/Alice Holt Forest (Hampshire) | 150km S.S.West |
| 1950 | Odell in Valley of the Ouse (Bedfordshire)<br>Overston (Northamptonshire)<br>Knightly Wood (Northamptonshire)<br>Kelmarsh /Weekley(Northamptonshire) | 25km North<br>38km N.N.West<br>50km N.West<br>50km North |
| 1951 | Wellsbourne Wood (Northamptonshire) | 60km N.West |
| 1952 | Forest of Arden (Warwickshire)<br>Stoke Lynne/Bicester (Oxfordshire) | 85km N.West<br>40km W.S.West |
| 1953 | Charnwood (Leicestershire)<br>Bayford (Hertfordshire) | 110km N.N.West<br>42km S.East |
| 1954 | Kilworth Hall (? escape from Croyland Abbey) | 60km N.West |
| 1955 | Outer Birmingham<br>Twyford (Berkshire) | 90km N.West<br>60km S.S.West |
| 1958 | Witheringtom Wood (Gloucestershire)<br>Dumbleton (Gloucestershire)<br>Enfield (Middlesex)<br>Hatley Wood (Cambridgeshire)<br>Watlington (in the Chilterns) | 95km S.West<br>147km West<br>50km S.S.East<br>60km N.East<br>48km S.West |
| 1963 | Tackley Wood (Oxfordshire)<br>Cheltenham (Gloucestershire) | 50km W.S.West<br>100km West |
| 1964 | Windsor Forest (Berkshire)<br>Swinley Forest, Ascot (Berkshire) | 58km South<br>65km South |
| 1966 | Highclere (Berkshire)<br>Barnes Common (Surrey)<br>South Downs (Sussex)<br>Trowbridge (Wiltshire) | 88km S.S.West<br>65km S.S.East<br>120km South<br>125km S.West |
| 1968 | East Dorset<br>Bury St. Edmunds (Suffolk)<br>Milford Common (Surrey) | 160km S.West<br>95km N.East<br>90km South |
| 1969 | Cranbourne (Dorset) | 155km S.West |
| 1970 | Thame (Oxfordshire)<br>Dorking (Surrey) | 15km S.West<br>80km South |
| 1971 | High Wycombe (Buckinghamshire)<br>Rickmansworth/Watford (Hertfordshire)<br>Henley (Oxfordshire)<br>New Market (Suffolk)<br>Cobham (Surrey) | 38km S.S.West<br>36km South<br>52km S.S.West<br>75km N.East<br>75km South |
| 1972 | Wantage (Berkshire)<br>Tingewick (Buckinghamshire) | 72km S.West<br>30km West |
| 1973 | Norwich (Norfolk) | 140km N.East |
| 1974 | Swaffham (Norfolk)<br>Mildenhall /King's Forest (Suffolk) | 120km N.East<br>88km N.East |

Table 1·1 A record of muntjac sightings from 1922 to 1974

## 1·4 Habitat Selection, Social Organisation and Mating Behaviour of Reeves' Muntjac

Reeves' muntjac are essentially forest dwelling animals although they have been found in a variety of habitats including scrubland, deciduous woodland and coniferous regions (Anderson & Cham 1987). Previously they were thought to be a very selective, dense forest species (Soper 1969, Barrette 1987), deciduous woods being preferred to coniferous woods as the latter impedes the growth of sufficient low cover (Soper 1969). Dansie (1970) also concluded that dense cover and available water was necessary for muntjac colonisation. However, studies in the King's Forest in Suffolk, a predominantly coniferous woodland with no standing water, have shown that variety of ground vegetation is more important than density of cover (Chapman *et al* 1985, Jewell & Bullock 1991). Reeves' muntjac is now known to be a flexible and opportunistic species and a generalist feeder. This ability to adapt to a range of habitats has no doubt enhanced their capacity for expansion.

Initial reports on muntjac social structure claimed that they lived in family associations of up to four individuals; early perceptions being that the male protected his female(s) and fawns by aggressive defensive behaviour (Dansie 1970). However, later work by Chaplin (1977b) and more recent work by Jewell and Bullock (1991) and Chapman *et al* (1993) has shown that muntjac are solitary and polygamous animals that come together for mating on a common feeding ground, but do not form social groups. Does are rarely seen with fawns due to the rapid development of fawns (weaning is at seven to eight weeks), and the habit of the fawn of lying apart from its mother (Chaplin 1977b). When a fawn is seen, a buck is occasionally also seen in their company (Jewell & Bullock 1991), but this is probably due to post-partum œstrus of the female rather than paternity (Chaplin 1977b). Radio tracking has indicated that particular males and females may associate for some years but this should not be taken to indicate a life-bond, as changes in association are known to occur (Jewell & Bullock 1991).

Muntjac are territorial in as much as individuals have home ranges. These home ranges however are not exclusive and frequently overlap (Chapman *et al* 1993). The regional spatial organisation of a group consists of overlapping doe ranges, the core areas of which are largely exclusive, but they may have some social cohesion (Chapman *et al* 1993). The buck's range is significantly larger than that of the doe and commonly extends over several doe's ranges, but is generally exclusive of the ranges of other adult bucks. There is some seasonal fluctuation in range boundaries, but overall the ranges are very stable (Harris *et al* 1990).

Muntjac are aseasonal breeders, producing young after a seven month gestation period (Chapman *et al* 1984). In captive animals the first conception can occur at five to six months, and as a result of post-partum œstrus, the doe can conceive within a few days of parturition. The female can therefore be almost continuously pregnant for all her reproductive life. The record to date in a captive female is 16 live births over a period of nine years and ten months (Jewell & Bullock 1991). Although less active during the summer period, bucks are fertile throughout the year (Chapman & Harris 1991). Unlike roe deer or fallow deer, muntjac exhibit minimal post natal contact between the female and her offspring, the young are quickly driven off to find their own territories (Chaplin 1977b, Anderson & Cham 1987). Jewell & Bullock (1991) report life expectancy to be in the order of 16 years for males and 19 years for females, although the age at which females cease to breed is not known.

### 1·5 The Dangers of Being a Small Population

Small populations are faced with a number of risks that singly, or in combination, increase the possibility of local extinction. Challenges to small populations can be classified as either intrinsic or extrinsic. Intrinsic factors encompass demographic processes arising from the random changes in birth, sex ratio and mortality rates. These processes may result in a loss of genetic variation, and hence decreased fitness as a consequence of founder effect, genetic drift and inbreeding. Extrinsic factors include environmental fluctuations due to changes in weather, food supply and the populations of competitors, predators, parasites etc. and natural catastrophes such as flood, fire, drought, famine etc., which may occur at random or regular intervals. These factors may have a profound influence on the dynamics of small populations. These factors also interact. For example, an epidemic may reduce the population size to a point where inbreeding depression becomes a major factor which decreases the demographic parameters of survival and reproduction.

Inbreeding in a small local population may be a positive genetic asset <u>if</u> even a very limited amount of emigration/immigration is likely to occur. Consider two separate inbreeding sub-populations that have become relatively homozygous and have a high average coefficient of relatedness. The progeny of a mating between two individuals from these sub-populations would be highly heterozygous - thus restoring a level of genetic variability within the sub-population. This dispersal-related heterosis may in some species be related to their social system, e.g. termites (Dawkins 1989).

However, in general the effect of any one individual on a population's viability is significantly less in large populations than in small ones. Shaffer (1981) and Ballou (1991) have both suggested that demographic risks will present but a minor challenge in all but very small populations, i.e. those composed of less than twenty individuals.

The distribution of genetic variability over the area encompassed by a population may significantly influence the evolutionary dynamics of that population (Lewontin 1974; Endler 1977; Brown 1979; Wright 1982). Thus, the geographic distribution of genetic variability has long been of primary interest in evolutionary studies. However there appears to be a lack of consensus on the spatial scale at which population genetic structuring is expected to develop (Dewey & Heywood 1988).

### 1·6 The Importance of Spatial Heterogeneity

Several studies have demonstrated the importance of spatial heterogeneity for the continued survival of the population (Rolf 1974; Lomniki 1980; Hastings 1982; Vance 1984) and for the maintenance of abundance (Myers 1976; Taylor & Taylor 1977; Hanski 1982, 1985). The continuous systems of interactive species, i.e. predator/prey, host/parasite or competitors can also depend, to a large extent, on habitat spatial heterogeneity (Levin 1974; Comins & Noble 1985; Holt 1985; Chesson & Murdoch 1986). The import role of spatial patchiness in maintaining the overall population is that, at any particular time, patches which have suffered a local population decline or extinction can be recolonised by dispersing individuals from nearby patches which have not suffered a decline or extinction. If we assume that the factors causing a local decline or extinction do not affect all patches in the same way at the same time, then there will always be patches containing colonisers and therefore, long-term, the probability of the survival of the population will be enhanced.

Models that include spatial patchiness do so by assuming that either (1) all dispersing individuals have equal access to all habitat patches or, (2) dispersing individuals from each habitat patch distribute themselves equally among all neighbouring patches (Fahrig & Paloheimo 1988).

The maintenance of a patchy distribution is dependent upon the continued existence of dispersal corridors connecting the various patches (Fahrig & Paloheimo 1988). Should these corridors become hindered by (for example) human activity, then over time, the spatial relationships among patches may cease to be obvious. The past patterns of dispersal among patches will still be evident though in the genetic profiles of those patches. Theoretically, genetic differentiation over short distances in the absence of gene flow, or in the presence of restricted gene flow, will occur as a result of localised genetic drift. Genetic drift, localised or otherwise, encompasses several sources of variation including;

(1) founder effect; in which the founders provide only a limited range of the total genetic variability available.

(2) Parametric variation; caused by differences among loci in mutation rates (Slatkin & Arter, 1991).

(3) Stochastic effects; demographic stochasticity and random effects of the environment.

(4) Sampling effects. Sampling of course has no relationship with any genetic process but the interpreted frequency of an allele, and thus our assessment of a genetic parameter, will ultimately be based on the method of sampling.

Commonly a species will be found to be distributed over a large area containing many populations, the population dynamics of which will be interactive (Ballou 1991, Foose 1991). One of the commonest causes of the creation of a metapopulation is habitat fragmentation. The resultant non-uniform spatial distribution is composed of habitable patches with intervening regions that are of poorer quality (Gilpin 1987). The dynamics of metapopulations depend to a large extent on the rate of extinction and re-colonisation between patches, the probability of extinction in a single patch being inversely proportional to the carrying capacity of that patch - all other parameters being equal (Gilpin 1987). A reduction in patch area will increase its populations vulnerability to environmental catastrophes and lead to a loss of genetic diversity due to a (probable) increase in inbreeding or drift. Another consideration of reduced patch size is the area to boundary ratio; the smaller the patch the proportionally larger the boundary. A reduced boundary can give rise to two categories of 'edge' effects, (1) a deterioration of habitat quality near the boundary and (2) the expulsion of individuals into unsuitable regions where they may fail to reproduce or perish prior to attaining another suitable habitat. The minimum patch size on which a population can continue to survive - the critical patch size - is determined as a function of the dispersal rate of individuals into unsuitable areas (Lande 1988).

There are many different patterns to metapopulation structure. For example, in some species the population structure consists of a core area with peripheral isolates, whilst in others the structure may consist of small clusters of more-or-less equal status. An example of the first structure may be seen in the grizzly bear (*Ursus arctos*) populations of North America. The core area is Glacier National Park and the surrounding area whilst peripheral isolates reside in the Cabinet-Yaak region of Idaho and in the Yellowstone ecosystem (Shaffer 1981, Gilpin 1987). A good example of the second structure mentioned above is the northern spotted owl (*Strix caurina occidentalis*) of North America. The total population of this species was estimated to be 2,500 breeding pairs; an effective population of 500 breeding pairs was considered to be appropriate to maintain sufficient genetic diversity for the population to be able to adapt to environmental change. Consequently protection for 500 pairs was afforded, but subsequent logging of unprotected areas produced a fragmented habitat in which suitable regions for the owl were too sparsely distributed to support the population (Lande 1988), with the concomitant formation of regional isolates (Gilpin 1987).

The probability of migration between patches is considered (as a first approximation) to decrease exponentially with distance between the patches. It follows from this that decreasing distance between patches may be expected to increase the long-term survival of the metapopulation by enhancing migration. However, the opportunity to migrate does not necessarily mean that migration will occur or that breeding will occur between migrants. The importance of this can be seen in the example of the grizzly bear, mentioned above. Isolation of the Yellowstone sub-population has led to inbreeding and the associated loss of fitness (Shaffer 1981).

### 1·7 The Effects of Population Structure

The amount of gene flow between populations will affect genetic structure and is therefore an important parameter in coming to an understanding of the genetic diversity and continuing genetic evolution of populations. Slatkin (1987) has identified two different groups of methods for estimating gene flow in populations. The first group of methods directly estimate gene flow by an analysis of dispersal and either effective population size or neighbourhood size. The second group of methods indirectly estimate gene flow by analysing the geographic distribution of allele frequencies. However, direct estimates describe current levels of gene flow whilst indirect estimates represent different levels of gene flow over time and must therefore include past events which could have altered the distribution of allele frequencies. Indirect estimates of gene flow must therefore be interpreted bearing in mind the stochastic events which will have affected the development of the population and led to its status at the time of sampling.

A mathematical model of genetic differentiation of populations was first proposed by Wright (1931, 1943, 1951, 1965, 1969), in which he examined the situation of a founder population split into a large number (theoretically infinite) of isolated populations. Within the confines of this model there is no migration and random mating occurs within populations.

Any deviation from Hardy-Weinberg proportions within these isolated populations may be measured by a single parameter (F), the fixation index, which can be positive, zero or negative. Selection may increase or decrease the F value; if heterozygotes have a lower fitness than both homozygotes then F will become positive. If the converse is true, F will become negative. However, a negative F does not necessarily mean that there is a heterozygous advantage (Lewontin & Cockerham, 1959). Inbreeding increases the frequency of homozygotes and, in the absence of other factors, F will always be positive. Assortive mating has a similar effect to inbreeding in that the frequency of homozygotes increase, (and therefore F will be positive) but the effect is limited to only those loci concerned with the character trait(s) with which assortive mating occurs (Karlin 1969, Wright 1969).

### 1·8 Subdivision of Populations

It is often very difficult to be certain of the breeding structure in divided natural populations. Even if Hardy-Weinberg equilibrium is maintained in each sub-population the overall genotype frequency of the entire population may deviate. In this case F would be 1 when populations are fixed for different alleles and equal to 0 when there is no subdivision. Wright measured the deviation of genotype frequencies in a subdivided population in terms of three parameters; $F_{IS}$ , $F_{IT}$ and $F_{ST}$.

In Wright's terms $F_{IS}$ describes the relationship between two gametes relative to the sub-population and $F_{IT}$ the same relationship relative to the total population. $F_{ST}$ is the correlation between two gametes drawn at random from each sub-population and estimates the degree of genetic differentiation of the sub-populations. These parameters are related by:-

$$1 - F_{IT} = (1- F_{IS}) \ / \ (1-F_{ST}).$$

Ralls, *et al* (1986), however have suggested caution in interpreting F, or any estimator of F because *"F must be defined with respect to some reference generation or population (or both) and thus is a relative rather than an absolute measure"*. The values of F will also be influenced by many factors such as; assortive mating, family groups or groups of siblings in the sample, differences in allele frequency between parents, selection acting on animals before the sample is taken and sex-biased dispersal. As Jacquard (1974) mentions, *"An inbreeding coefficient cannot be regarded as an estimate of any real quantity"*.

## 1·9 Natural Genetic Variation in Feral Populations

In most natural populations genetic variation exists in the form of multiple alleles at a locus. By the very nature of this variation, it is commonly not apparent at the level of the phenotype. To be able to quantify this variation in relation to population differences it is necessary to have some convenient metric of variation that can be universally applied. Most commonly a measure of heterozygosity is employed. Heterozygosity may be defined as the presence of two different alleles at a particular locus in a diploid individual, and provides a measure of the genetic variation either in a population (the frequency of individuals heterozygous at a locus) or in an individual (the proportion of loci that are heterozygous) (Mitton & Pierce 1980). There are however some serious difficulties with the concept of heterozygosity as generally applied to individual and population studies. Individual heterozygosity of the entire genome is effectively impossible to measure. The genome of *Drosophila* contains approximately 5,000 genes, whilst the human genome is estimated to contain some 50,000 genes (Mitton 1993). Electrophoretic and molecular genetic surveys may examine only a very small percentage of the genome. A study of 50 genes would be a major undertaking but would only investigate a random 1% or 0·1% respectively of the genomes of *Drosophila* or man. From data such as these an adequate estimate of individual heterozygosity across the entire genome cannot reasonably be made (Mitton & Pierce 1980, Chakraborty 1981). This problem can be overcome to some extent by the application of appropriate sampling procedures and statistical analysis.

Archie (1985) has provided a theoretical study that claimed that data from a few individuals ($n \leq 25$) but many loci ($\gg 40$) are required to assess low heterozygosity. But rather than sampling loci independently for detailed heterozygosity estimates, most population studies are generally concerned with detecting polymorphism at particular loci.

Sjögren and Wyöni (1992) have suggested that few authors (*viz.* O'Brien *et al* 1983, Lesica *et al* 1988) have considered their sampling procedures on the very low levels of electrophoretic variation they found, and present a mathematical model that demonstrates that the effect of sample size in detecting genetic variation may have been underrated. The model showed that small samples yielded a significant risk of scoring monomorphism in polymorphic populations.

The above notwithstanding, a positive correlation between various aspects of fitness and heterozygosity has been shown to exist; for example, a positive correlation between heterozygosity and fecundity in *Drosophila* (Serradilla & Ayala 1983), between heterozygosity and adaptation to various temperature regimes in killifish, *Fundulus heteroclitus* (Mitton & Koehn 1975), and between heterozygosity and both the lactate dehydrogenase (LDH) and phosphoglucose isomerase (PGI) loci and survival during starvation in atlantic cod, *Gadus morhua* (Mork & Sundnes 1985).

## 1·10 Population Bottlenecks, Founder Effect and Genetic Drift

Population bottlenecks occur when a population is drastically reduced in number to near extinction. The effect of a population bottleneck can be to severely reduce the genetic variation available for succeeding generations. A bottleneck can result from several causes such as; disease or environmental catastrophes, a reduction in range leading to the creation of isolated peripheral populations or founding events. A founder event occurs when a single individual or a very limited number of individuals originate a new isolated population. The founder(s) may represent a small sample of the original gene pool to which they belonged. Natural selection acting on the restricted sample of variation yields gene combinations that are different from those in the ancestral population. Within a small isolated population the random sampling of gametes will give rise to chance fluctuations in allele frequencies and consequential genetic drift (Hartl & Clark 1989). Genetic drift refers to random changes in allele frequency due to sampling variation of gametes from generation to generation. Genetic drift, which may be particularly rapid following a bottleneck event, will lead to fixation of some alleles in small populations, although not all genes or alleles may be equally affected. Carson (1975), from his work on Hawaiian *Drosophila* (Carson 1968), has suggested that, for the purpose of understanding the effects of population bottlenecks, founder effect and genetic drift, the genome can be divided into two parts. A 'closed' part that probably provides regulatory functions and thus is highly co-adapted and balanced, and an 'open' part that is relatively free of epistatic interactions and able to vary without having a major effect on the phenotype. This of course only applies to those genotypes that have a sufficient survival value.

When multiple alleles are present in a population, founder events and/or bottleneck events, particularly those of short duration, often have a greater effect on the number of alleles than on the level of heterozygosity (Allendorf 1986). Rare alleles, advantageous and disadvantageous, are especially susceptable to loss. Allendorf (1986) has calculated that an allele at a frequency of 0·01 has a 60% chance of being lost following a bottleneck of 25 individuals.

Species that are undergoing a bottleneck event are especially susceptible, at least until their number increases. The Florida panther (*Felis concolor coryi*) exists as a small relic sub-population of approximately 30 individuals in southern Florida (Roelke *et al* 1993). The recent drastic reduction in population size coupled with the occurrence of aberrant morphological features (Barone *et al* 1993, Roelke *et al* 1993), the presence of a single mtDNA haplotype (O'Brien *et al* 1990) and the presence of physiological and endocrine impairments that have led to the loss of reproductive success (Barone *et al* 1993) are all factors consistent with the occurrence of one or more bottlenecks events followed by inbreeding (Maher 1990, O'Brien *et al* 1990). It has been estimated that this subspecies will go to extinction within 25-40 years (Seal & Lacy 1989).

Species which have increased their numbers after a bottleneck event will continue to have high levels of homozygosity and may have associated fitness problems. There are several well known examples of species surviving natural bottleneck events, for example; those of the elephant seals (Bonnell and Selander 1974), the Scandinavian moose (Ryman *et al* 1977) the polar bear (Allendorf *et al* 1979), the cheetah (O'Brien *et al* 1983), the California Channel Island fox (Gilbert *et al* 1990), the lions of the Gir Forest of western India (Wildt *et al* 1987) and the lions of the Ngorongoro Crater (Packer *et al* 1991b).

Bonnell and Selander (1974) found no variation at 24 loci in the northern elephant seal (*Mirounga angustirostris*) and attributed the high level of monomorphism to a decimation of the population by 19th century hunters. The estimated 20 individuals that survived this onslaught have expanded to a present-day population estimated to be in the region of 30,000. A similar situation exists in the Scandinavian moose (*Alces alces*), 22 out of the 23 loci examined were found to be monomorphic, indicating a bottleneck event that may have lasted for 5 - 10 generations (Ryman *et al* 1977). Allendorf *et al* (1979) failed to uncover any variation at thirteen loci examined from polar bears (*Thalarcticus maritimus*), but made no comment on the historical occurrence of a bottleneck event. O'Brien *et al* (1983) found a total absence of genetic polymorphism at 47 allozyme loci in south African cheetahs (*Acinonxy jubatus*), this coupled with severe juvenile mortality (circa 70%), low sperm counts, a high incidence of sperm abnormality and the ability to tolerate reciprocal skin grafts between unrelated individuals, would be seen to indicate a severe population bottleneck followed by inbreeding (O'Brien *et al* 1985).

Using genetic fingerprinting, Gilbert *et al* (1990) found a high level of homozygosity within island populations of the California Channel Island fox (*Urocyon littoralis*). This low level to complete absence of minisatellite variability was attributed to founder effect and isolation coupled with, in one case, predation.

However, monomorphism at many loci need not be confined to natural populations that have suffered a bottleneck event. Heterozygosity in the Yellowstone Park elk (*Cervus canadensis*) was found to be not significantly different from that of elephant seals, Scandinavian moose or polar bears (Cameron & Vyse 1978). Yet the elk has probably not experienced a serious bottleneck in the recent past. Houston (1974) has estimated that herd size was in the thousands at the time when other game animals were being decimated across western North America. Cameron and Vyse (1978) questioned whether the low levels of genetic variation found in the elk were due to an adaptive genetic strategy or stochastic interaction between neutral mutations and genetic drift.

Nei *et al* (1975) have shown that bottlenecks can produce surprisingly little short term reduction in heterozygosity. For example, a sample of two individuals drawn from a randomly mating population should contain 75% of the heterozygosity found in the parent population and a sample of ten should contain 95% of the parent population variation. Nei's argument can lead to the conclusion that present day low levels of genetic variation found in the elephant seal, elk, moose and polar bear may not have been generated by bottleneck events but by population subdivision or inbreeding of long duration (Cameron & Vyse 1978).

The theory of neutral genetic variation (Kimura & Ohta 1978) predicts that heterozygosity will be a function of the mutation rate per cistron per generation and the effective population size. If the mutation rate is assumed to be comparable between homologous sets of loci for different species, the differences in the heterozygosity among species will reflect current and historical differences in population size and structure. Severe reduction in population number can drastically reduce heterozygosity, the subsequent recovery of which depends on the bottleneck size and the rate of population increase (Haigh & Maynard-Smith 1972, Nei *et al* 1975, Nei & Li 1976).

A population which has been reduced in number by some external influence such as human interference is unlikely to recover until the source of interference has been removed. The U. K. population of Reeves' muntjac has been fortunate in not suffering a debilitating bottleneck event and the concomitant reduced genetic variation. Instead, a small group of founders (from unknown backgrounds) have expanded rapidly due to the availability of suitable habitat, lack of competition and lack of predation. Similar successful introduction into the U. K. are those of the Chinese water deer, *Hydropotes inermis* (Putman 1988), and sika deer, *Cervus nippon* (Ratcliffe 1991).

### 1·11 Inbreeding

Inbreeding has been used to define any mating system from 'mating between related individuals' (Hartl & Clark 1989) to mating between relatives 'that share a greater common ancestry than if they had been drawn at random from an entire species' (Ralls *et al* 1986). The latter definition seems to be rather general since most populations could then be described as inbreeding. Clearly, the term 'inbreeding' has to be used to define a continuum of mating strategies which will then need quantification, which in turn will only be relevant to the species under study.

Here I use the term 'inbreeding' to refer to matings between closely related individuals, the primary effect of which is to cause allele frequencies to depart from Hardy-Weinberg expectations in the direction of homozygous excess. The population allele frequencies are not necessarily changed (in the absence of any selection pressure), only the relationships of alleles within a genotype are altered. If natural selection is operating, as is usually the case, then the resultant inbreeding depression can have profound effects on the viability of the population in which it is occurring, as well as on the course of evolution. Falconer (1981) has estimated that there is often a five to ten percent reduction in fitness (as a percentage of the non-inbred fitness) for each 10% increase in the inbreeding coefficient. This value is an average taken from comparisons of inbreeding coefficients and fitness reductions in a variety of plants and animals.

Although few data are available from natural populations, studies of captive populations and laboratory animals constantly confirm the magnitude of the effects of inbreeding (Ralls & Ballou 1983). Packer (1979) has listed three reasons why inbreeding depression should be expected to occur. First *'increasing homozygosity increases the chances of a detrimental recessive being expressed.'* Second, *'the heterozygote may be sometimes be fitter than either homozygote'*. Third, *'an increase in homozygosity decreases the variability between offspring, with the effect that the chance of any one of an individual's progeny surviving a sudden environmental change may be reduced'* (Williams 1966).

The occurrence of deleterious recessive mutations is a natural consequence of the evolutionary processes and can give rise to the genotypic (and hence phenotypic) variation necessary for adaptation. Because these rare recessive alleles can remain unexpressed in heterozygous individuals, inbreeding can have the effect of exposing these deleterious alleles to selection and thus raising the probability that they will be removed from the gene pool, or at least their incidence will be reduced.

Deleterious recessive mutations are rarely found at high frequencies in natural populations but, as a result of inbreeding and the consequent reduction in heterozygosity, may become common. The reduction in heterozygosity following inbreeding is accompanied by many congenital abnormalities in man (Cavalli-Sforza & Bodmer 1971, Seemanova 1971, May 1979 [but see Bittles (1979) for a criticism of these works]). Increased juvenile mortality and a reduction in fecundity in laboratory animal strains (Falconer 1981) and in captive populations (Ralls *et al* 1979, Ralls & Ballou, 1982a, b, Ballou & Ralls 1982, O'Brien *et al* 1985) have been reported. A high level of heterozygosity is positively associated with fecundity (Smith *et al* 1975), body size (Koehn *et al* 1973, Garten 1976, Cothran 1983), growth rate (Singh & Zouros 1978, Zouros *et al* 1980, Pierce & Mitton 1982, Ledig *et al* 1983, Cothran *et al* 1983) and in some cases, social dominance (Baker & Fox 1978). Despite reports of the levels of inbreeding depression in humans, work on a large sample of the Hindu, Christian and Muslim populations of Southern India has failed to detect any deleterious consequences of uncle/niece and first cousin marriages (Devi & Rao 1981). Within these religious groups 28·8%, 14·1% and 21·7% respectively of marriages are consanguineous (Ralls *et al* 1986).

Domestic animals can probably tolerate inbreeding to a greater extent than feral animals or recently captive populations because deleterious alleles will have been removed by mortality and selective breeding. However, during the early phases of domestication or selective breeding high juvenile mortality and reduced fertility can be major problems. A well known example of this is the short-horn breed of cattle, the history of which is known back to the period when the breed was founded (Wallace 1923). Of practical relevance is the pedigree of the 'Duchess' strain of which, according to Wallace (1923), 24 of the 58 females bred up to the time of the breeder's death (Thomas Bates) had failed to conceive, and the remaining 34 only produced 110 calves during their lifetimes. This low fecundity contributed greatly to the line's extinction (Wright 1923).

It has been suggested by Baker and Marler (1980) that interpretation of the results of studies on the effects of inbreeding should consider the origin of the founding population since many vertebrates may naturally inbreed to a limited extent. Such species that tend to inbreed naturally in the wild would show less deleterious effects when subjected to inbreeding in captivity or in small isolated populations. Very little is known about the breeding structure of Reeves' muntjac (re: section 1·3) but they are thought to be polygynous (Chapman *et al* 1993). Bengtsson (1978) and Smith (1979) have suggested that father/daughter matings should occur in polygynous species if the advantages of having a greater proportion of one's genes represented in the offspring are greater than the costs due to inbreeding. Those offspring that survive any deleterious effects of inbreeding (commonly reduced fecundity and juvenile mortality) will, by their very existence be 'fitter' and more likely to pass their genetic survival traits to future generations.

Ralls *et al* (1979) and Ballou and Ralls (1982) have presented the only investigations to date on the effects of inbreeding in a captive group of Reeve's muntjac. In summary they found (from zoo and collection records) that the juvenile mortality rate (JMR) in inbred offspring (33·3%) was almost double that of the outbred offspring (18·2%) (n = 40). Due to a lack of data, JMRs between outbred and inbred does could not be completed, but it was found that the JMR for outbred does producing inbred offspring was greater (40%) than for those producing outbred offspring (10%). In their study Ballou and Ralls (1982) considered factors that could confound the results, for example, birth season, birth order, management changes, population density, differences between captive-born and wild-caught, between inbred and non-inbred does and between primiparous does and multiparous does. None of these factors were found to account for the higher mortality rates of inbred offspring.

## 1·12 Previous Work on Reeves' Muntjac

The breeding habitat of muntjac deer in England is not uniformly distributed over space but occurs in discrete patches bounded by barriers that are generally of human origin (Chapman 1991a). These barriers - which are inhibitory rather than insurmountable - are formed by regions of human settlement, agriculture and associated infrastructure. The muntjac population is therefore considered as being divided into numerous sub-populations that are interconnected to varying degrees as a result of dispersal among habitat patches.

The distribution of *Muntiacus reevesi* may be considered as a metapopulation (a group of conspecific populations) consisting of a densely populated core in the Bedfordshire region and many small populations of varying density covering outlying regions in much of the south of England. Such a system may have the effect of enhancing long-term survival as patch populations may accumulate more genetic variance than would exist in a homogenous population of the same size. As a matter of random chance, sub-populations will tend to towards fixation for different alleles through genetic drift, founder effect and possibly because of different selection pressures. This of course will be balanced by the aggregation of deleterious alleles that could accumulate by the same mechanisms. The assessment of genetic differences between muntjac sub-populations is an important aspect of an investigation into the successful expansion of this species in the U. K.

Much work has been done on common indigenous U. K. species such as roe deer, *Capreolus capreolus* (Chaplin 1977b, Loudon 1987, Wermer 1987) and red deer, *Cervus elaphus* (Wermer 1987, Pemberton *et al* 1992) but relatively little has been done on the ecology and demography of Reeves' muntjac. Early reports on habitat choice and social organisation (Dansie 1970, Chaplin 1977b) were not always conclusive and only recently has karyological and morphometric work (Chapman & Chapman 1982, Chapman *et al* 1983) established the taxonomic status of the U. K. population.

Recent work, particularly on the King's forest population, has provided information on muntjac resource utilisation (Chapman *et al* 1985), social organisation and activity (Chapman *et al* 1993), ranging behaviour (Chapman *et al* 1993) and reproductive strategies (Chapman *et al* 1984, Chapman & Harris 1991, Chapman 1991b). The distribution and history of Reeves' muntjac has been closely documented by the Bedfordshire Natural History Association due to their interest in the animal, which is their emblem (Anderson & Cham 1987).

Many cytogenetic and genetic studies have been undertaken which have involved Reeves' muntjac in contributions to understanding chromosome evolution and gross structure in muntjac. The chromosome set of the Indian muntjac, *Muntiacus muntjak vaginalis,* (female 2n = 6, male 2n = 7) may have evolved from small acrocentric chromosomes such as those found in the complement of the Chinese muntjac, *Muntiacus reevesi* (2n = 46) by a series of tandem fusions and other rearrangements (Lee *et al* 1993). The origins of the unusually low chromosome number, for a mammal, found in the Indian muntjac has been the subject of much research (Schmidtke *et al* 1981, Johnston *et al* 1981, Lima-de-Faria *et al* 1984, Yu *et al* 1986, Scherthan 1990, Lee *et al* 1993 [plus many others]). The Indian muntjac and the Chinese muntjac are a species pair that show striking morphological similarity but extreme karyological differences. It has been suggested (Lee *et al* 1993) that the Indian species may have arisen from Chinese muntjac-like ancestors.

To date only one investigation has been made into muntjac population structure using molecular genetic techniques (Riegstra 1993). Other studies have been made at the taxonomic level (Wurster & Benirschike 1967, Chapman & Chapman 1982) and at the level of protein electrophoresis. Maughan and Williams (1967) detected differences in the electrophoretic mobility and separation of haemoglobin between Reeves' muntjac and the Chinese muntjac, and suggested that this difference could be taxonomically diagnostic. The only other investigation into the genetic variability of Reeves muntjac used inbreeding coefficients based on pedigree records (Ballou & Ralls 1982).

### 1·13 Organisation of the Thesis

The main body of this thesis is arranged in three main sections, one for each molecular genetic technique. These three main chapters (Chapters 3, 4 and 5) are each constructed as complete entities, that is, each chapter begins with a literature review which is followed by a methods section, then results, and finally a discussion. These chapters are preceded by a general introduction and proceeded by a general discussion. I decided on this format because I feel that the differences between the techniques and interpretation of results warrant separate consideration. Chapter 6, the general discussion, brings together the three separate sets of results.

Since the completion of the practical aspects and most of the written part of this thesis, new information has come to light regarding the origins of some of the U. K. sub-populations. This work has been published by Chapman *et al* (1994) and will be briefly discussed in relation to this thesis in Chapter 6.

Chapter 2

# SITE SELECTION, SAMPLE COLLECTION AND PREPARATION

## 2·1 Selection of Sampling Sites

The sample collection was organised by S. Harris and N. Chapman. The sites were chosen in consultation with S. Harris on two criteria, (1) the anticipated availability of samples and (2) the location with regard to Woburn Abbey, Bedfordshire - the putative origin of the muntjac population in England. The minimum number of samples required from each site was set at 18, mainly because this is the maximum number of individuals that can be run on one fingerprint gel, along with two flanking molecular weight markers. As the samples were coming primarily from animals being killed for sport, the number collected was, however, somewhat variable.

Eighteen sampling sites, arranged in three concentric circles around Woburn were chosen (Figure 2·1) to reflect successively established populations as determined by historical records of sightings (Lever, 1977).

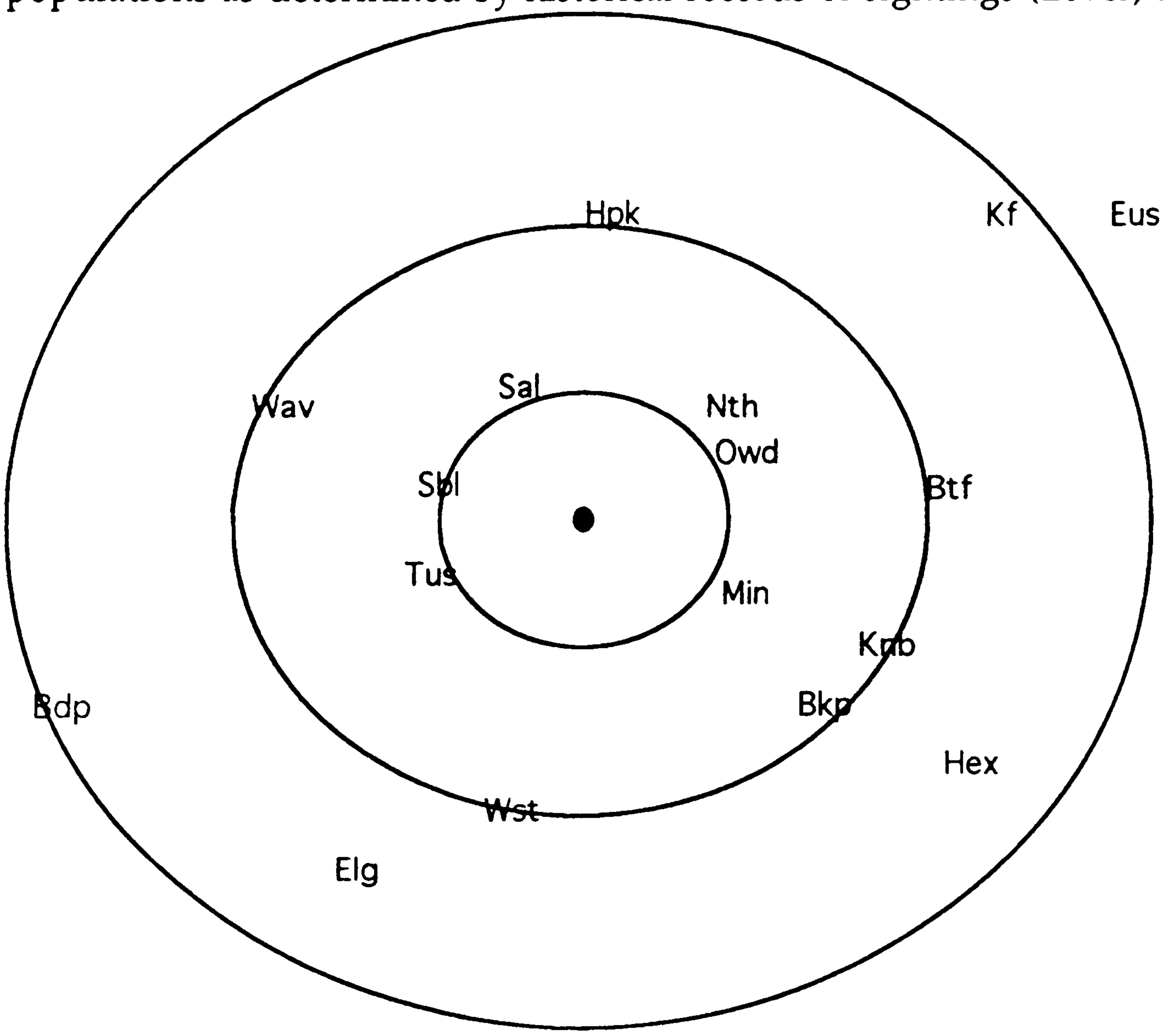


Figure 2·1 The arrangement of sampling sites centred on Woburn Abbey

It is anticipated that if the muntjac had radiated from Woburn, then a pattern to the loss of variability, with a loss of alleles which were rarest in the founding population followed by a reduction in heterozygosity, would be observed. The loss of alleles should follow a process of successive random subsampling of the allelic variability in the preceding (source) population as predicted by a 'stepping stone' colonisation model involving a process of sequential founding events (Maynard Smith 1989; Hanski 1991).

We were also able to obtain 31 samples from Taiwan (one putative source of the U.K. population), although six of these were too degraded to be of use. Figure 2·2 is a general map of the sampling sites. Table 2·1 is a list of the sample sites and the number of individuals taken. A complete list may be found in Appendix A.

| Population | | Grid Ref. | Number of samples | | | Total |
|---|---|---|---|---|---|---|
| | | | ♂ | ♀ | ? | |
| WAVERLY | (Wav) | SP 3570 | 3 | 2 | | 5 |
| TUSMORE | (Tus) | SP 5630 | 10 | 8 | | 18 |
| SOUTHBLOCK | (Sbl) | SP 6542 | 16 | 9 | | 25 |
| SALCEY FOREST | (Sal) | SP 7918 | 8 | 4 | 2 | 14 |
| WOBURN ABBEY | (Wob) | SP 9632 | 11 | 8 | 5 | 24 |
| HARRY'S PARK | (Hpk) | SP 9487 | 5 | 2 | | 7 |
| BOWDEN PARK | (Bdp) | ST 9368 | 8 | 2 | 3 | 13 |
| ELING COMMON | (Elg) | SU 5274 | 12 | 12 | 6 | 30 |
| WORMSLEY ESTATE | (Wst) | SU 7394 | 11 | 3 | 1 | 14 |
| OLDWARDEN | (Owd) | TL 1343 | 11 | 9 | | 20 |
| NORTHILL | (Nth) | TL 1446 | 13 | 12 | | 25 |
| MINSBURY HILL | (Min) | TL 1527 | 10 | 7 | | 17 |
| BROCKET PARK | (Bkp) | TL 2113 | 7 | 7 | 8 | 22 |
| KNEBWORTH | (Knb) | TL 2320 | 14 | 12 | 2 | 28 |
| HEXTON | (Hex) | TL 3010 | 15 | 12 | 2 | 29 |
| BUNTINGFORD | (Btf) | TL 3629 | 11 | 6 | 1 | 18 |
| KING'S FOREST | (Kf) | TL 8-7- | 20 | 12 | 3 | 35 |
| EUSTON | (Eus) | TL 9-7- | 18 | 7 | | 25 |
| TAIWAN | (Twn) | - | 12 | 11 | 2 | 25 |

Table 2·1 Sample Sites and Numbers of Individuals Collected

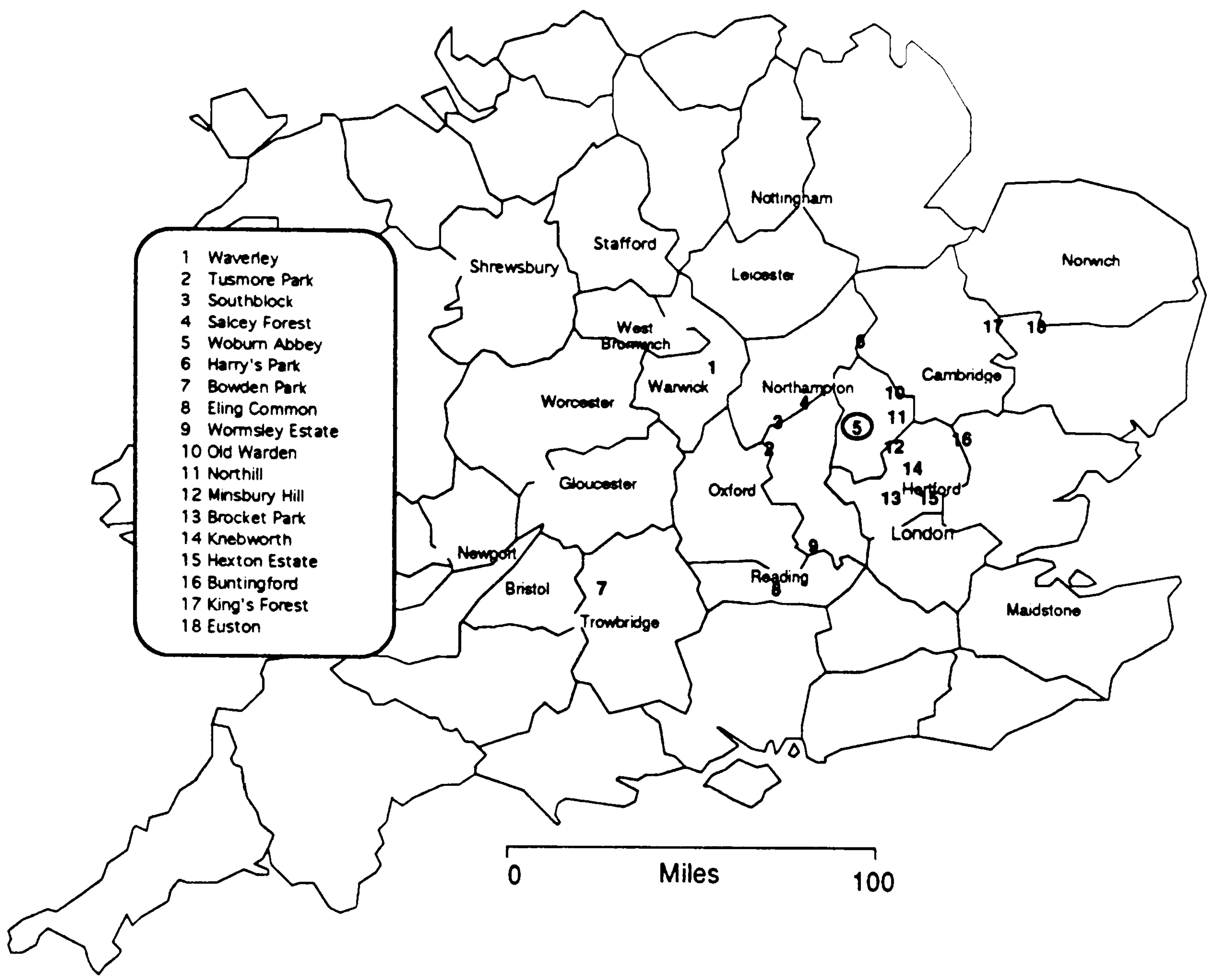


Figure 2·2 Sampling Sites

## 2·2 Tissue Collection and Storage

Three types of tissue were available, blood, ear notches from ear tagging, and tongue. The blood was collected in 5ml vials containing ACD buffer (citric acid : sodium citrate : glucose), the ear notches were stored in small vials and the tongue was collected in small containers. All were delivered to Bristol University and there stored at -70°C. Upon arrival at the Institute of Zoology the blood and ear notches were placed in storage at -70°C whilst the tissue was placed in the -20°C freezer prior to cutting into appropriate sized pieces, placing in screw capped micro centrifuge tubes and storing at -70°C. The tissues samples (tongue) were collected by hunters either during the course of their recreation or as part of a culling program. The majority of samples collected during the first part of this investigation were found to yield DNA that was too degraded to be of adequate quality for fingerprinting. This was probably due to the time interval between killing the animal, removing a portion of the tongue and posting it to the Department of Zoology at the University of Bristol, where it was placed in cold storage. The tissue samples were then packed in dry ice and sent to the Institute of Zoology in London. During the second stage of the investigation, 20ml scintillation counter vials (Gallenkamp) containing 10ml of DMSO solution (20% Dimethyl Sulphoxide, 0·25M EDTA, saturated with Sodium Chloride; pH8·0.) were sent to the collectors with instructions to place the cut up tissue samples in them as soon as the animal had been killed. The samples collected in this way were put into cold storage (-20°C) upon arrival at the Institute of Zoology and provided DNA of sufficient quality to allow fingerprinting.

## 2·3 DNA Extraction and Purification

### i) From Tongue Samples and Ear Notches.

The method followed is essentially that of Bruford *et al* (1992). Between 200mg and 300mg of tissue was taken for DNA extraction. Initially the tissue was macerated in liquid nitrogen but later this was found not to be necessary as adequate yields were obtained by mincing the tissue with a scalpel by cutting and shearing against the sharpened edge of a scissor blade. The macerated tissue was placed in an 11ml polythene extraction tube containing 4 ml of lysis buffer (100 mM NaCl, 50 mM Tris-HCl, 5 mM EDTA, p*H* 7·8). 200µl of 10% Sodium Lauryl Sulphate solution (SLS) was added to give a final concentration of 0·5%, followed by 20µl of 20mg/ml proteinase K (NBL). The extraction was incubated overnight at 56°C on an orbital shaker. Prior to purification a further 10µl of proteinase K solution was added and incubation allowed to continue for another hour. Total genomic DNA was purified by extraction once with an equal volume of phenol (USB; equilibrated to p*H* ≥8·0), twice with equal volumes of phenol : chloroform : isoamyl-alcohol (25:24:1) and once with an equal volume of chloroform : isoamyl-alcohol (24:1), followed by precipitation by the addition of 2·5 volumes of absolute ethanol. There followed two washes in 70% ethanol prior to drying in a rotary vacuum drier set on medium heat. The pellet of DNA was resuspended overnight at room temperature in 500µl of TE (10 mM Tris-HCl, 1mM EDTA; p*H* 7·5) prior to assay in a Beckman DU-65 U.V. spectrophotometer. DNA concentration was measured using ultraviolet absorbance spectrophotometry at 260 nm and 280 nm. At a wavelength of 260 nm an absorbance ($A_{260}$) of 1·0 corresponds to 50 µg of double-stranded DNA per ml (Brown 1990). $A_{260}$ and the ratio of $A_{260}/A_{280}$ was used to give an indication of purity. If this ratio was less than 1·7 the purification procedure was repeated. To assess the purity and quality (i.e. level of degradation) ca 500ng was run on a 1% TBE agarose gel (90mM Tris-borate, 90mM boric acid, 1mM EDTA, 1% agarose (Sigma type ll)) for 60 minutes at 5V/cm, stained with ethidium bromide (0·5µg/ml in 0·5x TBE) and viewed under long wave ultra violet light. 500ng of native λDNA (BRL, supplied at ca 500µg/ml) was also included in an adjacent lane as a standard. The DNA was stored at 4°C.

**ii) From Blood Samples.**

Five millilitres of blood/buffer solution was diluted with an equal volume of phosphate buffered saline (0·13 M NaCl, 2·7 mM KCl, 10 mM $Na_2HPO_4$, 1·4mM $KH_2PO_4$ p*H* 7·2) and centrifuged for 15 minutes at room temperature at 4000rpm in a Beckman GPR benchtop centrifuge with a swing-out rotor. The supernatant was discarded and the pellet resuspended in 2ml of extraction buffer (10 mM Tris-HCl; 0·1mM EDTA; 0·5% SLS, 10µl of ribonuclease A (RNase A) from 10mg/ml stock solution (NBL)). 200µg of Proteinase K was added and the extraction incubated overnight at 56°C on an orbital shaker. Purification was as above.

**iii) Rapid Method of DNA Extraction from Tissue and Blood.**

This method is based on that of McCabe (1990) for extraction of DNA from bacterial colonies. Approximately 1mm$^2$ of tissue was taken either from the -70°C store or from DMSO, added to 200µl of extraction buffer (20 mM Tris-HCl; 2 mM EDTA p*H*7·5-8·5; 1% Triton X-100) and incubated at 95°C for 10 minutes. The cell debris was spun down at in a bench-top centrifuge (13000rpm/10min/room temp.). Dilutions were made of 1:50 and 1:100 and 2µl of these dilutions plus 2µl of the undiluted stock were taken for use in the Polymerase Chain Reaction.

Chapter 3

# MULTILOCUS DNA FINGERPRINTING AS A TOOL FOR INVESTIGATING POPULATION DIFFERENTIATION AND GENETIC STRUCTURE IN FERAL MUNTJAC DEER.

## 3·1 Introduction

Until the late 1970s the majority of genetic investigations at the molecular level were limited to polymorphisms detected serologically or by gel electrophoresis. With the development of Southern blotting (Southern 1975) and DNA cloning it became feasible to investigate genetic variability at the level of genomic DNA by analysis of restriction fragment length polymorphisms (RFLP) (Jeffreys *et al* 1986[a]). The gain or loss of a single restriction endonuclease cleavage site produces dimorphisms whose information value is limited by their low heterozygosity. For any particular diallelic marker the maximum frequency of heterozygotes obtainable in a population (in the absence of selection) is 50% (Jeffreys *et al* 1986[a]). Nevertheless, these polymorphisms can be of considerable value, for example, variation in the human β-globin gene cluster provided useful codominant markers (Jeffreys 1979). The value of RFLPs as genetic markers has also been found in the construction of linkage maps.

Genetic analysis was simplified considerably with the fortuitous discovery by Wyman and White (1980) of a random human DNA segment which defined a multiallelic locus - the first demonstration that hypervariable regions existed in human DNA. Subsequently, and again by chance, several other hypervariable regions were discovered near the human insulin gene (Bell *et al* 1982), α-related globin genes (Higgs *et al* 1981, Proudfoot *et al* 1982, Goodbourne *et al* 1983) and the c-Ha-*ras*-1 oncogene (Capon *et al* 1983). In each case the variable region consisted of tandem repeats of a short sequence or 'minisatellite' and polymorphisms resulted from allelic differences in the number of repeats. Jeffreys *et al* (1979) described a short minisatellite consisting of four tandem repeats of a 33 base-pair sequence in an intron of the human myoglobin gene. In subsequent papers (Jeffreys *et al* 1985a, b) showed that this myoglobin 33bp repeat, and other similar repeat sequences, could be used to detect hypervariable minisatellite DNA in humans and to produce somatically stable individual specific DNA 'fingerprints'.

Since this technique of DNA fingerprinting was first established in 1985 (Jeffreys *et al* 1985a, b) it has become increasingly refined and widely used in many areas of scientific endeavour. Minisatellite DNA probes of human origin have been found to cross-hybridise to and reveal polymorphisms in a wide variety of eukaryotic species including domestic and feral mammals (Morton *et al* 1987 - dogs; Jeffreys & Morton 1987 - dogs and cats; Jeffreys *et al* 1987 - mice; Georges *et al* 1988 - cattle, horses, pigs, dogs; Laikre *et al* 1992 - wolves; Pemberton *et al* 1992 - deer), birds (Georges *et al* 1988 - chickens; Gibbs *et al* 1990 - red-winged blackbirds; Hillel *et al* 1989 - poultry; Brock & White 1991 - parrot; Burke and Bruford 1987 - several species, Hanotte *et al* 1992 - gallinaceous birds; Meng *et al* 1990 - swan), fish (Georges *et al* 1988 - cyprinid fish; Baker *et al* 1992 - orange roughy), some invertebrates such as corals and snails (Coffroth *al* 1992 - gorgonian coral, Jarne *et al* 1990, 1992 - *Bulinus*), rice (Dallas 1988), angiosperms (Rogstad *et al* 1988 - *Populus deltoides* (cottonwood) and *P. tremulodies* (aspen)) and yeast (Walmsley *et al* 1989).
This almost universal applicability of minisatellite probes has led to them being used in many scientific fields as diverse as; authenticating cell lines (Fey & Tobler 1991), forensic investigation (Gill *et al* 1987, Gill & Werrett 1987) conservation genetics (Wayne *et al* 1991a, Greth *et al* 1992,) and fungal taxonomy (Meyer *et al* 1991, 1992).

### 3·1a Organization of Minisatellite DNA

Minisatellite DNA consists of tandem arrays of short repeat units which are in the order of 16-64 base-pairs (Avise 1994). Minor sequence variations may exist among repeat units but, in terms of DNA fingerprinting, the most important variation is that of the number of repeat units per tandem array. Jeffreys *et al* (1985a, b) assessed variation in tandem arrays composed of approximately 33 base pairs per repeat unit and found fragments ranging in size from 200 to over 20,000 base pairs. Generally though, fragments below 2kb are not scored because of the difficulty in resolving individual bands, and fragments above 20kb are uncommon. The conventional nomenclature of tandem repeats is related to their overall size: satellites (hundreds of kilobases), midisatellites (tens of kilobases), minisatellites (a few hundred bases to tens of bases) and microsatellites (generally less than five bases). The term 'DNA fingerprinting' has become a generic term used to describe many DNA based genetic profiling techniques. Here, the term will be used only in conjunction with the genetic profiles generated from a class of moderately repetitive DNA sequences known as 'minisatellite' DNA (Jeffreys *et al* 1985a, b).

Various models have been proposed for the origins and functions of tandem repetitive DNA. Smith (1976) proposed that tandem repetitive DNA, whose sequence was not maintained by natural selection, would arise as a result of unequal crossing over if strand exchange in the recombinational process was assumed to depend on sequence similarity. He justified this proposal by a simulation of unequal cross-overs for two different rates of crossing over that were high relative to the mutation rate. Recently though, tandemly repeated non-coding DNA sequences have been localised to regions of the eukaryotic genome that are associated with different levels of recombination (Stephan 1989). Jeffreys *et al* (1985a) presented a speculative model for the origins and function of minisatellite DNA based on the similarity between the myoglobin 33bp repeat sequence and the chi sequence, a signal for generalised recombination in *E. coli.* The core region in the minisatellites they cloned was seen by Jeffreys *et al* to suggest that the sequence may help to generate minisatellites by either promoting the initial tandem duplication of unique sequence DNA, or by initiating unequal exchange and then amplifying the duplication into a minisatellite. The model predicted that the core or core-like sequence may be a 'hotspot' for human chromosome recombination (Jeffreys *et al* 1985b).

A model of sequence dependent unequal crossing over and gene amplification has been simulated by Stephan (1989) to account for some of the structural features of tandemly repeated DNA sequences. He demonstrated that regions of DNA (the structure of which is not under selective constraints) will show repetitive patterns over a wide range of recombination rates as a result of the interaction of unequal crossing over and slippage replication, both processes that depend on sequence similarity.

To date, the consensus of opinion seems to indicate that the increase and decrease in the length of a particular array of tandem repeats arises from high rates of unequal cross-over during meiosis. The speculation that minisatellite DNA sequence provides 'hotspots' for recombination continues (Jarman & Wells 1989).

Many pedigree analyses of humans (Jeffreys *et al* 1985c, 1986a) and other animals (Gibbs *et al* 1990, Hillel *et al* 1990, Westneat 1990) have shown that DNA fragments detected by minisatellite probes can be reliably used for studying the segregation of multiple heterozygous DNA fragments, even in a situation where one or the other parent is unavailable. Alleles at these extremely polymorphic loci are inherited as codominant Mendelian traits (Wong *et al* 1986) and such loci can be detected simultaneously by minisatellite probes. For example, using the Jeffreys' probes 33·15 and 33·6 upto 34 autosomal hypervariable loci are detected in humans (Jeffreys *et al* 1986a), between 20 and 32 loci can be detected in various avian species (Burke and Bruford 1987, Westneat 1990) upto 40 loci (>2kb) in cetaceans (Amos & Hoelzel 1990) and upto 35 in domestic dogs and 21 in domestic cats (Jeffreys & Morton 1987). Even in a system modified to resolve bands greater than 4kb (in a captive population of Guinea baboons) a mean of 26 bands per individual was scored (Bruford & Altmann 1993). The number of bands scored depends on several factors, the two most important of which are (1) the restriction enzyme/probe combination used (some restriction enzymes produce more bands than others, see Georges *et al* 1988; Bruford & Altmann, 1993), and (2) the stringency of hybridisation and subsequent washes.

The Jeffreys' probes 33·15 and 33·6, although commonly used are not the only minisatellite probes available. The wild type of bacteriophage M13 has been found to contain a sequence that detects hypervariable minisatellites in animals (vertebrates and invertebrates), plant and micro-organisms (Vassart *et al* 1987, Ryskov *et al* 1988, Ryder *et al* 1989). There have also been many reports of DNA fingerprinting successfully accomplished with species specific probes (Georges *et al* 1988, Gyllensten *et al* 1990) and synthesised oligonucleotide probes of various design (e.g. Buitkamp *et al* 1991a 1991b).

**3·1b The Main Applications of DNA Fingerprinting Relative to this Thesis**

DNA fingerprinting has become an important diagnostic technique in two areas of scientific investigation, those that require individual identification, (be it a plant, animal or cell line) and those where individual identification is secondary to a measure of group relatedness. It has taken a place of unique importance in the work of field biologists and behavioural ecologists not only because of the ability of this technique to resolve previously unanswerable questions about paternity, kinship, reproductive success and breeding strategies (for review see Burke 1989), but also because it is sometimes possible to resolve relatedness amongst populations with a greater accuracy than was previously attainable. Small, or isolated populations such as those in captivity or on islands often exhibit low levels of genetic variability. The large number of polymorphisms generally uncovered by DNA fingerprinting are therefore of great value in estimating the genetic structure of these populations. Pedigree analysis has revealed that the bands in DNA fingerprint profiles segregate in the normal Mendelian fashion (Wong *et al* 1986). This simple Mendelian nature of inheritance allows the identification of first order relatives (Jeffreys *et al* 1985c, Odelberg *et al* 1988, Yokoi 1990), an assessment of relatedness within family groups (Young 1990), and an assessment of relatedness both within and between isolated populations (Lehman *et al* 1992, Triggs *et al* 1992). This degree of certainty about relationships has opened up a large area of research into mating strategies, paternity and reproductive success (e.g. Gibbs *et al* 1990, Amos 1990, Pemberton *et al* 1992, Amos *et al* 1993a, b)

### 3·1c Paternity Studies, Reproductive Success and Breeding Strategies

Ely and Ferrell (1990) assessed paternity in a group of captive chimpanzees (*Pan troglodytes*). Even though band sharing was high among unrelated individuals (0·48), in all 21 cases involving multiple potential sires, the true sire was identified and other potential sires were excluded, with high probability. In a similar study of Guinea baboons (*Papio hamadryas*), Bruford & Altmann (1993) used multilocus DNA fingerprinting to determine reproductive success among seven active males from a captive population of 65 individuals. However the high levels of band sharing found in this isolated population - means of 0·84 for bands greater than 2 kb to 0·61 for bands greater than 4 kb - hindered attempts to elucidate paternity in all except (tentatively) four of the 33 offspring analysed.

The determination of the degree of genetic relatedness among individuals is important for many behavioural studies (Hamilton 1964). This is of special relevance in avian species where intraspecific brood parasitism (ISBP) and extra-pair fertilisations (EPF, or extra-pair copulation - EPC) occur. Quinn *et al* (1987), in study of the lesser snow goose (*Anser caerulescens caerulescens*) used a DNA marker to confirm field observations that suggested this species engaged in ISBP and EPF. Their study was one of the first to provide evidence for mixed maternity and paternity in a socially monogamous species. Westneat (1990) showed that 35% of offspring in indigo bunting (*Passerina cyanea*) families resulted from extra-pair fertilisations. The application of DNA fingerprinting in this study corroborated previous data from protein electrophoresis and wing-length heritabilities that indicated that 27%-42% of bunting offspring came from extra-pair matings.

Polygamy in the form of extra-pair fertilisation has also been discovered - by means of DNA fingerprinting - in many wild bird species such as the dunnock, (*Prunella modularis*, Burke *et al* 1989), blue tit (*Parus caeruleus*, Kempenaers *et al* 1992), red-winged blackbird (*Agelaius phoeniceus*, Gibbs *et al* 1990), and the house sparrow (*Passer domesticus*, Wetton *et al* 1987). The traditional measure of reproductive success is the number of young successfully raised, and as individuals are generally expected to maximise their reproductive success, these mating strategies of extra-pair fertilisation as well as intraspecific brood parasitism will have a major effect on the costs and benefits of parental care and on the intensity of sexual selection.

### 3·1d Kinship, Group Behaviour and Social Structure

Kinship and associated behaviours are areas of research that were previously the domain of behavioural observations. The unravelling of complex social groups that show extensive co-operative behaviour has only recently become possible due to DNA fingerprinting. One of the best examples of work of this nature is that of Packer *et al* (1991a) and Gilbert *et al* (1991). African lions are known to live in complex social groups and show extensive co-operative behaviour (Schaller 1975). Definitive parentage was assessed in eleven lion prides, permitting the precise genealogy of about 200 lions. Prior to this study it was believed that each male within a pride sired an equal proportion of offspring in the pride. But from this work it is now clear that this is only true for small groups of males within a pride. For larger groups, reproductive success is guaranteed for only a few males at the expense of other closely related males. Thus, in groups formed from close male relatives, some males act as non-reproductive helpers.

The dwarf mongoose (*Helogale parvula*), is a social, communally breeding carnivore in which the oldest male and female dominate reproduction while subordinates tolerate reproductive suppression and provide care for the offspring of the oldest pair (Creel & Waser 1991). By using Jeffreys' human multilocus probes (33·15 & 33·6) Keane *et al* (1994) were able to unravel a complex social structure and discover that, contrary to previous knowledge, 24% of young had subordinate fathers and 15% had subordinate mothers. Furthermore, multiple paternity was found within a single female's litter. The subordinates who reproduced were of 'high' social rank and tended to be distantly related to the same-sex dominant. Subordinate reproduction on this scale may be a means of 'power-sharing', in which dominant males and females concede some direct reproductive success to high-ranking subordinates in order to retain them in the pack as helpers and thereby increase their own reproductive success (Keane *et al* 1994).

Similarly, a study of male mating success and paternity in the grey seal, *Halichoerus grypus*, (Amos *et al* 1993b) found that the reproductive success of the dominant males was not as great as their behavioural domination of copulatory opportunities would suggest (judged from behavioural observations). In 36% of cases investigated the most likely candidate male was shown not to be the father.

### 3·1e The Ability of DNA Fingerprinting to Differentiate Between Populations

Recently, DNA fingerprinting has been used to determine genetic variation at the level of populations, essentially in three aspects of population demography, (1) the current dispersal of populations and species, (2) their migratory history and (3) their current social structure.

In association with mtDNA analysis, DNA fingerprinting was used in studies of three clusters of gray wolf packs to investigate social behavioural observations (Lehman *et al* 1992). Initially, dispersal was thought to be common among local groupings of packs and not sex-biased. A high level of similarity was found both within packs, and between packs that had overlapping ranges, thus confirming the first observation; but female short range dispersal was found to be more common than male. Band sharing coefficients calculated from unrelated individuals allowed the classification of individuals as either unrelated, moderately related or closely related. This classification, when applied to the packs and in addition to mtDNA analysis, demonstrated that the majority of wolf packs consisted of closely related individuals except for the breeding pair. However, some packs were found to be composed of unrelated non-breeding members.

In studies of other mammals DNA fingerprinting has shown that gene dispersal occurs between adjacent groups, for example, between pods of the Faroese pilot whale, *Globicephala melas* (Amos *et al* 1991). Investigations of inter- and intrapopulation variability in river populations of the blue duck (*Hymenolaimus malacorhynchos*) found a decrease in genetic similarity correlated with increased geographic distance and significantly higher genetic similarity within sub-populations. The high degree of similarity within sub-populations indicated that dispersal is very limited in this species and inbreeding common. Apparently these are natural characteristics of the blue duck social system (Triggs *et al* 1992).

Migrations on an historical scale have been examined by DNA fingerprinting analysis of isolated populations, such as the lions of the Ngorongoro crater in Tanzania (Packer *et al* 1991b, Gilbert *et al* 1991). Interpopulation genetic variability values suggest that the original Ngorongoro lions were founded from the adjacent Serengeti populations. The similarity, in terms of DNA fingerprinting, was found to be greater between Serengeti and Ngorongoro lions than between Ngorongoro lions and the other putative founder population - the Manyara lions.

From the population and sub-population point of view, levels of founder effect, inbreeding and genetic drift can sometime be of more consequence than familial relationships. The genetic differences defined by genetic fingerprinting can give an overview of the degree of gene flow between populations and an indication of the way in which a population, or sub-population first arose.

### 3·1f The Relevance of DNA Fingerprinting to Conservation Genetics

Ballou and Ralls (1982) have emphasised the need for monitoring genetic variability in small or isolated, captive and wild populations to assess inbreeding, effective population number and thus define the goals for management action. Knowledge of the genealogy among captive groups of animals - by means of DNA fingerprinting and other molecular techniques - can profitably be used in the genetic management of captive breeding colonies. By monitoring and (if possible) adjusting an individual's genetic contribution to succeeding generations a level of genetic diversity can be maintained. Managed breeding programs will thus be able to offset the negative aspects of inbreeding such as increased infant and juvenile mortality, reduced tolerance to infectious diseases, delayed sexual maturity and reduced fertility (Ralls *et al* 1986). Effective captive breeding programs therefore require the identification and attainment rate of successful males. DNA fingerprinting has been used in this respect by Ely and Ferrell (1990) and Ely *et al* (1991) to resolve ambiguous paternity in a captive population of chimpanzees (*Pan troglodytes*). The effects of social dominance on male reproductive success, variance in male and female reproduction, the rate of decline in genetic variability and an estimate of effective population size were assessed.

DNA fingerprinting has also been used in the captive breeding recovery program of the Puerto Rican Parrot (*Amazona vittata*) with the purpose of identifying parent pairs exhibiting low levels of band-sharing and thus a high potential for reproductive success (Brock & White 1992).

In small and/or isolated feral populations reduced levels of genetic variability many result from one or more of several circumstances. For example, the low genetic variability of the California Channel Island fox population (*Urocyon littoralis*) on the island of San Nicolas has been tentatively attributed to a bottleneck event created by predator pressure during the early 1970s (Gilbert *et al* 1990). Similarly, levels of heterozygosity and productivity in the lions (*Panthera leo*) of the Ngorongoro Crater region of Tanzania have declined since the mid 1960s due to inbreeding (Packer *et al* 1991a, Gilbert *et al* 1991). The current population are all descended from 15 founders that survived an epizoötic of biting flies (*Stomoxys calcitrans*) in 1962 (Fosbrooke 1963). Compared to the bordering, but isolated Serengeti population, the crater lions have significantly less allozyme heterozygosity, less restriction-fragment length polymorphism in major histocompatibility complex class I genes and more sperm abnormalities (O'Brien *et al* 1987, Wildt *et al* 1987, Yuhki & O'Brien 1990, Packer *et al* 1991b).

Combatting isolation and inbreeding are of prime concern for the successful development of small populations. These two factors result in the depressed level of genetic variability which is the precursor to population extinction as a result of stochastic incidents such as disease or famine.

The Florida panther (*Felis concolor coryi*) has, over the past 100 years or so, suffered severe range and demographic diminution, with a consequent severe cost in fitness (Roelke *et al* 1993). DNA fingerprinting of the Florida panther revealed a very low level of minisatellite variation, in the order of 85% less than that of the closely related western U.S. subspecies (Roelke *et al* 1993).

The application of DNA fingerprinting has also been used to determine whether isolated populations are sufficiently genetically distinct to warrant their continued protection. For example, the wolves on Isle Royale in Lake Superior, where a single gray wolf pair is thought to have founded the present population, were found to have limited genetic distinctiveness such that its conservation as a separate genetic entity was not warranted. Such distinctions are needed in order to make decisions concerning the introduction of mainland wolves to reinforce the isolated population that is in danger of becoming extinct (Wayne *et al* 1991a).

### 3·1g Applications to this Study

The two minisatellite probes used here are those known as 33·15 and 33·6, which were originally isolated from a four tandem repeat 33bp sequence within an intron of the human myoglobin gene (Jeffreys *et al* 1985a). Probe 33·15 consists of a cloned human minisatellite comprising 29 repeats of a 16bp variant of the 'core' sequence and probe 33·6 is composed of 18 repeats of an 11 bp sequence. The sequences of the core and probes are (Jeffreys *et al* 1986b):

| | |
|---|---|
| core sequence | $\text{GGAGGTGGGCAGCA}\,{}^{\text{A}}_{\text{G}}\,\text{A}$ |
| 33·15 | $(\text{AGAGGTGGGCAGGTGG})_{29}$ |
| 33.6 | $[(\text{AGTGGCTGGAGG})_{3}]_{18}$ |

In this chapter I have used the technique of DNA fingerprinting to investigate the genetic relationships within and between four muntjac populations. My approach (in terms of DNA fingerprinting) has assumed that the relative distribution of similarity values in related and unrelated individuals will be similar in different sub-populations. This clearly may not be the case in sub-populations that are small and isolated or have suffered a recent population bottleneck or are exhibiting the consequences of a founder event (Gilbert *et al* 1990, Packer *et al* 1991b).

An estimator of gene flow between populations ($F'_{ST}$) will be obtained through an analysis of between population genetic variance (Lynch 1991). In both the experimental design and the analysis of data I have attempted to control for the theoretical and practical limits of DNA fingerprinting technology (Burke 1989; Lander 1989; Lewin 1989), as well as applying the appropriate mathematical corrections to the interpretation of the data (Lynch 1988, 1990, 1991).

## 3·2 MATERIALS AND METHODS

### 3·2a Restriction Enzyme Digestion

DNA was extracted from the samples as previously described. Twenty micrograms of high molecular weight total genomic DNA (as determined by agarose gel electrophoresis) was digested to completion with 60 units of the restriction endonuclease *Hin* fI, according to the manufacturer's recommendations. The incubation was allowed to continue overnight.

An aliquot of about 500ng from the digestion was taken and run on a 14cm 1% agarose gel at 5v/cm for one hour with a molecular weight marker in an adjacent lane (λ/*Hin* dIII), followed by visualisation with ethidium bromide, to assess whether digestion was complete. If digestion was incomplete, 30 units of enzyme and an appropriate amount of buffer was added and incubation continued for another 6 - 8 hours. The level of digestion was again assessed. Completed digests were then purified with phenol : chloroform : isoamyl alcohol, chloroform : isoamyl alcohol, precipitated by the addition of 2·5 volumes of 100% ethanol, washed once in 70% ethanol, dried and re-suspended in 20μl of dd$H_2O$. One microlitre was taken from those samples to be run on a single fingerprint gel, mixed with an appropriate amount of gel loading buffer (15% Ficoll, 0·2% bromophenol blue, 0·2% xylene cyanol) and electrophoresed on a 14cm 1% agarose gel at 5v/cm for one hour to assess the relative concentration of each sample. Adjustments were made to the concentrations and further gels run until all the samples were seen to be of similar concentration. Five nanograms of a 1kb ladder, extending from 1kb to 12kb, was included in each sample.

### 3·2b Gel Electrophoresis

The samples were mixed with 5μl of 1 x loading buffer (15% Ficoll, 0·2% bromophenol blue in 1x TBE p*H* 7·8) and loaded into a 30cm 0·9% agarose gel (dissolved in 1 x TBE p*H* 7·8) and run in 1 x TBE (p*H* 7·8) at between 1V/cm and 2V/cm, at room temperature, until the 2021bp fragment of a molecular weight marker (λDNA digested according to the manufacturer's recommendations with the restriction enzyme *Hin* dlll), loaded into the two flanking lanes, had migrated 20cm. The low (<2kb) molecular weight DNA was run off the end of the gel because this region is known to contain an almost irresolvably complex mixture of shorter hybridising DNA fragments, which may be derived from short and relatively monomorphic minisatellites and/or small fragments from larger arrays (Jeffreys *et al* 1986).

### 3·2c Southern Blotting

Southern blotting was performed as per Bruford *et al* (1992), using a 20cm x 22cm membrane. The gel was prepared for blotting by (i) depurination (0·25M HCl) for 20 minutes - to break up the DNA fragments so that they can migrate with less impedance from the gel matrix, (ii) denaturation for 20 minutes (1·5M NaCl; 0·5M NaOH) - to separate the DNA strands, and (iii) neutralisation for 20 minutes (1·5M NaCl; 0·5M Tris-HCl p*H*7·5). Transfer buffer was 20x SSC (3M NaCl; 0·3M Na-citrate) and the membrane, 20 cm x 22cm Hybond-Nfp (Amersham). For convenience, the transfer was usually allowed to proceed overnight. After transfer the membrane was washed briefly in 2x SSC and the DNA was fixed by crosslinking to the membrane by exposure to long wavelength U.V. light (circa 312nm) for four minutes - as recommended by Amersham.

### 3·2d Probe Preparation and Labelling

Both Jeffreys' probes 33·6 and 33·15 were supplied (via M. Bruford) inserted into the plasmid PUC18. Stocks were produced by standard transformation, harvesting and purification procedures (Sambrook *et al* 1989). Probe fragments were excised from the host plasmids by digestion with *Eco* RI and *Hind* III and purified by agarose gel electrophoresis (Sambrook *et al* 1989). The concentration of the final product was assessed by spectrophotometry at $OD_{260}$ and adjusted to give a concentration of about 40ng/µl.

The probe labelling method followed was based on that of Feinberg and Vogelstein (1983). Forty nanogrammes of probe (33·15 *or* 33·6 *or* the molecular weight marker) was boiled in 16µl of dd$H_2O$ for five minutes and then plunged into ice for two minutes, followed by centrifugation at 12,000rpm for 5 seconds. The following reagents were added, in this order: 4µl oligo-labelling buffer (OLB), 0·6µl of 10mg/ml Bovine Serum Albumin (BSA), 4µl [$\alpha^{32}$P] dCTP (3000Ci/mmol - Amersham), 1·5µl of Klenow fragment DNA polymerase 1 and the mixture incubated at 37°C for three hours.

The OLB was made by mixing the following solutions, A:B:C, in the ratio of 10:25:15.

Solution A. To 1ml of 1·25 M Tris-HCl, 0·125 M $MgCl_2$ p*H* 8·0, add 18µl of 14·3 M β-mercaptoethanol and 5µl of a 0·1 M solution of dATP, dGTP, dTTP.

Solution B. 2 M HEPES (titrated to p*H* 6·6 with 4 M NaOH).

Solution C. Hexadeoxyribonucleotides (Pharmacia). 50 OD units evenly dissolved in 550µl TE (3 mM Tris-HCl, 0·2 mM EDTA, p*H*7·0).

After incubation the labelled probe was spun briefly and the following reagents added: 50µl of stop solution (20 mM NaCl, 20 mM Tris-HCl p*H* 7·5, 2 mM EDTA, 0·25% SLS), 20µl of 5 M NaCl, 500µl of 100% ethanol, and 3µl of 10µg/ml tRNA. The mixture was spun at 13000rpm for 15 minutes at room temperature, the liquid transferred to a fresh eppendorf tube and the activity of the pellet and liquid compared to give an approximation of the degree of incorporation of [$\alpha^{32}$P] dCTP. The pellet was resuspended in 500µl of $ddH_2O$ and boiled for 5 minutes prior to adding to the hybridisation buffer.

### 3·2e Prehybridisation and Hybridisation

Prehybridisation and Hybridisation were done in a Hybaid 'Maxi Oven' with the orbital shaking table set at medium speed.

Prehybridisation was at 62°C for 45 minutes in 0·5M $Na_2HPO_4$ (p*H*7.2), 1% SLS in a total volume of 500ml.

Hybridisation was at 62°C for 16 hours in 0·2M $Na_2HPO_4$, 1% SLS, 1% BSA, 6% polyethylene glycol 6000 and $ddH_2O$ to give a total volume of 110ml.

At the end of the hybridisation period the membranes were washed twice, for 30 minutes each time, in 500ml of 0·2M $Na_2HPO_4$ (p*H*7·2), 0·1% SLS. Throughout the prehybridisation and hybridisation processes a small 'blank' of fresh Hybond Nfp membrane was included and when the activity on this blank had been reduced to the local background, the fingerprint membranes were deemed to have been sufficiently washed.

They were then rinsed in 500ml of 3x SSC at room temperature, dried with a paper towel and wrapped in Saranwrap.

Exposure was on Kodak X-omat AR scientific imaging film, or Fuji RX x-ray film for a length of time determined by the activity on the membrane as measured with a Minimonitor 6000 (Mini-Instruments). For example, a membrane carrying 20 cps in the 'hottest' area would initially be exposed for 48 hours. The length of time of further exposures being dependent upon the results of the first. Each membrane went through the following process: probed with 33·6, stripped and re-probed with 33·15, stripped and re-probed with the labelled molecular weight marker.

### 3·2f Stripping and Re-Probing

Successful blots were stripped and re-probed. Stripping was done by immersing the membrane in 0·1% SLS that had been boiled and allowed to cool to 80°C and then placing on an orbital shaker set at low speed until the solution came down to room temperature.

### 3·2g Interpretation of Fingerprints

Autoradiographs were exposed for 1 to 14 days and at least two exposures of differing time intervals were made from each probing of each blot to avoid the problem of strongly hybridising bands overmarking proximal weaker bands. The autoradiograph was scored by eye and the data recorded in the following way.

The fingerprint autoradiograph and its associated molecular weight marker autoradiograph were aligned and overlaid with an acetate sheet. A fine line (0·2mm) was drawn across the acetate sheet at the 12kb and 2kb levels (3kb for Kf 33·6) and the molecular weight autoradiograph removed without disturbing the orientation of the fingerprint autoradiograph and acetate overlay. Fingerprint bands were marked on the acetate overlay by drawing through the most intense region of each band, as estimated by eye. A graph of molecular weight against distance migrated was constructed for each population (on linear mm graph paper) based on the appropriate molecular weight marker autoradiograph. The molecular weight of each band in the fingerprint was estimated by measuring the distance a particular band had migrated with reference to the 12kb level, and applying this measure to the graph. The information was stored in a data base program (Microsoft Excel v 4·0), which had the advantage of presenting the data in an easily interpretable way as well as allowing the straightforward accumulation of data and statistical manipulation. This method of interpretation allows comparisons to be made across many autoradiographs because the coincidence of bands between individuals is based on molecular weight rather than electrophoretic mobility, which almost certainly will not be concordant across autoradiographs. Figure 3·1 is the autoradiograph of the individuals from Brocket Park, digested with *Hin* fI and probed with Jeffreys' probe 33·15.

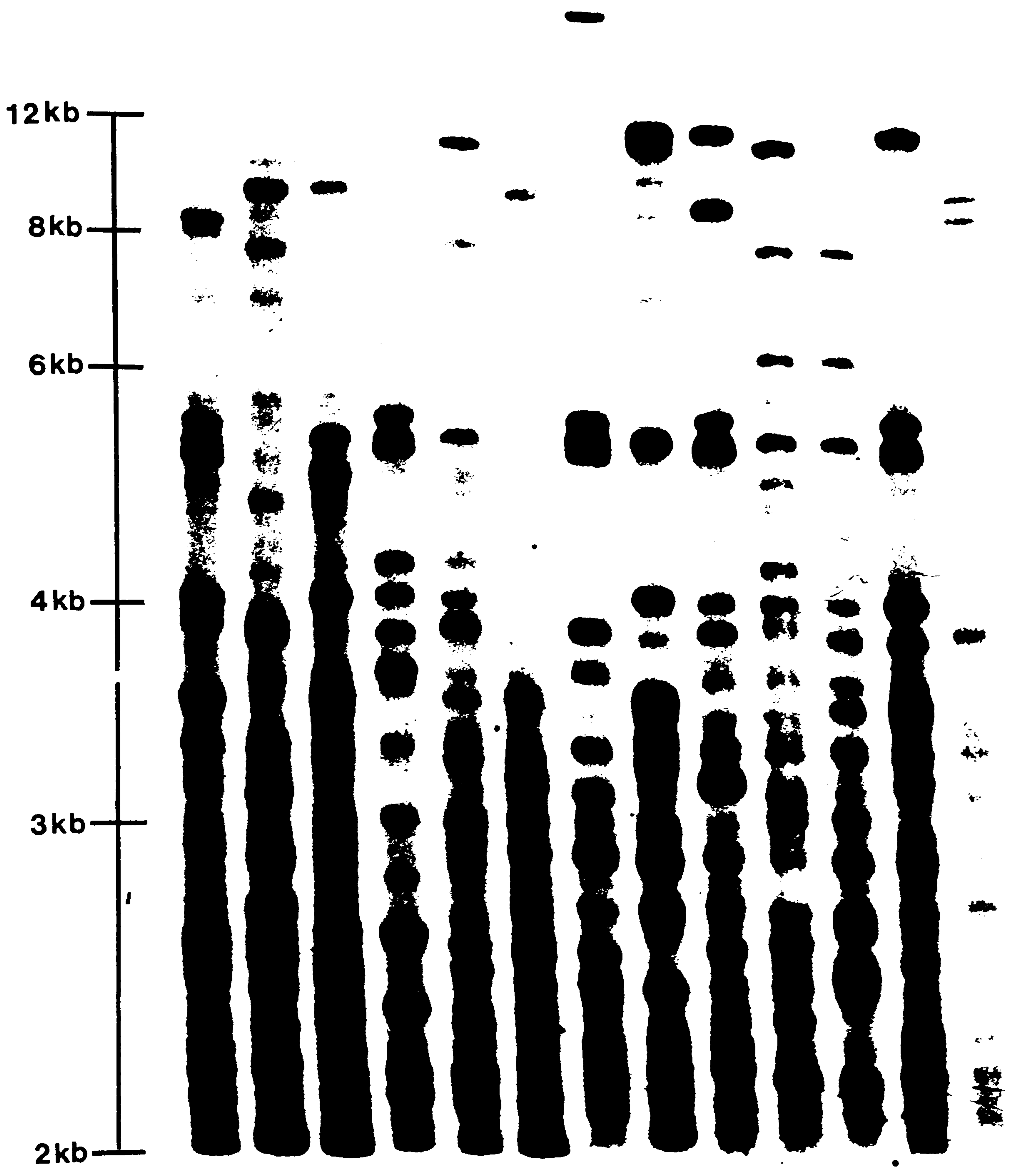


Figure 3·1 The DNA fingerprint from the Brocket Park population, digested with *Hin* fI and probed with 33·15. Lane six was not scored because of the poor quality of this sample.

### 3·2h Band Matching Criteria and Sources Of Error.

The decision that a particular band in individual A is equivalent to a particular band in individual B is based on two visual criteria, those of position and intensity. Unfortunately these criteria are not discrete entities but part of a continuum, and so the bands we see and our ability to discern individual bands are going to be influenced by a number of factors. These factors can be divided into two broad groups, neighbourhood effects and over-all running conditions of the gel. Neighbourhood effects are of course a result of the gel running conditions but I use the term here to discriminate between some of the gross effects that are easily seen and the more subtle effects which are not so apparent. As an example of a neighbourhood effect, consider two bands of similar molecular weight within the same lane. The DNA fragments for these two bands should be perpendicular to the electric field, but on many occasions bands have been observed to run at an angle to the field. Unless these bands are of low molecular weight, and thus well separated, they could appear on the final fingerprint as one intense band with equally less intense bands above and below, or one wide intense band, or one intense band with regions of differing intensity above and below. The appearance depending on the relative intensity of the two bands. Perpendicular bands will of course appear as two separate bands.

Assuming that these bands were not linked, this problem of interpretation should not be difficult to resolve, and probably be of no consequence in a paternity study. But in a population study the possibility of mis-scoring would lead to overestimating the level of band sharing (if both bands were scored as one). Another neighbourhood effect is the differential migration rates sometimes seen between the centre and edge of a gel.

The procedures from running the gel to hybridisation are slightly different for each fingerprint. Minor differences in electrophoretic conditions, gel concentration and homogeneity, concentration of buffers, strength, duration and orientation of the electric field, transfer conditions and crosslinking may all have an effect. These variables usually do not effect interpretation within a gel, but across several gels the degree of differences can become more important.

Band intensity is (bearing in mind the previously noted possible sources of error) directly related to the amount of radioactively labelled probe that binds to its target. Its target however is not a single sequence but rather a population of similar sequences to which it will bind with more or less efficiency as the stringency of the hybridisation process is respectively lowered or raised. This continuum of binding affinity will fade quite rapidly at a particular level of stringency, however, bands derived from loci that lie adjacent to this barrier may exhibit large fluctuations in intensity due to small random changes in hybridisation stringency. For this reason it is good practise to process as many membranes as possible at one time within the hybridisation chamber.

In order to make comparisons across several gels each band has to be assigned a molecular weight by reference to a known molecular weight marker, in this case at 1kb intervals up to 12kb. Errors will accumulate from measuring the molecular weight marker, measuring the fingerprint band and transferring both of these measures onto a graph. However, these interpretative errors are likely to be random in direction, hence unbiased.

Bands were scored as 'shared' if their molecular weight fell into the same molecular weight category, as determined from the graph. Because the graph of molecular weight against distance migrated is logarithmic, the magnitude of errors in reading from the graph will also be logarithmic. A band was assigned to either the top, middle or bottom of a millimetre square on the graph paper. Thus the error, in terms of positioning on the graph paper, is ±0.5mm. At the 10kb level, 0·5mm on the X axis represents 1kb on the Y axis; at this level the error was then ± 1kb (10%). At the 3kb level 0·5 mm represents 25b; at this level the error was ±25b (0·8%).

### 3·2i Methods of Analysis

As an assessment of the variation between individuals, the frequency distributions of the number of bands scored per individual were calculated for both probes individually and combined and the distribution of the frequencies calculated. Statistical calculations were performed using the statistics software included in the data-base program Microsoft Excel v 4·0. The incidence of co-detected bands and the individual specificity of fingerprints were calculated as per Jeffreys *et al* (1991).

The calculations for relatedness as determined by band sharing, with corrections for non-independent data were performed as described by Lynch (1990, 1991).

Analysis of variance (anovar) was calculated using the statistics software in Microsoft Excel v 4·0. Anovars were calculated for two data sets; (1) the complete data set, i.e. all possible combinations of pairs of individuals, and thus non-independent. (2) A partial data set derived from comparisons of unique non-overlapping individuals, and thus independent.

The relationships between populations were expressed as dendrograms based on the average percentage difference (APD) (Gilbert *et al* 1990) between populations.

Finally, the level of population sub-division was assessed by calculating heterozygosity as well as applying Nei's (1987) estimator of gene flow between populations.

## 3·3 RESULTS

The assumption in this and proceeding chapters is that data have been accumulated from individuals drawn at random from the various populations. There is however the caveat that deer were shot by individuals or small groups of hunters working within a small area at any one time and thus it is likely that the deer killed over a period of a few days could be more closely related than would otherwise be expected from a true random sample.

### 3·3a Number of Bands

To investigate the variation between individuals in terms of the numbers of bands scored with probes 33·15 and 33·6, the frequency distributions of the bands scored were determined. The number of bands scored per individual ranged from 11 to 23 (33·6), from 15 to 33 (33·15) and from 29 to 50 for both probes combined. Figure 3·2 graphically displays the frequency relationship between the two probes, and Figure 3·3 is of the combined probe distribution. The complete data set may be found in Appendix B.

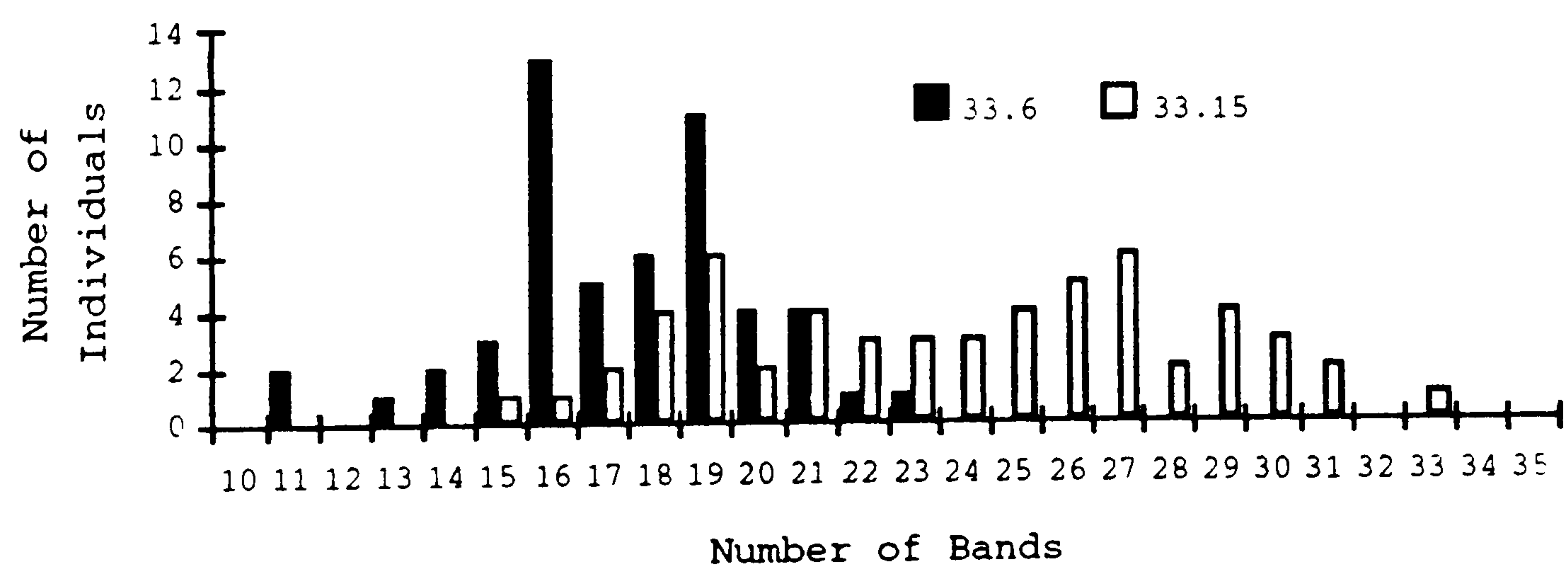


Figure 3·2 Individual variation in the number of bands scored with the two probes

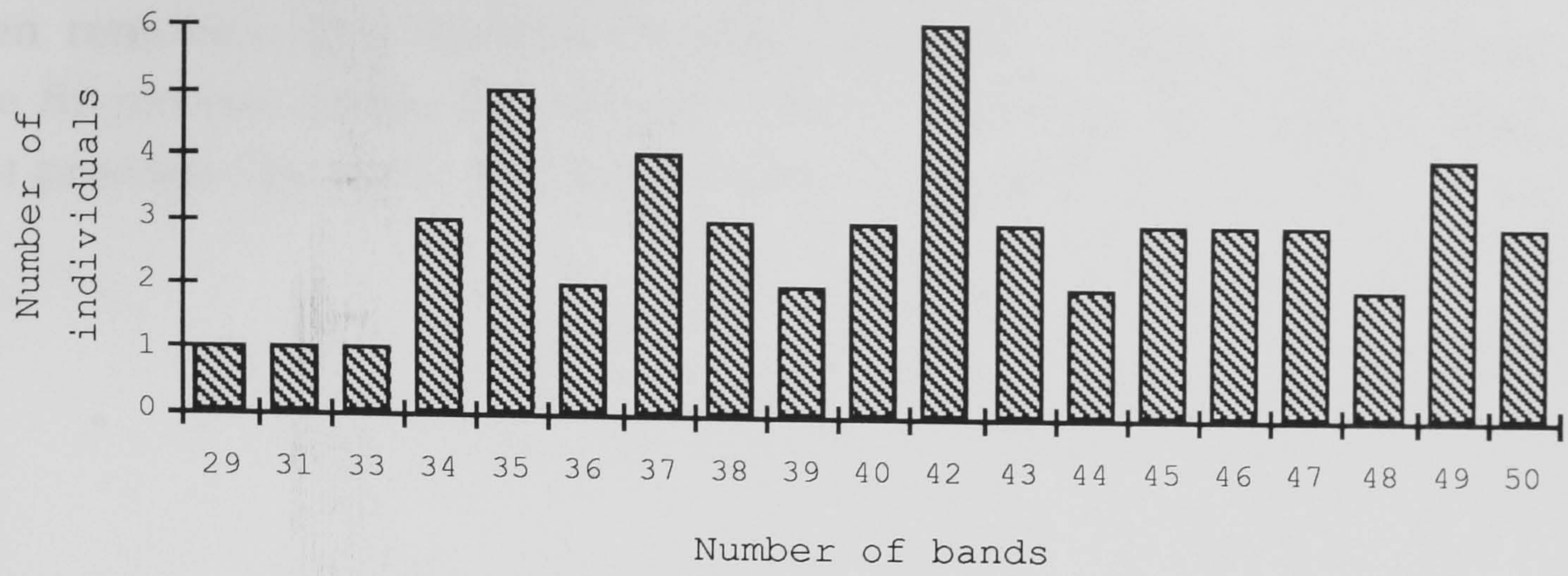


Figure 3·3 Individual variation in the number of bands scored at 33·15 + 33·6

These findings broadly coincide with those of Jeffreys *et al* (1991) in which a much larger data set was investigated. The mean number of bands recorded for muntjac is also broadly similar to that found in many other outbred species.

Table 3·1 lists some examples.

| | Enzyme | mean number of bands 33·15 | mean number of bands 33·6 | Reference |
|---|---|---|---|---|
| Muntjac (>2kb) | *Hinf*1 | 23·75 ± 4·5 | 17·33 ± 2·6 | this work |
| Red deer | *Alu* 1 | 15 - 30 | - | Pemberton *et al* 1992 |
| Human (>3·5kb) | *Hinf*1 | 17·07 ± 3·3 | 18·06 ± 3·5 | Jeffreys *et al* (1991) |
| Grey seal (>3kb) | *Hinf*1 | 20 ± ? | - | Amos *et al* (1993 b) |
| Domestic Dog (>3kb) | *Hinf*1 | 19·4 ± 1·4 | 15·9 ± 1·6 | Jeffreys & Morton (1987) |
| Dwarf Mongoose (>?kb) | *Alu*I | 16·1 ± 4·05 | 16.5 ± 3·95 | Keane *et al* (1994) |
| Wild Muskrats (>3·4kb) | *Hae*III | 14·4 ± 2·8 | - | Marinelli *et al* (1992) |

Table 3·1 The mean number of bands recorded for various species

### 3·3b Co-detection of Bands

Inspection of the autoradiographs indicated that there may be some restriction fragments that were detected by both probes. These co-detected fragments are defined as bands with equivalent electrophoretic mobility and similar intensity. In a population study such as this, an examination of autoradiographs can only give a first approximation of the level of co-detection of bands. Many bands that appear to be similar may have co-migrated coincidentally and, in terms of chromosomal origins, be completely dissimilar. Figure 3·4 is a 'compressed' and intercalated representation of two autoradiographs produced by the probes 33.15 and

33.6, from the same population. Compressed in that blank regions have been removed. This example is taken from the Knebworth population. The fingerprint produced by probe 33.15 on individual A lies alongside that produced by probe 33.6 for the same individual.

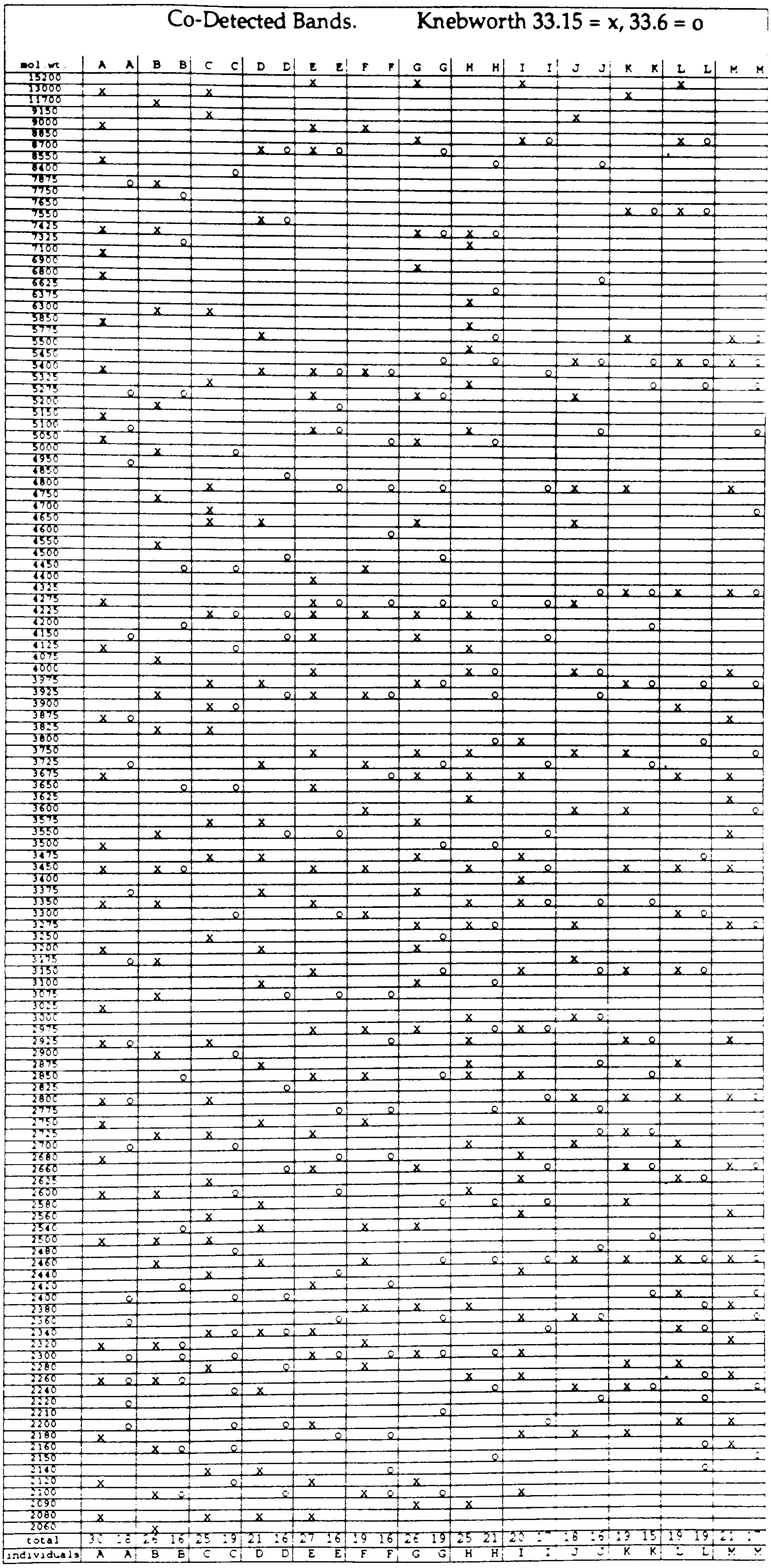


Figure 3·4 Co-detected bands from the Knebworth sub-population, probed with 33.15 & 33.6

This arrangement of the data allows comparisons to be easily made. Of the total number of bands scored, 1259 for 33.15 and 966 for 33.6, there were 192 incidences of co-detection involving 77 bands (39·7%) out of a total of 194 bands (i.e. molecular weight intervals) ranging from 2000bp to 16300bp. The numbers of co-detected bands per population and mean per individual are arranged in Table 3·2. The range refers to the maximum and minimum numbers of shared bands per population and is calculated as per equation 3·1 (see below), x100.

| Population | Total number of co-detected bands | Mean number of co-detected bands per individual | Standard Deviation | Range per individual |
|---|---|---|---|---|
| | | | | |
| Knb | 59 | 4·5 | ± 1·8 | 13% - 42% |
| Bkp | 50 | 4·1 | ± 1·9 | 0·6% - 44% |
| Btf | 42 | 3·5 | ± 1·7 | 0 - 24% |
| Kf | 41 | 2·6 | ± 1·4 | 0 - 30% |
| Total | 192 | 3·6 | ± 1·8 | 0 - 44% |

Table 3·2 Numbers of co-detected bands per population

The figures quoted in Table 3·2 seem to be high and a more realistic approach may be gained by examining a large number of parent offspring segregation patterns, with both probes. A segregation analysis will demonstrate if the co-detected bands are co-segregating or artifactual. In Jeffreys *et al* (1991), 36 families containing 5 to 9 offspring were analysed; a total of 1077 parental bands were scored with 33·6 and 970 scored with 33·15 and 'there were only 12 instances of apparently co-detected bands'. This is equivalent to 1·2% (cf with the 'range' in the above table). Jeffreys *et al* concluded that 'the degree of overlap between loci scored in the 33·6 and 33·15 fingerprints is negligible'.

### 3·3c Estimation of Individual Specificity of Muntjac DNA Fingerprints

The procedure I have used is that which is generally adopted for estimating individual specificity (Jeffreys *et al* 1991). The mean number of resolvable bands per individual is calculated, as is the mean probability that a band from individual 'A' is matched by a band in individual 'B'. (See Materials and Methods section **3·2h** for a discussion of band matching criteria). Implicit in this estimation is complete statistical independence of bands and uniformity of the mean probability over all bands.

The probability that all of 'A's bands are included within the fingerprint of 'B' is given by $x^n$ where :-

$x$ = the mean probability that a band in 'A' is matched by a band in 'B'
$n$ = the mean number of resolved bands per individual.

Table 3·3 gives those probabilities for fingerprint fragments between 20kb and 2kb for all the populations investigated (20kb to 2·6kb for Kf/33·6).

| | 33·6 | | |
|---|---|---|---|
| Pop. | mean band sharing prob.(x) (±SD) | mean No of bands (n) (±SD) | Probability |
| Knb | 0·153 (± 0·09) | 17·31 (± 1·75) | $7{\cdot}71 \times 10^{-15}$ |
| Bkp | 0·330 (± 0·124) | 16·33 (± 1·97) | $1{\cdot}37 \times 10^{-8}$ |
| Btf | 0·253 (± 0·129) | 16·50 (± 3·13) | $1{\cdot}42 \times 10^{-10}$ |
| Kf | 0·160 (± 0·127) | 13·38 (± 2·00) | $2{\cdot}24 \times 10^{-11}$ |
| Total | 0·220 (± 0·120) | 18·50 (± 3·46) | $6{\cdot}84 \times 10^{-16}$ |

| | 33·15 | | |
|---|---|---|---|
| Pop. | mean band sharing prob. (±SD) | mean No of bands (n) (±SD) | Probability |
| Knb | 0·186 (± 0·079) | 23·00 (± 4·02) | $1{\cdot}58 \times 10^{-17}$ |
| Bkp | 0·301 (± 0·109) | 19·00 (± 2·33) | $1{\cdot}24 \times 10^{-10}$ |
| Btf | 0·237 (± 0·076) | 27·67 (± 3·17) | $5{\cdot}00 \times 10^{-18}$ |
| Kf | 0·255 (± 0·096) | 25·50 (± 3·10) | $7{\cdot}35 \times 10^{-16}$ |
| Total | 0·244 (± 0·100) | 24·00 (± 4·00) | $1{\cdot}98 \times 10^{-15}$ |

Table 3·3 Estimations of Individual Specificity of Muntjac Fingerprints

The above probabilities refer only to the chance that all of 'A's bands are matched by all of 'B's and make the assumption that B is unrelated to A. The possibility of unmatched bands in B not being present in A is not taken into account.

### 3·3d Relatedness As Determined by Band Sharing

**i) Within Sub-Populations** The similarity index (or band sharing coefficient) between any two individuals was calculated as twice the number of restriction fragments coincident between two individuals, divided by the total number of fragments scored for the same two individuals (Lynch, 1990).

$$S_{xy} = \frac{2n_{xy}}{(n_x + n_y)} \tag{3·1}$$

Where: $S_{xy}$ = Similarity index of individuals x and y,
$n_{xy}$ = the number of coincident bands between x and y,
$(n_x + n_y)$ = the total number of bands from x and y.

The mean band sharing coefficients and standard deviations, both within and between populations, were calculated as:

$$\bar{S} = \frac{1}{n}\sum_{i=1}^{n} xi \qquad \text{and} \qquad \sigma = \sqrt{\frac{\sum x^2 - \left(\frac{\sum x}{n}\right)^2}{n-1}} \tag{3·2) \& (3·3}$$

Where: $\bar{S}$ = mean similarity index and $\sigma$ = standard deviation.
However the usual equation for estimating the standard deviation assumes independence of data. The pairwise comparisons presented here are not independent and so would produce downwardly biased estimates. To overcome this problem I have used the equation derived by Lynch (1990, 1991),

$$\sigma' = \left[\frac{2\bar{S}(1-\bar{S})(2-\bar{S})}{\bar{n}(4-\bar{S})}\right]^{0.5} \tag{3·4}$$

in which: $\sigma'$ = unbiased standard deviation,
$\bar{S}$ = mean of all pairwise values,
$\bar{n}$ = mean number of bands per individual.

Table 3·4 provides a summary of the statistics derived from all four populations.

| | | 33·15 | 33·6 | 33·15 | 33·6 |
|---|---|---|---|---|---|
| Pop | $n_i$ (33·15/33·6) | $\bar{n}$ | $\bar{n}$ | $\bar{s}$ (±SD) (σ′) | $\bar{s}$ (±SD) (σ′) |
| Knb | 13/13 | 23·0 ±4·0 | 17·3 ±1·7 | 0·187 (±0·08) (±0·08) | 0·153 (±0·09) (± 0·084) |
| Bkp | 12/12 | 19·0 ±2·1 | 16·3 ±1·9 | 0·301 (±0·11) (±0·101) | 0·343 (±0·118) (± 0·11) |
| Btf | 12/12 | 27·7 ±3·2 | 16·5 ±3·2 | 0·237 (±0·076) (±0·08) | 0·253 (±0·13) (± 0·101) |
| Kf | 16/18 | 25·5 ±3·1 | 13·4 ±2·0 | 0·255 (±0·096) (±0·08) | 0·160 (±0·13) (± 0·098) |

Table 3·4 Summary statistics within all four sub-populations. $n_i$ = number individuals from each population probed with Jeffreys' probes 33·15/33·6 respectively, $\bar{n}$ = mean N° of bands per individual, $\bar{s}$ = mean of all pairwise values and σ′ = unbiased standard deviation

The band sharing coefficients are broadly similar for 33·15 and 33·6 per sub-population with the exception of Kf 33.15/Kf 33.6, in which there appears to be a larger difference. To determine how concordant the mean band sharing coefficients are, a single factor analysis of variance was performed to test the hypothesis that the mean from both probings are equal. The results are presented in Table 3·5.

| Groups | n | Mean BSC | Var. | Source of Variation | SS | df | MS | F | F crit 5% | Prob. of occurrence |
|---|---|---|---|---|---|---|---|---|---|---|
| Knb33·15 | 78 | 0·19 | 0·01 | Between Groups | 0·044 | 1 | 0·044 | 6·042 | 3·903 | 0·0151 |
| Knb 33·6 | 78 | 0·15 | 0·01 | Within Groups | 1·113 | 154 | 0·007 | | | |
| | | | | Total | 1·157 | 155 | | | | |
| | | | | | | | | | | |
| Bkp33·15 | 66 | 0·33 | 0·02 | Between Groups | 0·005 | 1 | 0·005 | 0·342 | 3·914 | 0·5596 |
| Bkp 33·6 | 66 | 0·34 | 0·01 | Within Groups | 1·907 | 130 | 0·015 | | | |
| | | | | Total | 1·912 | 131 | | | | |
| | | | | | | | | | | |
| Btf33·15 | 66 | 0·24 | 0·01 | Between Groups | 0·008 | 1 | 0·008 | 0·742 | 3·914 | 0·3906 |
| Btf 33·6 | 66 | 0·25 | 0·02 | Within Groups | 1·464 | 130 | 0·011 | | | |
| | | | | Total | 1·472 | 131 | | | | |
| | | | | | | | | | | |
| Kf 33·15 | 120 | 0·26 | 0·01 | Between Groups | 0·664 | 1 | 0·664 | 50·58 | 3·876 | 1·0E-11 |
| Kf 33·6 | 153 | 0·16 | 0·02 | Within Groups | 3·557 | 271 | 0·013 | | | |
| | | | | Total | 4·22 | 272 | | | | |

df = degrees of freedom  MS = mean square  F = variance ratio

Table 3.5 Summary of results of single factor Anovar, where n = number of comparisons within a sub-population, BSC = band sharing coefficient and F = variance ratio.

The probability of occurrence represents the probability of the means being equal in two samples drawn at random from the populations. Therefore, with 99·5% certainty, the means of Knb33·15/Knb33·6 are dissimilar, those of

Btf33·15/Btf33·6 and Bkp33·15/Bkp33·6 are similar and those of Kf33·15/Kf33·6 are dissimilar.

**ii) Between Populations** A first analysis of the band sharing coefficient between populations is presented in Table 3·6. These measures have been made by comparing individual 1 (population A) with individuals 1, 2, 3, etc. (population B), for all combinations of populations, following the same procedures detailed in the previous section.

| | | 33·15 | 33·6 | 33·15 | 33·6 |
|---|---|---|---|---|---|
| Pop. | $n_i$ 15/6 | $\bar{n}$ | $\bar{n}$ | $\bar{s}$ (±SD) ($\sigma'$) | $\bar{s}$ (±SD) ($\sigma'$) |
| Knb/Bkp | 25/25 | 21·0 ±3·8 | 16·8 ±1·9 | 0·126 (± 0·07) (± 007) | 0·129 (± 0·08) (± 0·08) |
| Knb/Btf | 25/25 | 25·0 ±4·3 | 16·9 ±2·5 | 0·141 (± 0·07) (± 0·07) | 0·138 (± 0·08) (± 0·08) |
| Knb/Kf | 29/31 | 24·0 ±3·7 | 15·0 ±2·7 | 0·141 (± 0·07) (± 0·07) | 0·091 (± 0·08) (± 0·07) |
| Bkp/Btf | 24/24 | 23·0 ±5·3 | 16·4 ±2·6 | 0·154 (± 0·06) (± 0·07) | 0·144 (± 0·08) (± 0·08) |
| Bkp/Kf | 28/30 | 23·0 ±4·4 | 14·6 ±2·5 | 0·146 (± 0·07)( ± 0·07) | 0·161 (± 0·11) (± 0·09) |
| Btf/Kf | 28/30 | 26·0 ±3·3 | 14·6 ±2·9 | 0·175 (± 0·07) (± 0·07) | 0·128 (± 0·09) (± 0·09) |

| | |
|---|---|
| $n_i$ = number individuals (33·15/33·6) | $\bar{n}$ = mean N° of bands per individual |
| $\bar{s}$ = mean of all pairwise values | $\sigma'$ = unbiased standard deviation |

Table 3·6 Summary statistics between all four populations

Again there are some discrepancies between similarity as determined by probes 33·15 and 33·6. Nevertheless, all combinations of populations show a very low level of mean band sharing coefficient and therefore a low level of similarity. Further analysis is obviously needed to define more accurately the relationships among populations.

### 3.3e Analysis of Variance

Graphical representation of the band sharing coefficients from all populations indicates that the data may not follow a standard normal distribution. Figures 3·5 to 3·8 are graphical representations of the frequency of band sharing coefficients produced by the two probes used. Figure 3·5 is derived from the total band sharing coefficients of the four populations, as produced by Jeffreys' probe 33·6. Figure 3·6 relates to 33·15, Figure 3·7 is of both data sets combined and Figure 3·8 is of both data sets summed.

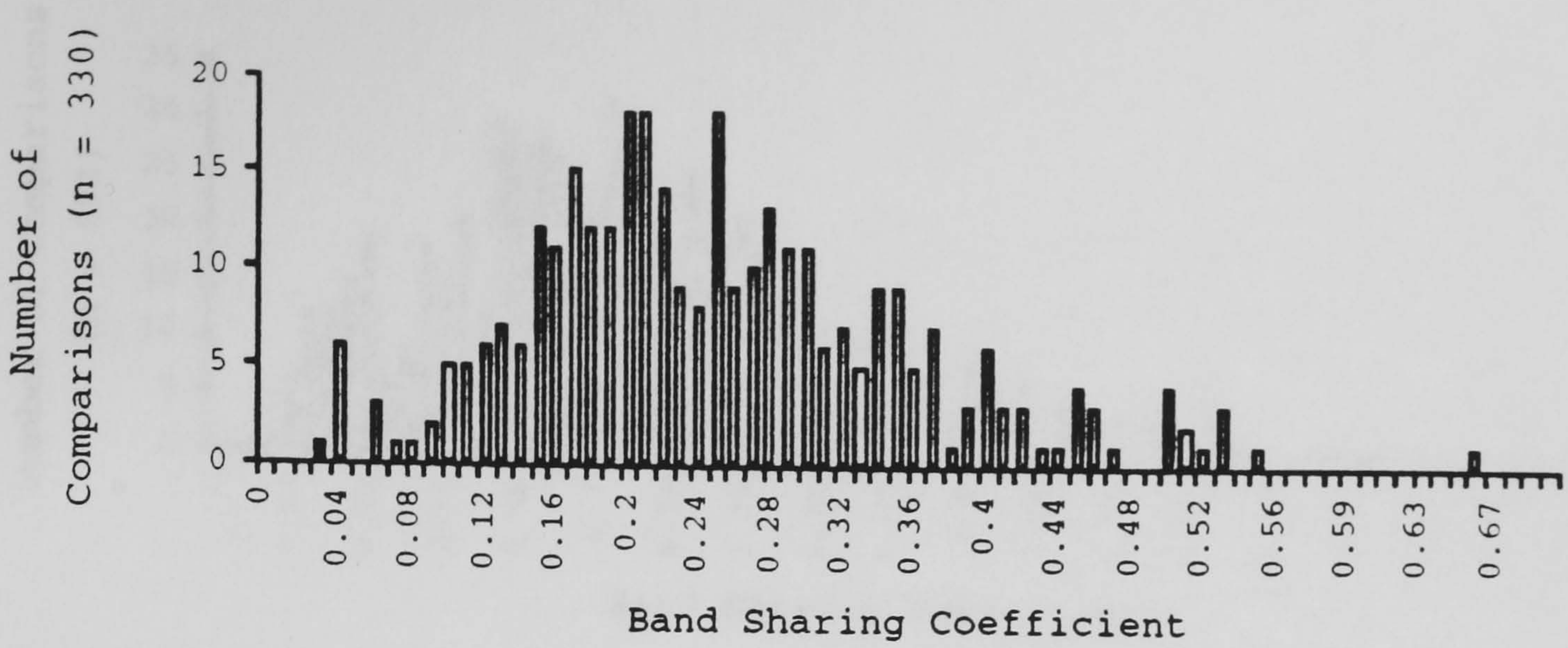


Figure 3·5 Frequency of band sharing coefficients - 33·6

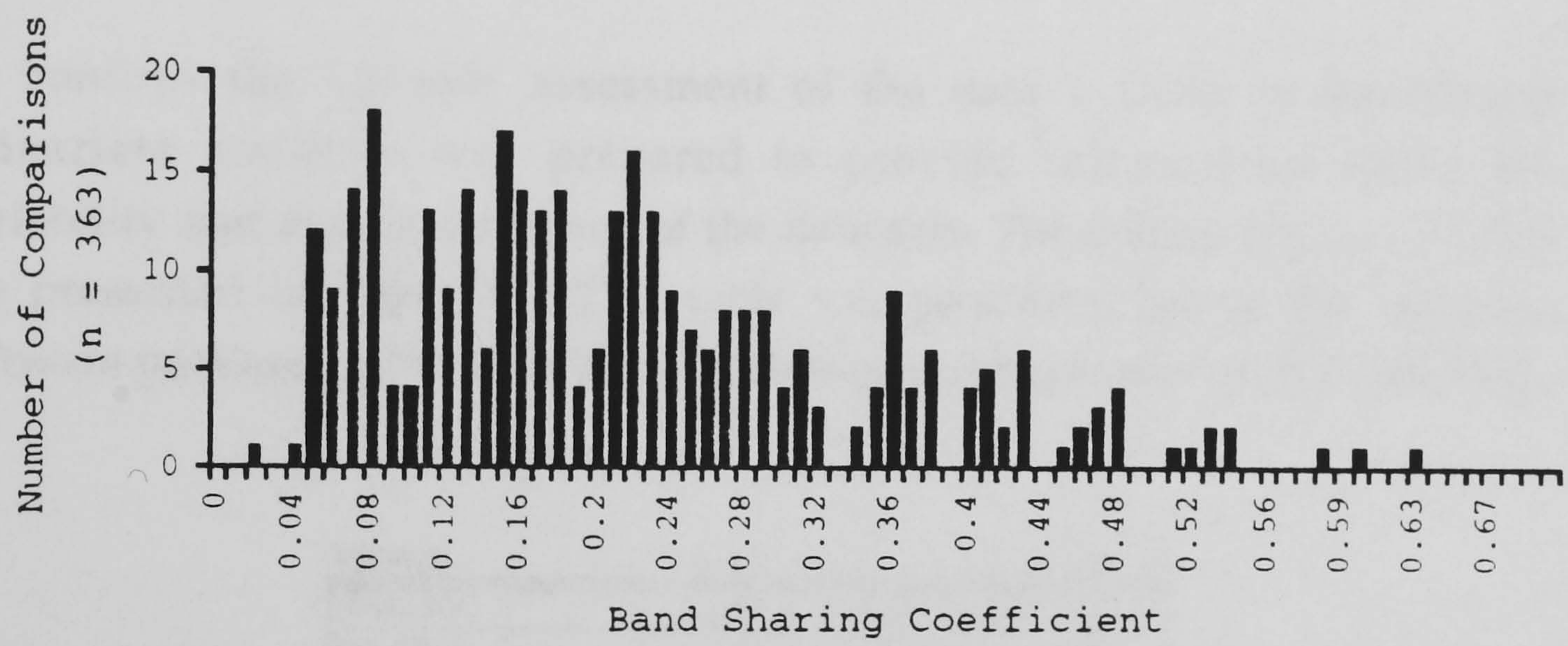


Figure 3·6 Frequency of band sharing coefficients - 33·15

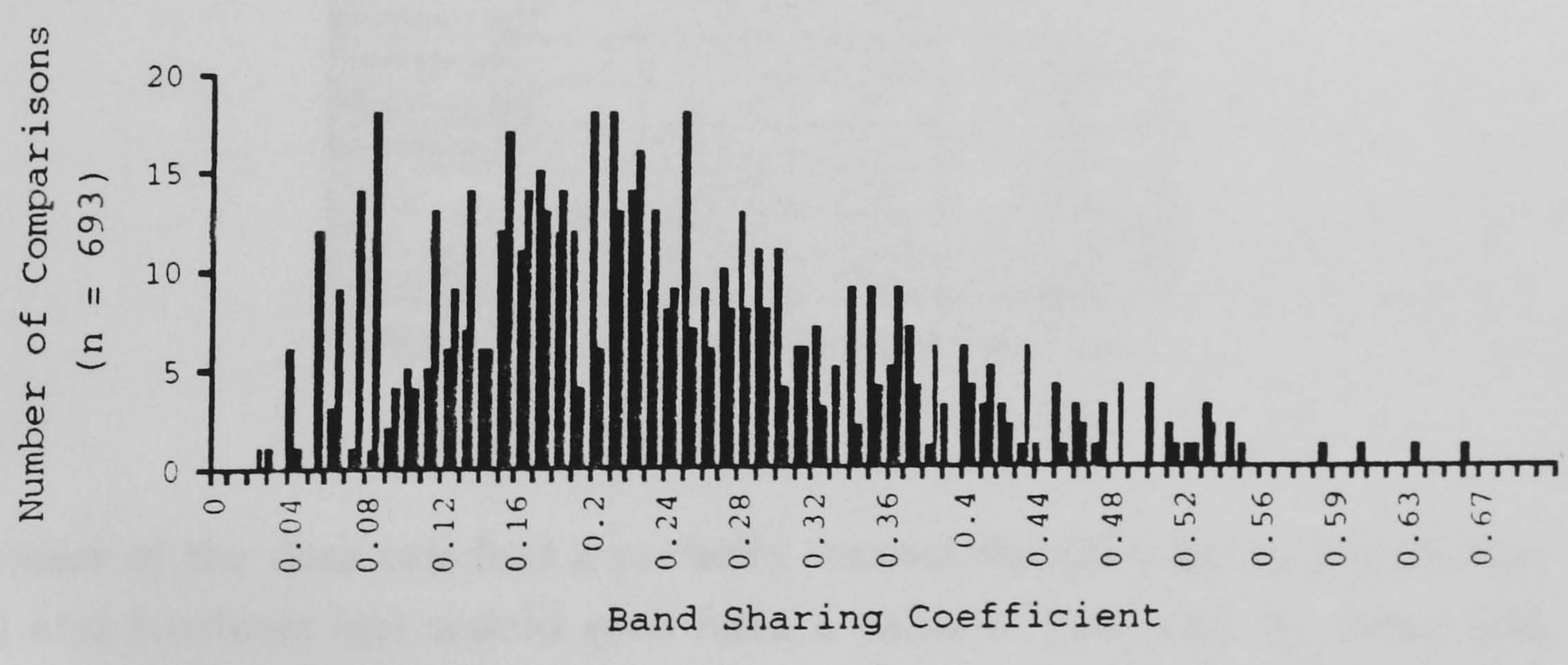


Figure 3·7 Frequency of band sharing coefficients - data from 33·6 and 33·15 combined

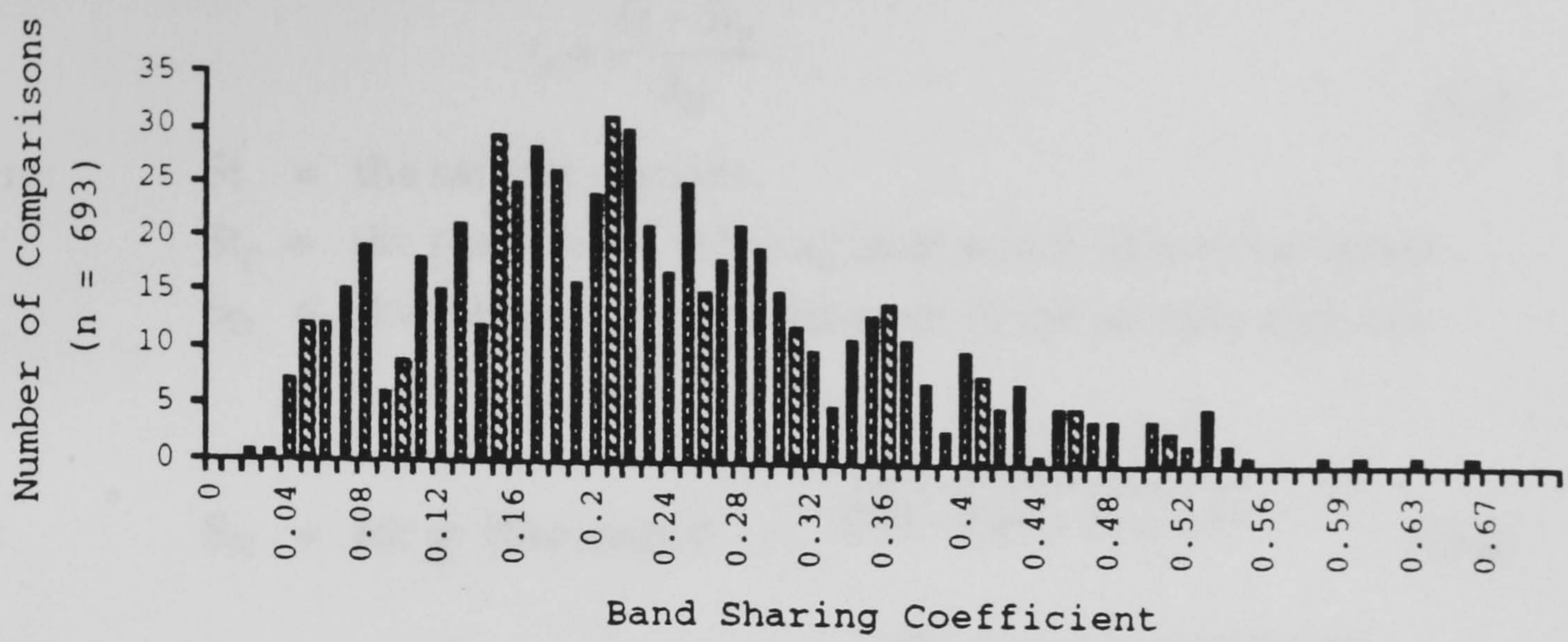


Figure 3·8 Frequency of band sharing 33·6+33·15 (summed)

To confirm the 'by eye' assessment of the data a table of descriptive univariate statistics was prepared to provide information about the variability and central tendency of the data sets. These descriptive statistics are presented in Table 3·7. This table was produced using the statistics software package included in the data base programme Microsoft Excel V4·0.

| Statistic | 33.15 | 33.6 | 33.15+33.6 |
|---|---|---|---|
| Mean BSC | 0.25 | 0.199 | 0.225 |
| Std. Error | 0.006 | 0.007 | 0.005 |
| Median | 0.24 | 0.18 | 0.21 |
| Mode | 0.25 | 0 | 0 |
| Std. Deviation | 0.106 | 0.132 | 0.123 |
| Variance | 0.011 | 0.017 | 0.015 |
| Skewness (g1) | 0.637 | 0.676 | 0.473 |
| Kurtosis (g2) | 0.53 | 0.129 | 0.032 |
| Minimum BSC | 0.03 | 0 | 0 |
| Maximum BSC | 0.66 | 0.63 | 0.66 |
| Range | 0.63 | 0.63 | 0.66 |
| Sum | 82.62 | 72.32 | 155.76 |
| Count (n) | 330 | 363 | 693 |

Table 3·7 Descriptive Statistics of Raw Data

If either of the data sets had a perfectly normal distribution then Skewness ($g_1$) and Kurtosis ($g_2$) would each have a value of zero and the mean and median would coincide. To assess the significance of the deviation from the expected value of zero for the parameters $\gamma_1$ (Skewness) and $\gamma_2$ (Kurtosis) a significance test ($t_s$) was applied (Sokal and Rohlf, 1981).

$$t_s = \frac{St - St_p}{S_{St}} \tag{3·5}$$

where: St = the sample statistic,

$St_p$ = the parametric value against which St is to be tested,

$S_{St}$ = the estimated standard error of the sample statistic.

and: $S_{St}$ = for $g_1$ (Skewness) $= \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}}$ (3·6)

$S_{St}$ = for $g_2$ (Kurtosis) $= \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$ (3·7)

where: n = total number of measures (the 'count' from the above table)

| | | 33.15 | 33.6 | 33.15+33.6 |
|---|---|---|---|---|
| ts ($g_1$) | = | 4·726 | 5·25 | 5·09 |
| ts ($g_2$) | = | 1·966 | 0·502 | 0·17 |

from a table of Student's t-distributions (Campbell, 1992. Table A12) the appropriate critical values are:

$t\,0{\cdot}5_{[\infty]} = 1{\cdot}960$ $t\,0{\cdot}1_{[\infty]} = 2{\cdot}576$ $t\,0{\cdot}001_{[\infty]} = 3{\cdot}291$

The conclusion drawn from these significance tests are:

1. For all three data sets, $t_s(g_1)$ is greater than the t-critical value at the 0·01% level, therefore $\gamma_1 \neq 0$; $g_1$ is positive, indicating that $\gamma_1 > 0$, and therefore the distributions are skewed towards higher values.
2. For both data sets $t_s(g_2)$ is equal to (33·15) or less than (33·6 & 33·15+33·6) the t-critical value at the 5% level, indicating the acceptability of taking $\gamma_2 \approx 0$, and thus the distributions are not significantly kurtotic.

Because the above data is derived from multiple comparisons (i.e. all possible combinations of pairs of individuals), the data are not independent. Similarity measures will tend towards positive correlation because individuals that happen to share bands that are relatively common in a population will tend to have high similarities with most other individuals. The opposite situation will apply to those individuals that exhibit rare bands. As a comparison and to test whether another data set would be more appropriate for analysis of variance, band sharing coefficients were drawn from the raw data tables in such a way that each similarity estimate was derived from a unique pair of non-overlapping individuals. I.e. the coefficients used were those between individuals A-B, C-D, E-F etc. from within populations and A-A, B-B, C-C, etc. from between populations (Lynch 1991). The data is graphically represented in Figures 3·9 to 3·12; as previously described.

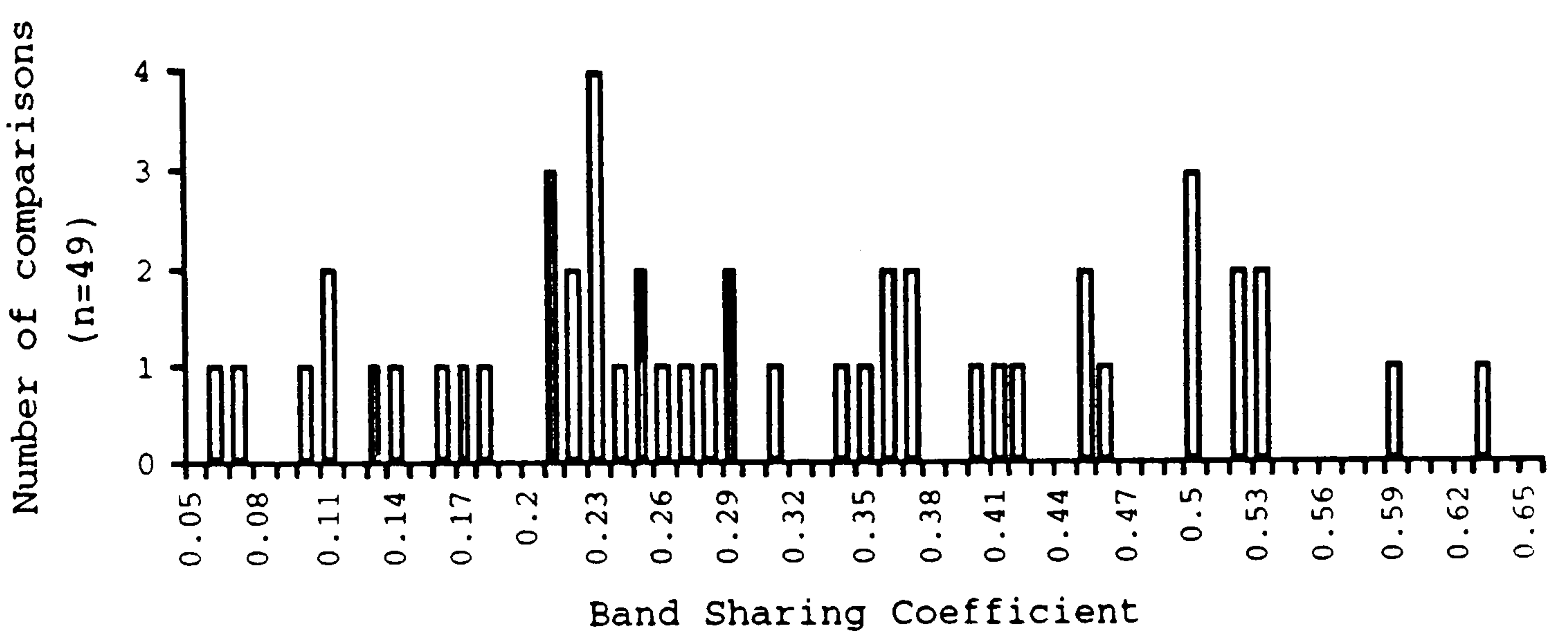


Figure 3·9 Frequency of band sharing coefficients (independent) 33·6

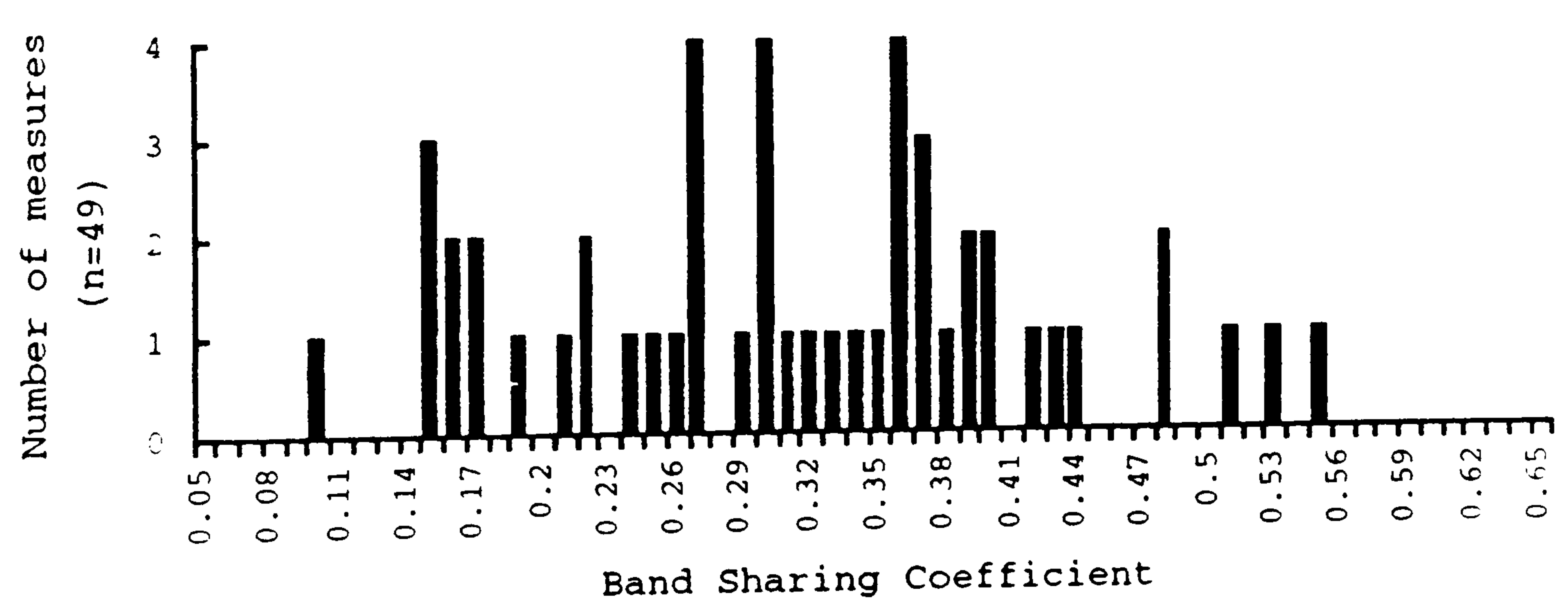


Figure 3·10 Frequency of band sharing coefficients (independent) 33·15

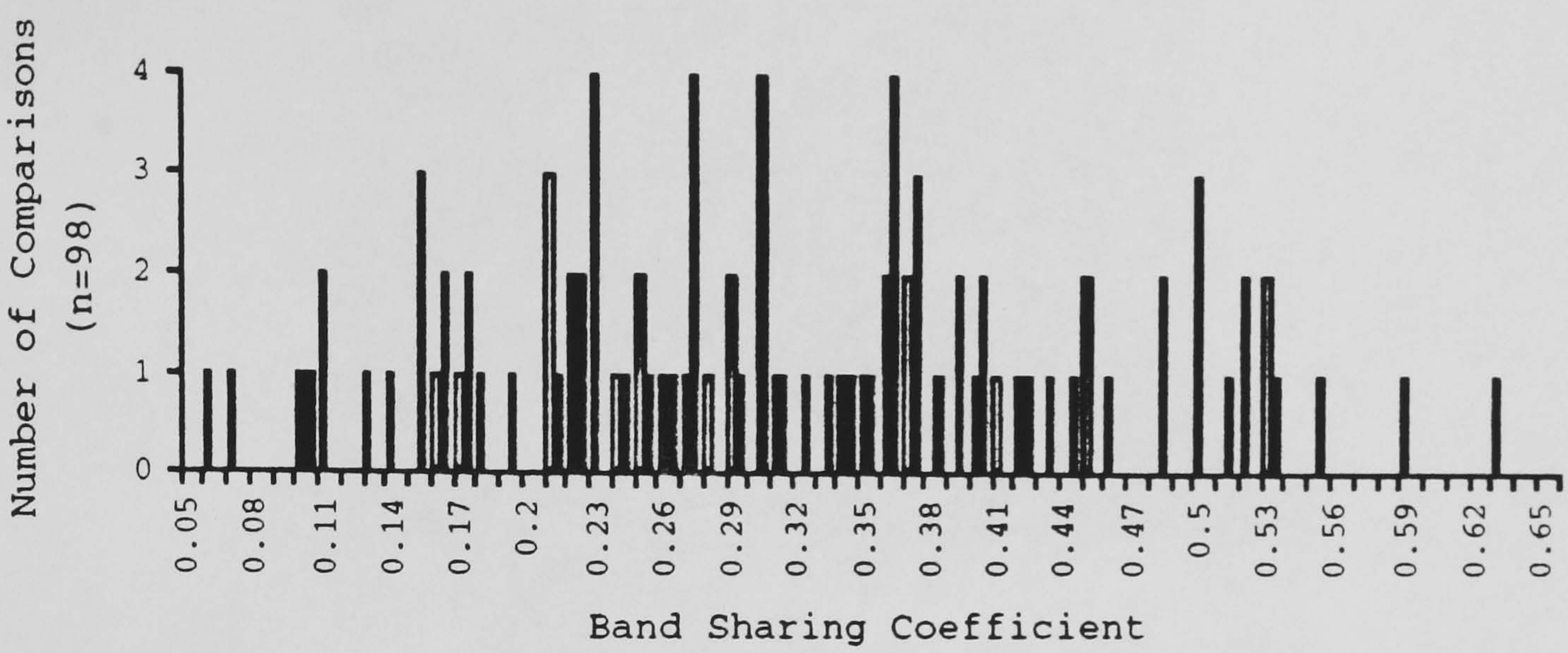


Figure 3·11 Frequency of band sharing coefficients (independent) 33·15/33·6 combined

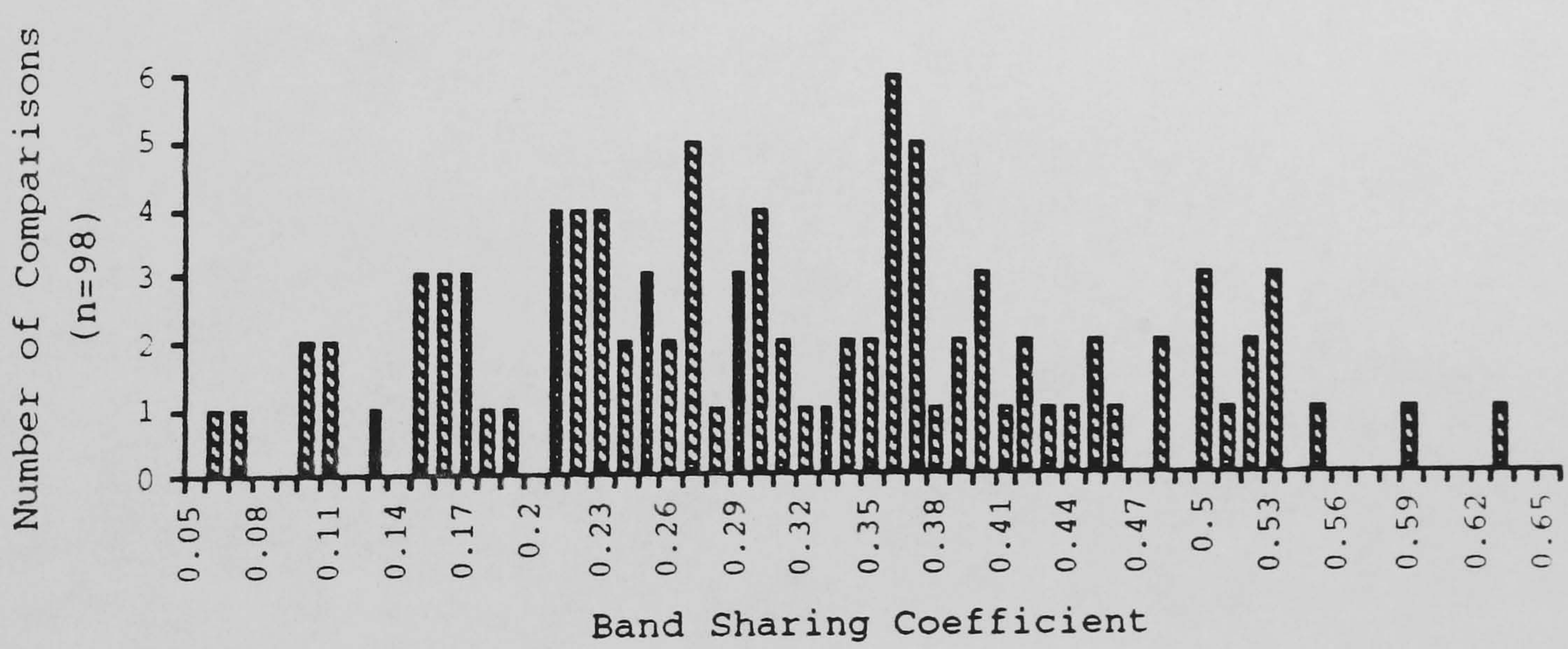


Figure 3·12 Frequency of band sharing coefficients (independent) 33·15+33·6 summed

Descriptive univariate statistics and the appropriate significance tests were produced for these data sets in the same manner as previously described.

| Statistic | 33.15 | 33.6 | 3.15+33.6 |
|---|---|---|---|
| Mean BSC | 0.314 | 0.311 | 0.312 |
| Std. Error | 0.016 | 0.021 | 0.013 |
| Median | 0.31 | 0.28 | 0.3 |
| Mode | 0.3 | 0.23 | 0.36 |
| Std. Deviation | 0.109 | 0.151 | 0.132 |
| Variance | 0.012 | 0.023 | 0.017 |
| Skewness (g1) | 0.109 | 0.235 | 0.19 |
| Kurtosis (g2) | -0.51 | -0.89 | -0.62 |
| Minimum BSC | 0.1 | 0.05 | 0.05 |
| Maximum BSC | 0.55 | 0.63 | 0.63 |
| Range | 0.45 | 0.58 | 0.58 |
| Sum | 15.37 | 15.78 | 31.15 |
| Count (n) | 49 | 51 | 100 |

Table 3·8 Descriptive Statistics of Independent Data

| | | 33.15 | 33.6 | 33.5+33.6 |
|---|---|---|---|---|
| $t_s\ (g_1)$ | = | 0·32 | 0·70 | 0·79 |
| $t_s\ (g_2)$ | = | -0·76 | -1·36 | -1·29 |

from a table of Student's t-distributions (Campbell, 1992. Table A12) the appropriate critical values are:

$t0{\cdot}5_{[50]} = 2{\cdot}09$ $\qquad$ $t\,0{\cdot}1_{[50]} = 2{\cdot}68$ $\qquad$ $t0{\cdot}001_{[50]} = 3{\cdot}50$

The conclusion drawn from these significance tests are:

1. For all three data sets, $t_s(g_1)$ is less than the t-critical value at the 0·01% level, therefore $\gamma_1 \approx 0$; and the distributions are not significantly skewed.
2. For all three data sets $t_s(g_2)$ is less than the t-critical value at the 5% level, indicating the acceptability of taking $\gamma_2 \approx 0$, and thus the distributions are not significantly kurtotic.

An analysis of variance was performed on the data sets described above. In Table 3·9 a comparison is made between the two groups of data; original and independent for 33·15, 33·6 and both probes combined.

| ORIGINAL DATA | | | | | | INDEPENDENT DATA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Groups | n | Mean BSC | F | F crit @ 5% | Prob. | Prob. | F crit @ 5% | F | Mean BSC | n | Groups |
| Knb33.15 | 78 | 0.186 | 52.27 | 3.907 | 3E-11 | 0.014 | 4.325 | 7.141 | 0.246 | 12 | Knb33.15 |
| Bkp33.15 | 66 | 0.301 | | | | | | | 0.354 | 11 | Bkp33.15 |
| Knb33.15 | 78 | 0.186 | 15.00 | 3.907 | 2E-04 | 0.231 | 4.325 | 1.52 | 0.246 | 12 | Knb33.15 |
| Bft33.15 | 66 | 0.237 | | | | | | | 0.295 | 11 | Bft33.15 |
| Knb33.15 | 78 | 0.186 | 27.67 | 3.889 | 4E-07 | 0.014 | 4.242 | 6.933 | 0.246 | 12 | Knb33.15 |
| Kf33.15 | 120 | 0.255 | | | | | | | 0.353 | 15 | Kf33.15 |
| Bkp33.15 | 66 | 0.301 | 14.97 | 3.914 | 2E-04 | 0.175 | 4.351 | 1.979 | 0.354 | 11 | Bkp33.15 |
| Bft33.15 | 66 | 0.237 | | | | | | | 0.295 | 11 | Bft33.15 |
| Bkp33.15 | 66 | 0.301 | 8.596 | 3.892 | 0.004 | 0.982 | 4.26 | 5E-04 | 0.354 | 11 | Bkp33.15 |
| Kf33.15 | 120 | 0.255 | | | | | | | 0.353 | 15 | Kf33.15 |
| Bft33.15 | 66 | 0.237 | 1.765 | 3.892 | 0.18 | 0.182 | 4.26 | 1.889 | 0.295 | 11 | Bft33.15 |
| Kf33.15 | 120 | 0.255 | | | | | | | 0.353 | 15 | Kf33.15 |
| Knb33.6 | 78 | 0.153 | 119.2 | 3.907 | 2E-20 | 1E-07 | 4.325 | 60.93 | 0.213 | 12 | Knb33.6 |
| Bkp33.6 | 66 | 0.343 | | | | | | | 0.443 | 11 | Bkp33.6 |
| Knb33.6 | 78 | 0.153 | 29.62 | 3.907 | 2E-07 | 0.11 | 4.325 | 2.786 | 0.213 | 12 | Knb33.6 |
| Btf33.6 | 66 | 0.253 | | | | | | | 0.312 | 11 | Btf33.6 |
| Knb33.6 | 78 | 0.153 | 0.033 | 3.882 | 0.855 | 0.124 | 4.210 | 2.527 | 0.213 | 12 | Knb33.6 |
| Kf33.6 | 153 | 0.156 | | | | | | | 0.289 | 17 | Kf33.6 |
| Bkp33.6 | 66 | 0.343 | 17.43 | 3.914 | 5E-05 | 0.041 | 4.351 | 4.78 | 0.443 | 11 | Bkp33.6 |
| Btf33.6 | 66 | 0.253 | | | | | | | 0.312 | 11 | Btf33.6 |
| Bkp33.6 | 66 | 0.343 | 104 | 3.885 | 3E-20 | 0.004 | 4.225 | 9.962 | 0.443 | 11 | Bkp33.6 |
| Kf33.6 | 153 | 0.156 | | | | | | | 0.289 | 17 | Kf33.6 |
| Btf33.6 | 66 | 0.253 | 26.6 | 3.885 | 6E-07 | 0.732 | 4.225 | 0.12 | 0.312 | 11 | Btf33.6 |
| Kf33.6 | 153 | 0.156 | | | | | | | 0.289 | 17 | Kf33.6 |
| 33.6 + 33.15 | | | | | | | | | | | 33.6 + 33.15 |
| Knb | 156 | 0.17 | 162.8 | 3.874 | 8E-30 | 9E-08 | 4.062 | 40.99 | 0.23 | 24 | Knb |
| Bkp | 132 | 0.322 | | | | | | | 0.398 | 22 | Bkp |
| Knb | 156 | 0.17 | 44.05 | 3.874 | 2E-10 | 0.041 | 4.062 | 4.443 | 0.23 | 24 | Knb |
| Btf | 132 | 0.245 | | | | | | | 0.303 | 22 | Btf |
| Knb | 156 | 0.17 | 7.043 | 3.863 | 0.008 | 0.04 | 4.052 | 4.473 | 0.23 | 24 | Knb |
| Kf | 273 | 0.2 | | | | | | | 0.295 | 24 | Kf |
| Bkp | 132 | 0.322 | 31.7 | 3.877 | 5E-08 | 0.014 | 4.073 | 6.579 | 0.398 | 22 | Bkp |
| Btf | 132 | 0.245 | | | | | | | 0.303 | 22 | Btf |
| Bkp | 132 | 0.322 | 89.69 | 3.865 | 2E-19 | 0.022 | 4.027 | 5.543 | 0.398 | 22 | Bkp |
| Kf | 273 | 0.2 | | | | | | | 0.319 | 32 | Kf |
| Btf | 132 | 0.245 | 12.96 | 3.865 | 4E-04 | 0.686 | 4.027 | 0.166 | 0.303 | 22 | Btf |
| Kf | 273 | 0.2 | | | | | | | 0.319 | 32 | Kf |
| n = number of samples | | | | | | F = variance ratio | | | | | |

Table 3·9. Comparison between the two groups of data; original and independent for 33·15, 33·6 and both probes combined.

In the original data, the probabilities of the means being equal are not significant except for Btf/Kf (33·15) and Knb/Kf (33·6). The independent data generally demonstrate a much higher significance for pairs of populations.

### 3·3f The Relationships Between Populations

To examine the relationships between these four populations, dendrograms were constructed in the following way. The mean band sharing coefficients between pairs of populations (Table 3·6) were taken and the average percentage difference (APD) calculated as 100(1- mean band sharing coefficient) (Gilbert *et al* 1990). The data were entered into a matrix and a dendrogram derived based on 'nearest neighbours'. That is; the two locations sharing the lowest APD being the two most closely related. These two (now combined) locations were then compared with the remaining localities and the lowest APD again taken to indicate the nearest neighbour.

Figure 3·13 is the dendrogram constructed from the data for both probes combined.

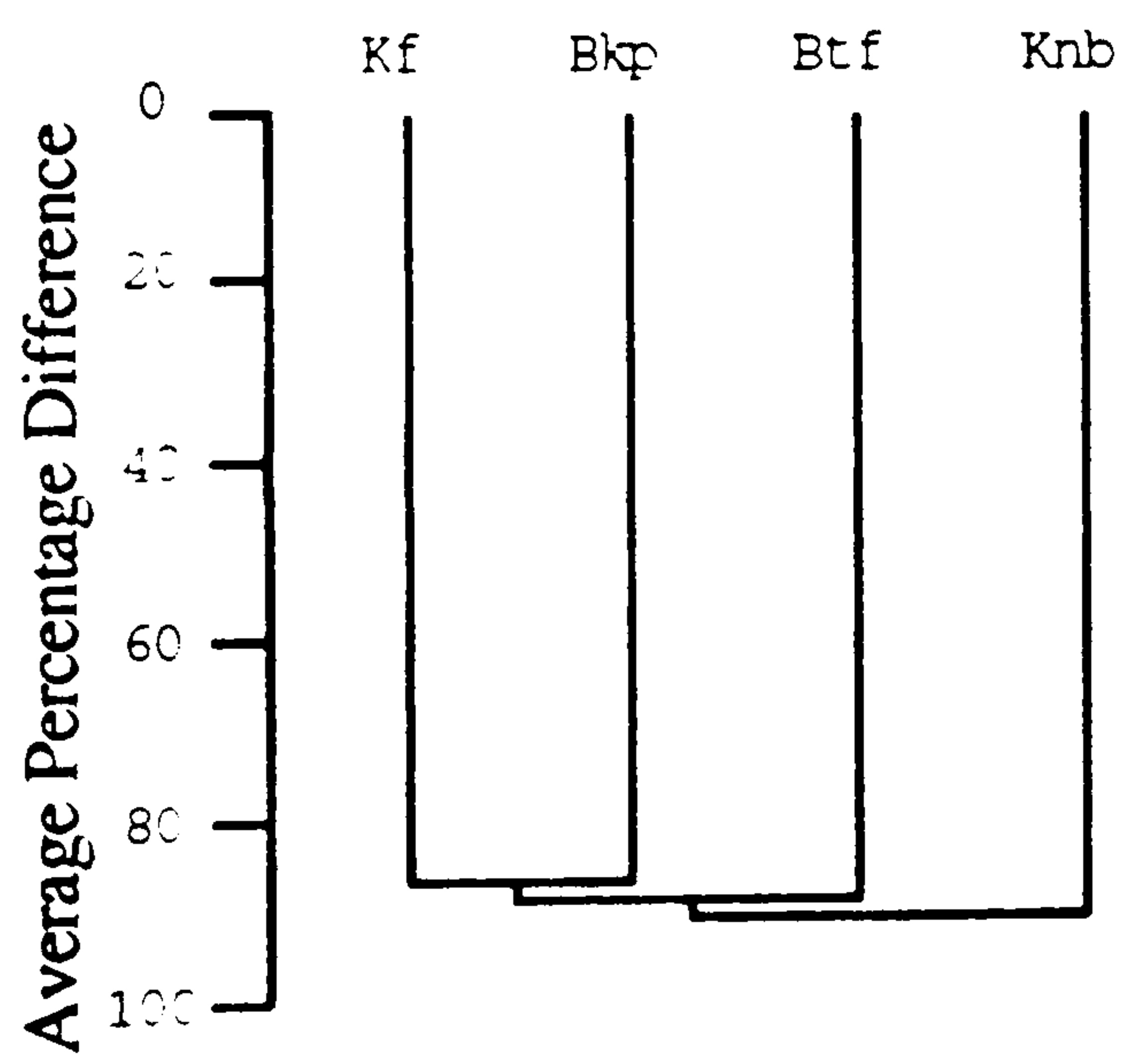


Figure 3·13 Dendrogram constructed from the APD for both probes combined

### 3·3g Population Subdivision

To test further the hypothesis of population subdivision an estimator ($F'_{ST}$), analogous to Wright's (1951) F statistic, was computed for all pairwise comparisons of populations. Measures of gene diversity resulting from single locus analysis are frequently used to measure population subdivision (Nei 1987). However, because explicit estimates of gene frequency are not available in this study, the usual formulae can not be applied. To determine whether there is significantly more variation between two populations than expected from the within-sample similarity, a between-population similarity index corrected by a within-population similarity index can be constructed (Lynch 1991). This index is calculated as:-

$$\bar{S}_{ij} = 1 + \bar{S}'_{ij} - \frac{\bar{S}_i + \bar{S}_j}{2} \qquad (3{\cdot}9)$$

where: $\bar{S}'_{ij}$ = mean similarity between random pairs of individuals from populations i and j

$\bar{S}_i$ = mean similarity in population i

$\bar{S}_j$ = mean similarity in population j

From this equation it can be seen that when $\bar{S}'_{ij}$ equals the mean similarity in the two populations, $\bar{S}_{ij}$ will equal one, indicating that the populations are homogenous. The sampling variance of $\bar{S}_{ij}$ is calculated as per Lynch (1990).
Heterozygosity, or between-population gene diversity, may be calculated from the similarity index (equation 3·1), even though it does not provide an unbiased estimate of population homozygosity. If $\bar{S}'_{ij}$, $\bar{S}_i$ and $\bar{S}_j$ are all biased to the same degree then the biases should cancel out in equation (3·9) and so $1-\bar{S}'_{ij}$ should provide a nearly unbiased, or slightly downwardly biased, estimator of heterozygosity.
Wright's (1951) index of population subdivision ($F_{ST}$) is defined as the fraction of total gene diversity attributable to population differences. Therefore Lynch's (1991) $F_{ST}$, an approximation of $F_{ST}$, is

$$F'_{ST} \approx \frac{1 - S_b}{2 - S_w - S_b} \qquad (3{\cdot}10)$$

Where: $S_b$ = the mean value of $\bar{S}_{ij}$ over all pairs of populations,

$S_w$ = the mean value of $\bar{S}_i$ over all i populations.

The data used in this estimate of population subdivision is the independent data previously used. The results are presented in Table 3·10

33·6

| Pop | Mean BSC | Var | $\bar{S}_{ij}$ | VAR($\bar{S}_{ij}$) | H | $F'_{ST}$ |
|---|---|---|---|---|---|---|
| Knb | 0·22 | 0·0099 | | | | |
| Bkp | 0·44 | 0·0054 | 0·76 | 0·0099 | 0·23 | 0·26 |
| Knb/Bkp | 0·1 | 0·0061 | | | | |
| Knb | 0·22 | 0·0099 | | | | |
| Btf | 0·32 | 0·0505 | 0·86 | 0·0203 | 0·14 | 0·16 |
| Knb/Btf | 0·13 | 0·0052 | | | | |
| Knb | 0·22 | 0·0099 | | | | |
| Kf | 0·26 | 0·018 | 0·84 | 0·0123 | 0·16 | 0·17 |
| Knb/Kf | 0·09 | 0·0054 | | | | |
| Bkp | 0·44 | 0·0054 | | | | |
| Btf | 0·32 | 0·0505 | 0·77 | 0·0170 | 0·23 | 0·27 |
| Bkp/Btf | 0·15 | 0·0031 | | | | |
| Bkp | 0·44 | 0·0054 | | | | |
| Kf | 0·26 | 0·018 | 0·8 | 0·0164 | 0·2 | 0·23 |
| Bkp/Kf | 0·15 | 0·0106 | | | | |
| Btf | 0·32 | 0·0505 | | | | |
| Kf | 0·26 | 0·018 | 0·84 | 0·024 | 0·16 | 0·18 |
| Btf/Kf | 0·13 | 0·0069 | | | | |

33·15

| Pop | Mean BSC | Var | $\bar{S}_{ij}$ | VAR($\bar{S}_{ij}$) | H | $F'_{ST}$ |
|---|---|---|---|---|---|---|
| Knb | 0·26 | 0·0078 | | | | |
| Bkp | 0·34 | 0·0094 | 0·84 | 0·0080 | 0·16 | 0·19 |
| Knb/Bkp | 0·14 | 0·0038 | | | | |
| Knb | 0·26 | 0·0078 | | | | |
| Btf | 0·27 | 0·014 | 0·86 | 0·0102 | 0·14 | 0·16 |
| Knb/Btf | 0·12 | 0·0048 | | | | |
| Knb | 0·26 | 0·0078 | | | | |
| Kf | 0·4 | 0·0141 | 0·79 | 0·0082 | 0·21 | 0·24 |
| Knb/Kf | 0·12 | 0·0027 | | | | |
| Bkp | 0·34 | 0·0094 | | | | |
| Btf | 0·27 | 0·014 | 0·86 | 0·0098 | 0·14 | 0·17 |
| Bkp/Btf | 0·16 | 0·004 | | | | |
| Bkp | 0·38 | 0·0094 | | | | |
| Kf | 0·4 | 0·0141 | 0·76 | 0·0111 | 0·24 | 0·28 |
| Bkp/Kf | 0·12 | 0·0053 | | | | |
| Btf | 0·27 | 0·014 | | | | |
| Kf | 0·4 | 0·0141 | 0·8 | 0·0139 | 0·2 | 0·23 |
| Btf/Kf | 0·13 | 0·0069 | | | | |

Table 3·10 Calculations for Analysis of Population Subdivision (H = heterozygosity)

If the average similarity between populations ($\bar{S}'_{ij}$) is less than the average similarity within populations ($\bar{S}_i$ + $\bar{S}_j/2$) then the value for the similarity index ($\bar{S}_{ij}$) will be less than one - indicating a low degree of genetic similarity between the two populations. The converse holds if $\bar{S}_{ij}$ is equal to, or greater than, one.

Lynch's (1991) $F'_{ST}$ (analogous to Wright's index ($F_{ST}$ ) of among population variance) is a measure of the inbreeding found in a subdivided population and represents the fraction of total gene diversity that is attributable to population differentiation. $F'_{ST}$ takes a value of one when populations are fixed for different alleles (i.e. subdivided) and zero when there is no subdivision. The mean values calculated here, ranging from 0·17 - 0·27 (33·6) and 0·16 - 0·28 (33·15) are similar, confirming that population subdivisions exist; despite both probes showing dissimilarities in earlier measures.

## 3·4 DISCUSSION

The purpose of using DNA fingerprinting in this study has not been the usual one of elucidating family relationships, reproductive success or parentage, but rather to quantify genetic similarity at the level of sub-populations. The minisatellite probes used in this work have revealed - in terms of the probe/restriction enzyme combination used - that a useful number of polymorphic loci exist in muntjac deer. The fundamental measure taken from fingerprints is the similarity index, and whilst this is very useful in defining familial relationships, it is somewhat weaker in population genetic studies. The concordance of the mean band sharing coefficient, as determined by single factor anovar (Table 3·5) reflects to some extent the number of co-detected bands per population (section 3·3b). The population with the greatest incidence of co-detection between 33·6 and 33·15 (Brocket Park) is the population with the greatest probability of the mean bands sharing coefficients being equal. The population with the lowest number of co-detected bands (King's Forest) is the population with the lowest probability of the mean band sharing coefficients being equal.
Bearing in mind that this is a population study, not a study of family groups, the band sharing coefficients must be taken as a indication of the level of similarity within the local populations. The degree of local population similarity will be affected by many factors, the more important of which will be inbreeding, founder effect and bottlenecks. The mean band sharing coefficients found here are broadly similar to those of other outbreeding populations (e.g. Jeffreys & Morton 1987, Amos *et al* 1993a, Keane *et al* 1994), with the exception of Brocket Park, where the mean band sharing coefficient is rather higher than the others (this and other studies) but similar to that found in the (outbreeding) population of red deer (*Cervus elaphus*) on the Isle of Rum, Scotland (Pemberton *et al* 1992).

The traditional measure of the level of population differentiation is the mean homozygosity. Similarity indices are representative of the average identity-in-state rather than the more useful identity-by-descent, for random pairs of individuals. Nevertheless, if a large number of polymorphic loci are examined then the similarity index ($S_{xy}$) may be treated as a normally distributed variable with approximate mean ($\bar{S}$) and approximate variance ($Var(S_{xy})$. As a normally distributed variable it would be possible to use the standard errors to identify populations that were lacking in genetic variation.

This would of course only be valid under the assumption that the degree of variation found in genetic fingerprints closely reflects the degree of variation found in the remainder of the genome. When the between population similarity index ($\bar{S}_{ij}$) is treated as a normally distributed variable, it also becomes possible to test for population subdivisions.

The analysis of variance of the total data showed very large variance ratios and correspondingly low probabilities of occurrence of the means being equal. This may be due to the non-independent and non-normal nature of the data. However the independent data gave lower variance ratios and correspondingly higher probabilities of occurrence of the means being equal. By taking these results diagrammatic representations of the (genetic) distances between populations were constructed.

From the data derived by using probe 33·15, Btf and Kf are closely related whilst Bkp is distantly related and Knb even more so. From the data derived by using probe 33·6, Kf & Bkp are closely related, Btf quite far removed and Knb the most distant. From the combined data Kf & Bkp are closely related whilst Btf is distantly related and Knb the most distant.

At first sight this data may appear to be contradictory but, as only a few animals have been sampled from each population, some conflict may be expected.

From Figure 3·14, a plan of the sample sites centred on Woburn Abbey, it can be seen that three of the sites, Brocket Park, Knebworth and Buntingford are geographically quite close to each other. Brocket Park and Knebworth are separated by 8km and Buntingford is about 15km from Knebworth. King's Forest is 70km from Buntingford and 90km from Brocket Park.

It may be intuitively expected that geographically close locations would be genetically close. However, the incompleteness of the data due to only a small proportion of the total number of samples collected being of sufficient quality to produce fingerprints, may be obscuring some relationships. The population at Minsbury Hill, being in close proximity to Knebworth (10km) and Brocket Park (20km) may have provided genetically dissimilar founders for these populations. As may have the population at Hexton which is only 10km to the south of both localities. Greater genetic differences than were originally present could have arisen by drift or some agent of selection (i.e. the culling of all the (probably closely related) deer in a small area). A significant observation is the degree of similarity between Brocket Park and King's Forest. These two localities are found to be genetically closer than other localities, even though they are geographically the farthest apart.

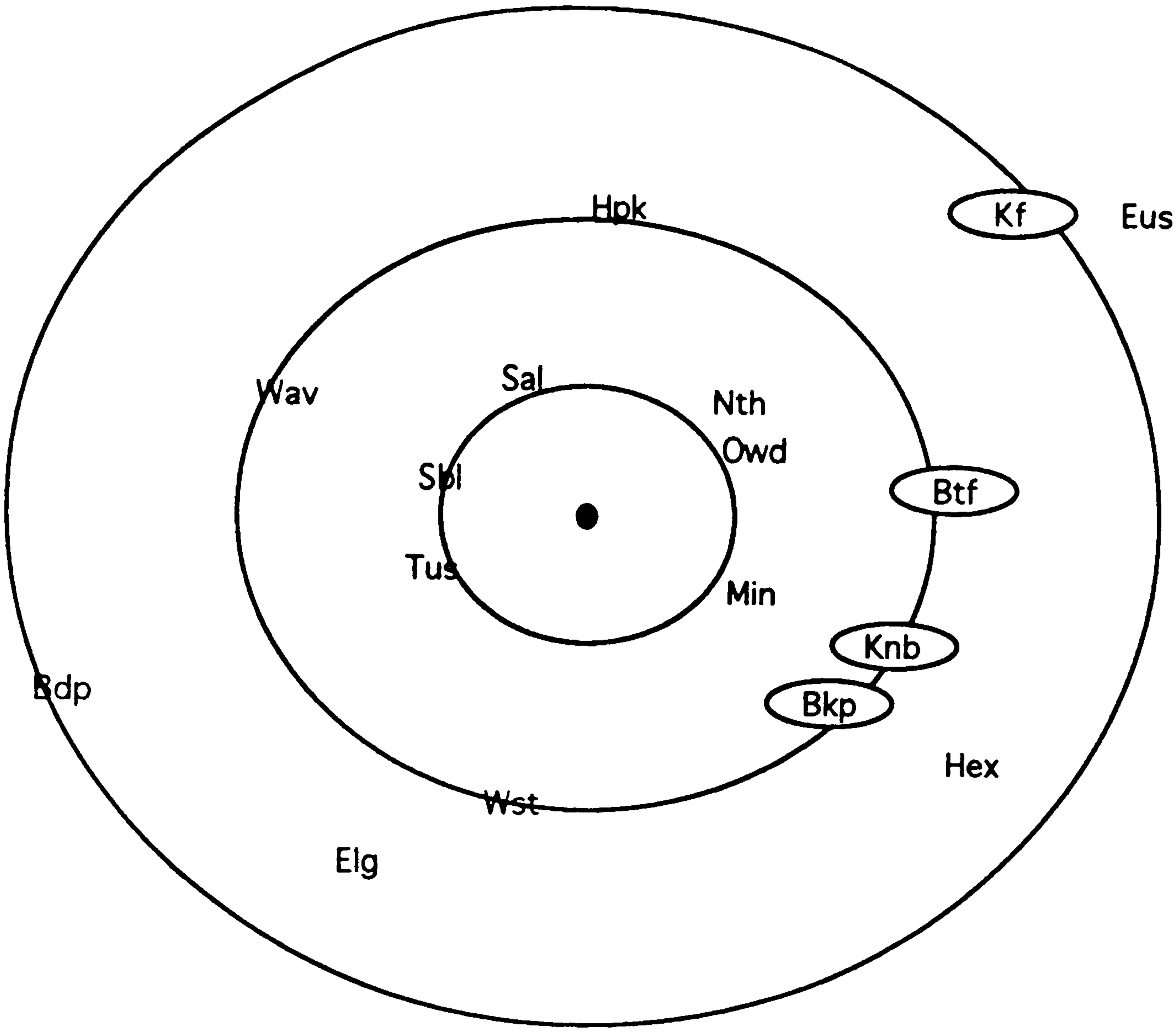


Figure 3·14 Plan of the sample sites, centred on Woburn Abbey. The circled sites are those referred to in this chapter.

Theoretically, the variance in a relatedness class should depend only on the mean similarity index (Lynch, 1988), however the technical limitations of fragment size categorization, especially across several gels severely limits the accuracy of the variance. Ideally the variance should be obtained empirically. To improve the technique - in terms of band matching - individuals from the two populations to be compared should be run in alternate lanes on the same gel. This would have the disadvantage of becoming very time consuming in a large group of populations. However, a great advantage would be gained from deriving band sharing estimates from animals of known genealogy; comparisons of band sharing estimates between individuals of unknown background could then be accurately interpreted. Bearing in mind the above caveats, DNA fingerprinting has revealed a degree of similarity among the populations investigated but, as previously noted, the quality of the majority of samples was too poor to produce interpretable fingerprints. For this reason only four sites were available, Brocket Park, Knebworth, Buntingford and King's Forest. For the King's Forest population to be closely related to the

Brocket Park population, (geographically most separated localities) yet for Knebworth and Brocket Park to be distantly related (geographically closest localities) there must be some movement of deer between the former group and less between the latter. It is probable that there would only be a very remote chance of these relationships developing if the populations of deer found at these localities were a consequence of natural spread and colonisation. The only conclusion to be drawn from these data is that either (a) deer have been taken from Brocket Park and deposited in King's Forest or, (b) Brocket Park is being restocked from King's Forest. Knebworth, being more-or-less equally (genetically) distantly related to the three other sites, probably does not have any immigration or emigration of deer. The limited data derived here support the idea of human mediated migration.

Inter- and intra-population comparisons of restriction fragments indicate that the level of genetic differentiation between populations is higher than within populations, the average percentage difference (1 - similarity index) within populations ranges from 66% (Bkp) to 84% (Knb) compared with a range of 82·5% (Btf/Kf) to 91% (Knb/Kf) for between populations. Further analysis of the data reveals a low heterozygosity for all population comparisons (range 0·14 to 0·24) and $F'_{ST}$ values for pairwise comparisons ranging from 0·16 to 0·28. Collectively these data show that there is a measurable amount of sub-division between populations.

This, taken with the results from the analysis of variance, demonstrates that the populations are identifiably distinct units - 20% of (fingerprint band) diversity is attributable to population differences and yet there is some migration between populations.

The paucity of data is a great hindrance and prohibits any observations about gene flow, migration and expansion amongst the localities sampled. And, from a study such as this it is not possible to put these migrations in an historical context or to determine the pattern of migration. Nevertheless, this investigation has demonstrated that multilocus DNA fingerprinting can be applicable to investigating the similarity of localities within a subdivided population.

Chapter 4

# mtDNA RESTRICTION ANALYSIS OF GEOGRAPHICALLY SEPARATED GROUPS OF FERAL MUNTJAC DEER

## 4·1 Introduction

Mitochondrial DNA (mtDNA) is a single, duplex, covalently closed circular DNA molecule ranging from 15·7 kb in sea urchins (*Strongylocentrotus* sp. and *Lytechinus* sp.) (Brown, 1981) and fruitflies (*Drosophila neohydie*) (Fauron & Wolstenholme, 1976), to 19·5 kb in *Drosophila melanogaster* (Fauron & Wolstenholme, 1976). The diploid genome of multicellular animals varies greatly in size, ranging from $4 \times 10^8$ to $4 \times 10^{11}$ base pairs (bp) subdivided into from 4 to over 250 discrete chromosomes (Brown 1983). By comparison, the animal mitochondrial genome is in the order of $4 \times 10^{-5}$ times smaller than the smallest animal nuclear genome, and varies in size by a factor of three, compared with a factor of $10^3$ in the nuclear genome (Brown, 1985; Gjetvaj *et al* 1992). Complete sequences of the mtDNA have been determined for a number of vertebrates, including human (Anderson *et al* 1981), mouse (Bibb *et al* 1981), cattle (Anderson *et al* 1982), *Xenopus* (Rowe *et al* 1985), dolphin (Southern *et al* 1988), rat (Gadelata *et al* 1989), chicken (Desjardins & Morais, 1990), fin whale (Arnason *et al* 1991) and the harbour seal (Arnason & Johnsson, 1992). Mitochondrial gene order has been found to be highly conserved (Wallace, 1982; Sederoff, 1984). Mammalian mitochondrial DNA has been sequenced from three different orders; rodent - house mouse (Bibb *et al* 1981), primate - man, and artiodactyl - cattle (Anderson, *et al* 1981, 1982) and the gene order found to be identical, indicating a stability that as been maintained since these lineages diverged some 80 to 100 million years ago. The only exception so far noted is the non-mammalian one found in gallinaceous birds, where the $tRNA^{Glu}$ and the $ND_6$ genes are located immediately adjacent to the control region, instead of between the cytochrome b and $ND_5$ genes (Desjardins & Morais, 1991).

### 4·1a Organization of the mtDNA Molecule

The molecule consists of two strands, known as the heavy (H) strand and the light (L) strand. This nomenclature derives from the pronounced bias in the G+T composition which causes a difference in the buoyant densities of the strands in CsCl gradients. There is a strict asymmetry in the strand distribution of the structural genes, with only one protein gene (ND6) and eight tRNA genes encoded by the light (L) strand (Wallace, 1982).

The mammalian mitochondrial genome, being a non-recombining molecule, seems to be lacking in introns as well as being less susceptible to sequence re-arrangements, which are commonly found in the more complex and relatively fluid sequence organization of the nuclear genome. The genome contains genes encoding 13 proteins including cytochrome b, cytochrome c oxidase subunits l, ll and lll, ATPase subunit 6; two ribosomal RNAs (12S and 16S rRNA), 22 transfer RNAs (tRNAs), and a non-coding region known as the control region (Avise & Lansman 1983; Brown, 1985; Clayton, 1982 and 1984) which is functionally similar but structurally different between major animal groups (Brown, 1985). No variation is known in the gene content of mammalian mtDNA (Vawter & Brown, 1986), the sequences of which occur only once per mitochondrial genome (Brown, 1985). Among cattle, primate and mouse sequences, conservation of the first and second positions is higher than that of the third position, reflecting both the degeneracy of the genetic code between lineages, and the higher rate of silent substitutions (Aquadro *et al* 1984). Comparison of nucleotide substitutions between Japanese quail and chicken protein genes also found that most substitutions occurred at the third codon position, so amino acid replacement is low (Desjardins & Morais, 1991). Codon usage in mtDNA genes is nonrandom, with a strong bias for codons ending in A or T in *Drosophila* (Clary & Wolstenholme, 1985), and a strong bias against the use of codons ending in G (Desjardins & Morais, 1991).

Mitochondrial DNA, being much simpler than nuclear DNA, is therefore not prone to producing the many interpretational ambiguities and complexities that nuclear DNA produces, and its evolution appears to proceed in a much more straight-forward manner than that of nuclear DNA.

Although microvariation in the size of the mitochondrial genome has been found to occur between conspecific individuals (Crews *et al*, 1979; Upholt & Dawid, 1977; Aquadro & Greenberg, 1982; Greenberg *et al*, 1983 ), the gross size appears to be relatively stable within species and over wider taxonomic groupings. There is no correlation between genome size and taxonomic group (Brown, 1983). Rare exceptions to the above can be found within the *Drosophila* species complex where the size of the mitochondria in *D. neohydei* is 15·7 kb, that of *D. melanogaster* (Japan) is 18·7 kb and D. *melanogaster* (ORE-R) is 19·5 kb; and lizards from one population of *Cnemidophorus sexliniatus* have a mtDNA genome that is 1·2 kb larger than individuals from six other populations (Brown & Wright, 1979; Brown, 1981).

#### 4·1b Replication and Size Variation

The control region, which lies between the genes for proline and phenylalanine tRNAs, contains a triple-stranded displacement loop (D-loop) which functions as a site for the origin of replication and RNA transcription of the H and L-strands (Sederoff, 1984; Chang & Clayton, 1985; Mayhook *et al* 1992). The origin of H-strand replication is situated in the midpoint of the D-loop, although there are no obvious primary sequences or secondary structure elements that are essential for this function. Priming for L-strand replication occurs in a T-rich loop of potential secondary structure, although this origin of replication only becomes functional when the template becomes single-stranded by displacement when H-strand synthesis is 67% complete (Sederoff, 1984; Chang *et al* 1985). In mammalian mtDNA replication the initial event must therefore be initiation of H-strand synthesis (Chang & Clayton, 1985). Synthesis on both strands is continuous, highly asymmetric and initially unidirectional (Clayton, 1982). When strand synthesis is complete the molecule closes, becomes supercoiled and then forms a relaxed superhelix as a consequence of D-loop formation. Replication pause sites have been reported (Mayhook *et al* 1992) which may promote initiation of DNA synthesis on the lagging strand, or have a role in transcriptional control. Although the origin of L-strand replication is conserved in mammalian and amphibian sequences, it appears to be absent from Japanese quail and chicken mtDNAs. L-strand replication in these species may be initiated in an equivalent region in another tRNA gene (Desjardins & Morais, 1991). DNA synthesis is terminated by stop points, which have been limited to a single nucleotide position on the template strand in humans (Doda *et al* 1981).

With so few genes, the coding capacity of mtDNA is severely limited, most of the protein components involved in replication, transcription and translation are encoded for by the nucleus (Chang *et al* 1985). Consequently, mitochondrial metabolism relies on close and co-ordinated interaction between nuclear genes and mtDNA (Tzagoloff & Myers, 1986; Moritz *et al* 1987). This situation is seen by some as supporting the hypothesis that mitochondria are derived from a more complex genome by the transfer of genetic functions to the nucleus, in response to intense selection for small genome size (Wallace, 1982). The selective advantages of a reduction in size and complexity may be faster replication (Rand & Harrison, 1986) and greater efficiency through dedication.

The occurrence of size variations can be taken as unequivocal evidence for the occurrence of deletions and additions during replication. Macrovariation in the size of mtDNA appears to result from large deletions or additions confined exclusively to the D-loop and adjacent non-transcribed regions (Brown, 1983). Evidence for this comes from high resolution electron microscopy of heteroduplex molecules which showed that large deletion or addition events are clustered near the ends of the D-loop (Upholt & Dawid, 1977). Microvariation in mtDNA size results from deletions and additions of one to a few base pairs (Crews *et al*, 1979; Brown *et al* 1982; Greenberg *et al* , 1983; Aquadro & Greenberg, 1982) in all regions of the mtDNA although the frequency of occurrence is variable, and low in genes that code for proteins (Brown, 1983). Because the origins of replication and overall structure are so precisely known, mtDNA provides a very good system in which to investigate the evolutionary genetics of this class of DNA sequence.

### 4·1c Evolutionary Rate of Change

Although the mitochondrial gene order is highly conserved, the rate of evolution of the genome has been estimated to be 5-10 times faster than single-copy nuclear DNA in vertebrates (Brown *et al* 1979; Vawter & Brown, 1986), although rates vary in different parts of the mitochondrial genome (Hoelzel *et al* 1991). This may not be applicable to all taxa, for example, both *Drosophila* sp. and sea urchins (*Strongylocentrotus* sp.) have been found to show a similar rate of evolution of both mitochondrial and single copy nuclear DNA (Brown, 1985; Vawter &

Brown, 1986). Britten, (1986) has attributed this to variation in the rate of evolution of the nuclear DNA rather than a change in rate of the mtDNA.

Aquadro *et al* (1984) found that only a small portion of the genome was evolving rapidly, and that a significant fraction of the molecule was under strong selective constraints. The evolutionary rate of the unconserved regions is due to an unusually high mutation rate (Brown *et al* 1979). The factors responsible for a higher mutation rate may include: (a) greater exposure to oxidation radical damage, as mtDNA has been reported to accumulate 16 times more oxidative damage than nuclear DNA (Wallace, 1992). Although this may be true, it should be remembered that free radicals are the normal enzymatically bound metabolic intermediates in the highly efficient and controlled process of oxidative phosphorylation and the possibility of their being free to undergo random reactions must be quite small. (b) An inefficient system of replication (Brown *et al* 1982), mammalian mitochondrial DNA is replicated by $\gamma$-polymerase, the only polymerase present in the mitochondria and an enzyme which both lacks the ability to edit and repair (Hübscher *et al* 1979) and is prone to a high rate of nucleotide misincorporation in comparison to the nuclear genome (Brown *et al* 1979, 1982, Clayton 1984, Wilson *et al* 1985). (c) A higher rate of turnover (Brown *et al* 1982). Transitions and length variants are the most common form of replication error (Topal & Fresco, 1976), which leads to a bias toward transitions that is most evident in close genealogical comparisons. With increasing divergence times, highly variable sites would receive many transitions, a number of which would be back mutations (Hixson & Brown, 1986) and so the bias towards transitions would become reduced.

Nuclear DNA has a generation time identical to that of the cell type in which it occurs, whereas mitochondrial DNA replicates autonomously, and its generation time is much shorter due to the continual replacement of defective and damaged mitochondria (Wiesner, 1992). If mtDNA undergoes more rounds of replication than nuclear DNA, this may increase the number of errors produced per cell generation (Brown, 1985). A bias in the mutation process causes a high predominance of transitions over transversions. The predominance of transitions is greatest in comparisons of more closely related sequences (Aquadro & Greenberg, 1982), and decreases as sequences diverge, until no bias can be detected

(Brown *et al* 1982). The decrease in the proportion of transitions is due to a highly biased substitution process (Aquadro *et al* 1984). This implies that convergent events, parallelism, and character state reversals increase in occurrence between mtDNA of more divergent taxa (Berg & Ferris, 1984).

Several studies have suggested that the most rapidly evolving region of the mitochondrial genome is the non-coding control region, which contains the D-loop (Upholt & Dawid, 1977; Walberg & Clayton, 1981; Chang & Clayton, 1985). Among animal species the size of this region varies between ~200 to 4,100bp (Brown 1985), and the substitution rate in this region in humans has been estimated to be from 2·8 (Cann *et al* 1984) to 5 (Aquadro & Greenberg, 1982) times the rate found in the remainder of the mitochondrial genome. Contrary to the above though, the rates determined from interspecific D-loop sequence comparisons for primates - 0·5% - 1·0%/Myr (Foran *et al* 1988), rodents - 0·5% - 1·0% (Brown *et at* 1986) and cetaceans - 0·5% (Hoelzel *et al* 1991) suggest little difference in comparison with the remainder of the mammalian mitochondrial genome. More evidence that the preceding substitution rates may not be relevant to all species comes from Heolzel *et al* (1991), who found the substitution rate in the cetacean D-loop region to be an order of magnitude lower than that suggested for human D-loop region, yet the remainder of the cetacean mitochondria genome had a substitution rate no higher than that estimated (Southern *et al* 1988).
Taking the mitochondrial genome in its totality, the range of substitution rates is 0·1% - 2·0% change per million years in animals: *Drosophila* 0·6% - 1·8%; Echinoids 0·13% - 1·1%; primates 0·5% - 1·0%; other mammals 2·0% (Moritz *et al* 1987, Table 1; Wilson *et al* 1985). These figures are notably higher than the substitution rate for single-copy nuclear DNA in vertebrates, for example 0·15% - 0·25% (Vawter & Brown, 1986).

**4·1d Mode of Inheritance**
In higher animals, mtDNA appears to be maternally inherited by transmission to the progeny through the egg cytoplasm (Avise *et al* 1979; Avise & Lansman, 1983; Gyllensten *et al* 1987). This statement is supported by several lines of enquiry. In an examination of the progeny of crosses between parents differing in their mtDNA restriction fragment patterns, it is the maternal line that is seen to be inherited. For example, donkey (*Equus onager*) and horse (*E. caballus*) differ in the position of the *Hae* lll

restriction sites (Avise & Lansman, 1983). Mules (the progeny of a cross between a female horse and an male donkey) exhibit the horse mtDNA type whilst hinnes (the reciprocal cross) exhibit the donkey mtDNA type. Similar evidence for maternal inheritance has been reported from experimental crosses of *Peromyscus* (Avise *et al* 1979), *Rattus* (Franscisco *et al* 1979; Hayashi *et al* 1978), *Xenopus* (Dawid & Blackler, 1972), humans (Giles *et al* 1980).

These early results are to be expected from the microscopical observations that sperm carry a very low number of mitochondria - in the order of 50 per sperm, while eggs may have in the order of $10^5$; although this ratio varies from species to species, for example the exceptionally large *Xenopus* egg is estimated to contain $10^8$ mtDNA molecules whilst the sperm may carry as few as 100 (Dawid & Blacker 1972). In sperm, the mitochondria are to be found concentrated in the energy-using region of the tail and, as the head of the sperm penetrates the body of the egg and makes its way towards the egg nucleus, (generally) the tail becomes separated and remains in the ooplasm. The fate of the contents of the tail region remain unknown, in some species the mitochondria are seen to disperse into the zygote cytoplasm (Gresson 1940; Friedlander 1980) and in other species the sperm mitochondria are reported to not penetrate the egg (Ursprung & Schabtach 1965).

If however, paternal mtDNA was being transmitted then, if we assume that the full content of one sperm were to be mixed with the full content of one egg, the species of mtDNA would be evident in the offspring in the ration of 1:2000, well below the conventional level of detection unless the paternally transmitted mtDNA bore some unique characteristics.
In insects, sperm mitochondria have been observed within the zygote (Friedlander, 1980; Chapman *et al* 1982), although their fate remains unknown. So, in an attempt to detect low levels of paternally transmitted mtDNA Lansman *et al* (1983) formed backcrosses for 91 consecutive generations between the fertile female hybrid progeny from two species of *Heliothis* (tobacco budworm) and males of one of the species. These species differed in their *Xba* 1 and *Eco* RI restriction profiles. If 'paternal leakage' of mtDNA was occurring then after 91 backcross generations enough paternally inherited mtDNA would have accumulated to become

detectable by autoradiographic assay. However, absolutely no paternally inherited mtDNA was detected.

Recent studies have demonstrated that paternal inheritance of mtDNA does in fact occur, albeit at a very low level, in mice and *Drosophila* by multiple backcrossing to a paternal strain (Kondo *et al* 1990; Gyllensten *et al* 1991). In mice the paternal contribution was estimated at $10^{-4}$ (i.e. 1:10 000). A more recent study by Zouros *et al* (1992) has reported a conservative estimate of 0·1 for the paternal contribution in two species of *Mytilus* from wild populations.

The concept of paternal leakage could be very important, because even low levels of leakage may have important evolutionary consequences for mitochondrial gene flow between female lineages of mtDNA that would otherwise be isolated from one another. As yet only a few species have been investigated and a much wider search is needed before any concrete statements can be made about the existence or not of paternally inherited mtDNA and the evolutionary consequences it would have. At the present state of our knowledge the contribution of paternal mtDNA is so small that its effect would be very difficult to discern. With the exception of the *Mytilus* report, paternal inheritance has only been demonstrated in artificial breeding systems and therefore, for the purposes of this study I will assume strict maternal inheritance.

Taking strict maternal inheritance to be the case, mtDNA genotypes will be clonally transmitted from the female to the progeny without recombination. The mtDNA of an individual will therefore provide definitive information about the female lineage to which it belongs (Avise & Lansman, 1983). Uniparental transmission means that short term mtDNA dynamics will be strongly influenced by population-shaping processes such as genetic drift (Chapman *et al* 1982).

**4·1e Heteroplasmy**

Heteroplasmy - the occurrence of two or more mitochondrial DNA genotypes within the same individual - has been observed in a number of species including *Drosophila* (Solignac *et al* 1984, 1986), scallop (Snyder *et al* 1987; Gjetvaj *et al* 1992), various fish (Beckwitt & Petruska, 1985; Bentzen *et al* 1988; Buroker *et al* 1990), lizards (Moritz & Brown, 1987), galliform birds (Desjardins & Morais, 1991), lagomorphs (Mignotte *et al* 1990; Biju-Duval *et al* 1991), evening bats (Wilkinson & Chapman, 1990), the harbour seal (Arnason & Johnsson, 1992) and humans (Wrischnik *et al* 1987). Recent studies of invertebrate and poikilotherm vertebrate DNAs have found considerable size variation, with the length of mtDNA varying by a factor of three (14·2 kb-43 kb) (Gjetvaj *et al* 1992). In the majority of cases, the length polymorphism was shown to be located within the D-loop; the most rapidly evolving region of the mtDNA (Buroker *et al* 1990). In two species of scallop (*Pecten*), the length polymorphisms were found to be in several different regions of the genome (Gjetvaj *et al* 1992).

There could be many reasons for the occurrence of heteroplasmy, but it is most likely to be the result of errors during replication (Moritz & Brown, 1987). A partial displacement of the D-loop strand during replication could easily lead to the formation of stable repeat units by slipped mispairing (Mignotte *et al* 1990). Length variation occurring within the D-loop may be a result of dynamic competitive equilibrium between the D-loop strand and the heavy (H) strand for base pairing with the light (L) strand (Buroker *et al* 1990). Wrischnik *et al* (1987) has found that heteroplasmy in humans is caused by the loss of one of two copies of a 9bp sequence during the evolution of modern types of mtDNA. In cattle, fixation of mtDNA sequence variants appeared to occur within the space of one generation, with no detection of a heteroplasmic intermediate. Koehler *et al* (1991) has postulated that this may be the result of a very high mutation frequency, or a gene conversion mechanism which rapidly spreads the variant to other molecules. Systematic heteroplasmy, as observed in many lagomorphs, may be the result of a high rate of molecular rearrangement localised in the non-coding region of the mitochondrial genome (Biju-Duval *et al* 1991). Heteroplasmy may of course be caused by paternal leakage of mtDNA; high levels of heteroplasmy in two species of mussel being found due to paternal mtDNA contributions (Zouros *et al* 1992).

Heteroplasmy is generally considered to be a transient stage of evolution in the mitochondrial genome, after the mutation of a female germ line cell (Solignac *et al* 1984). In some organisms, the occurrence of heteroplasmy has been observed in the cellular lineage giving rise to a tissue or organ (Biju-Duval *et al* 1991). In order for the mutation to become fixed in a population, the variant must segregate from the majority of parental mtDNA molecules to become the sole mitochondrial genome in the germ cells of successive generations (Koehler *et al* 1991). In general, populations derived from heteroplasmic individuals will arise as a result of weak selection pressure, genetic drift or a founding event, which will result in predominantly mutant or normal mtDNAs (homoplasy) (Solignac *et al* 1984; Wallace, 1992).

**4·1f Evidence for Selection**

Mitochondrial genes code for tRNA, rRNA as well as proteins that are involved in energy production or that are sub-units of enzymes for which the other sub-units are encoded for by the nucleus. Therefore, some mtDNA mutations must be highly visible to selection (Avise *et al* 1987), and others provide the potential for strong cytonuclear associations (Brown, 1985; Singh & Hale, 1990). Although Forbes & Allendorf (1991a) think there appears to be little evidence for cytonuclear associations.

Evidence for selection acting on mtDNA has been reported in crickets (*Gryllus* sp.) by Rand & Harrison (1986). However, the process of selection has been attributed to different sizes of mtDNA having a different selective advantage which would act at the intracellular level, rather than the level of the organism (Singh & Hale, 1990). Further evidence for selection acting on mtDNA has been reported in *Drosophila* by MacRae & Anderson (1988), yet it seems that the selection reported here is not selection on the mtDNA but selection due to a mating incompatibility system between the two species of *Drosophila* under investigation (Nigro & Prout, 1990; Singh & Hale, 1990).

A restriction enzyme study of 16 species of land snails (*Partula* sp.) from the Society Islands suggested that common genotypes in these populations may represent selectively favoured stable states, with individual restriction sites departing from and returning to the norm (Murray *et al* 1991). The action of positive selection on an advantageous mtDNA type has been suggested as an explanation for the apparent rapid branching of the Caucasian mtDNA tree. However, this explanation was considered unlikely, as the rapid branching only occurred in non-African populations, and neutrality tests have shown control region variation in these populations to be neutral (DiRienzo & Wilson, 1991). A recent study of *Salmo clarkii*, (the cutthroat trout) found that mtDNA was effectively neutral (Forbes & Allendorf, 1991b).

There is very little concrete evidence of selection acting on the fully functional mitochondrial DNA molecule. The only definite evidence of selection is found in defective mtDNA, for example in some inherited diseases where one sequence variant is partially functional or non-functional (Forbes & Allendorf, 1991b, Koehler *et al* 1991). In myoclonic epilepsy and ragged red fibre disease (MERRF), within a family, maternal relatives sequentially aquire different symptoms of the disease, the severity of these increasing with the percent of mutant DNAs and age (Wallace, 1992).

### 4·1g The Use of mtDNA in Population Studies

Total mtDNA has been widely used as a tool for measuring inter- and intraspecific divergence in a number of species. For example, the pocket gopher (Avise *et al* 1979), tree frogs (Lamb & Avise, 1986), deer (Miyamoto *et al* 1990), kangaroo rats (Thomas *et al* 1990), gray wolves and coyotes (Lehman *et al* 1991), killer whales (Hoelzel & Dover, 1991), the channel island fox (Wayne *et al* 1991b), humans (Merriwether *et al* 1991; Pesole *et al* 1992) and cichlid fish (Sturmbauer & Meyer 1992).

The most widely used technique has been that of extracting whole mtDNA and subjecting it to cleavage with a number of restriction enzymes, usually enzymes with a four, five or six base pair recognition site. The resulting fragments are separated by gel electrophoresis and visualised by ethidium bromide staining or autoradiography, depending on the amount of mtDNA available. Differences in restriction patterns due to loss or gain of a restriction site or length mutation are known as restriction fragment length polymorphisms (RFLPs) (Upholt 1977). Six base pair recognition site restriction enzymes ('6-cutters') typically produce 1-8 fragments whilst four base pair recognition site restriction enzymes ('4-cutters') typically produce 20 or more fragments (Avise & Lansman 1983). The advantages of the increased data produced by the 4-cutters are sometimes offset by the inherent difficulties in scoring the gels and determining fragment homology (Avise & Lansman 1983).

The genetic similarity between mtDNAs from different organisms can be assessed as the proportion of fragments shared in their digestion profiles. Estimations of the degree of sequence divergence can be improved by restriction mapping the mtDNA molecule. If mtDNA is digested with a mixture of two restriction enzymes and the resulting banding pattern compared with the individual patterns from each enzyme, the relative location of restriction sites can be deduced (Brown & Vinograd, 1974; Wilson *et al* 1985). Phylogenetic inferences can then be made from the restriction site data (Nei & Li, 1979; Templeton, 1983; Li, 1986). The first interspecific restriction mapping comparisons of mtDNA were performed on mtDNA from six hominoid species (Brown *et al* 1979). Comparison of the maps produced an estimation of sequence divergence and thus allowed the phylogenetic relationship to be calculated between the six species. The rate of subsitutions was found to increase rapidly up to 15% sequence divergence, after which the rate rapidly reduced. During the initial rapid phase, substitutions occur in nonessential sequences, such as the control region. The slowing of the rate occurs when those nucleotide positions under relatively weak functional constraint have been substituted (Moritz *et al* 1987).

The estimation of phylogenetic relationships by mtDNA restriction mapping has become a valuable procedure in historical zoogeographic reconstruction and the estimation of divergence times between different groups (Bonne-Tamir *et al* 1986). Recent studies have shown that present-day human mtDNAs had an African origin, with the Asian and Caucasian mtDNA lines diverging soon after the African founding event (Cann *et al* 1987; Merriwether, *et al* 1991). Furthermore, the evidence of Horai *et al* (1986) has shown that founding human populations were considerably polymorphic in terms of the mitochondrial genome before the divergence of the three major races. Data from these studies suggest that multiple founder events have occurred in human history, including the founding of the Americas, Australia, and of course many island populations.
These multiple founding events have resulted in a departure from equilibrium in human populations, with regard to the gain of variation by mutation, and the stochastic loss of variation by drift. These findings are supported by results from a detailed restriction enzyme study of two populations of native Americans, which showed that these populations

had been founded by two independent migrations from Asia (Torroni *et al* 1992).

Major factors influencing the geographic distribution of mtDNA haplotypes are those of the historical pattern of gene flow in terms of the dispersal capabilities and opportunities of the species, and geographic impediments to movement (Bermingham & Avise 1986; Avise *et al* 1987). Evidence has been found in some populations for geographic groupings of mtDNA haplotypes, suggesting that a local structuring of mtDNA exists (Avise *et al* 1979; Hale & Singh 1987; Murray *et al* 1991), whereas in other, more widely distributed species, there appears to be much less geographic partitioning (Tegelstrom 1987). Highly mobile species such as *Megaptera novaengliae* (humpback whale) show distinct partitioning of mtDNA types between populations that is no doubt due to a strong maternal tradition in migratory distribution (Baker *et al* 1990). Locally co-existing populations of killer whales (*Orcinus orca*) have been found to partition into reproductively isolated sub-populations (Hoelzel & Dover 1991). An observation consistent with earlier speculations on the existence of reproductive isolation based on colour morphology (Evans *et al* 1982; Baird & Stacey 1988), acoustical behaviour (Ford & Fisher 1982; Hoelzel & Osbourn 1986) and foraging strategy (Berzin & Vladimirov 1983; Bigg *et al* 1990). Observations such as these demonstrate the importance of considering behavioural strategies as well as dispersal opportunities when investigating both the amount of, and pattern of, phylogenetic differentiation among mtDNA haplotypes (Lansman *et al* 1981), and the macro and micro-geographic distributions of mtDNA phylogenetic groupings (Kessler & Avise, 1985; Avise, 1986, 1989).

### 4·1h mtDNA Analysis Techniques

The majority of mtDNA analysis techniques have employed restriction endonuclease digestion to produce restriction fragment length maps as a basis for the comparison of genetic variation between individuals at many levels in a taxonomic hierarchy. Restriction fragment length polymorphisms (RFLPs) have been successfully used in many studies ranging from an investigation of evolution within the genus *Equus* (George & Ryder 1986), phylogeny and evolution within the Cervidae (Miyamoto *et al* 1990), the origins of tetrapods (Meyer & Wilson 1990), to confirmation of directional hybridisation in tree frogs (Lamb & Avise 1986). In at least two cases RFLPs from the mitochondrial genome have been found to be sufficiently variable to permit individual diagnostic genetic 'fingerprints' by conventional restriction digest analysis (Avise *et al* 1979).

Despite the many successful applications of mtDNA there are still a number of limitations that need to be born in mind when undertaking a mtDNA investigation.

- RFLPs will only detect a small sample of the possible variation between DNA sequences.
- Fragment length comparisons tend to be unreliable if there is a small length variation or if the overall sequence divergence exceeds 10 - 15% (Moritz *et al* 1987; Kocher *et al* 1989; Wilson *et al* 1989).
- The effects of nucleotide sampling can have a considerable effect on the measurement of nucleotide diversity (Lynch & Crease 1990).
- Restriction sites provide a biased sample of base sequences (Kocher *et al* 1989).
- Absence of standardisation in the restriction enzymes used, gel separation techniques and visualisation among laboratories can make comparison of results from different laboratories unreliable.
- RFLP analysis usually requires intact mtDNA to be extracted from tissue with a relatively high ratio of mtDNA to nuclear DNA. This ratio is important because small quantities of nuclear DNA contamination will reduce and may even eliminate resolution of mtDNA fragments. Tissues such as oocytes (in animals that spawn), gonadal tissue, liver, brain and placentae are the common choice.
- Acquisition of these tissues requires the animal to be killed. However the recent development of DNA amplification by the polymerase chain reaction (Mullus & Faloona 1987), and subsequent refinements, now permit DNA amplification from minute tissue samples such as single hair (Vigilant *et al* 1989).

The finest resolution of genetic variation is obtained by sequencing, and this has been used to assess the accuracy of sequence divergence estimates from restriction map estimates of the total mitochondrial genome (Wilson *et al* 1985). Sequence analysis of representative individuals from various populations has both confirmed initial theories and given further insights into evolutionary histories. For example, the demographic history of human populations (DiRienzo & Wilson 1991; Ballinger *et al*

1992). The association between six hominid species was first clarified by RFLP analysis (Brown *et al* 1979), and later confirmed by sequencing data (Brown *et al* 1982). Furthermore, sequence information is not subject to some of the ambiguities that can arise in restriction analysis. In the $ND_4$ gene a single transition from T to C was found to cause a conformational change which enhanced the electrophoretic mobility of the restriction fragment corresponding to that region (Vigilant *et al* 1989).

### 4·1i mtDNA Regions Commonly Used in Investigations

Some regions of the mtDNA genome evolve at a greater rate than others, as previously noted. It follows from this that measuring the genetic variation found in short regions of the mitochondrial genome could be potentially useful in many areas of investigation. The choice of which variable region to investigate will depend on the type of population or evolutionary study being undertaken. Brown (1981) has shown that each of the thirteen protein coding genes has its own characteristic rate of evolution, and that the rate of change amongst these genes seems to be the same among vertebrates. The control region has been widely reported as being the most variable region of mtDNA as, in distinction to the remainder of the mitochondrial genome, it contains sequences that (as yet) appear to be 'functionally nonessential' (Brown 1985).

Due to their function in the respiratory pathway, most protein-coding genes must be highly conserved. mtDNA contains seven non-overlapping open reading frames (ORFs) that in total represent about 60% of the informational content of animal mtDNA (Attardi 1985). These regions have been assigned as components of the nicotinamide adenine dinucleotide dehydrogenase complex in the respiratory assembly - an integral part of the inner mitochondrial membrane. They are known as ND genes (Burger 1985). Their function is to create a hydrophilic environment within the mitochondrial inner membrane, so their structure may be less constrained than if they were subunits of a catalytic enzyme (Chomyn *et al* 1985). With the exception of $ND_1$, the ND genes exhibit a lower degree of sequence conservation than other protein-coding genes within the mitochondrial genome; $ND_6$ being the least conserved (Attardi 1985). A comparison between human and bovine mtDNA found that average amino acid conservation was in the order of 70% (Anderson *et al* 1981), and Johanssen *et al* (1990) found amino acid identities of 76%

and 46% respectively in comparisons of the $ND_1$ and $ND_4$ genes in *Xenopus* and *Gadus*. Although the general pattern of coding genes suggests that, in general, internal deletions are rare, short deletions have been observed in ND genes from interspecific comparisons (Garesse 1988; Arnason & Johnsson 1992). The low sequence convergence seen in these genes may be a consequence of the functionally important region being located on nuclear encoded ND subunits - the mitochondrially coded subunits may only have a limited regulatory role (Johanssen *et al* 1990).

Cytochrome b is a transmembrane electron transporting protein that comprises an iron-protoporphyrin (hæme) prosthetic group (Stryer 1975). It is an essential part of the electron transport chain and has distinctive properties and structure, and being such an essential part of a sequential chain of proteins it might have very little tolerance to change. The gene coding for cytochrome b is well known and has been used in a number of phylogenetic studies to determine the evolutionary relationships between taxa. Conserved regions and conserved amino acid sequences mean that sequences for distantly related taxa can be aligned (Irwin *et al* 1991). Sequence analysis of the entire 1140 base pairs of the cytochrome b gene for twenty mammalian species has, for example, been used as evidence to strongly support the theory that there has been three main radiations within the mammalia (Irwin *et al* 1991).

Sequence data from the 12S rRNA gene (304bp) and partial sequence data from the cytochrome b gene have been obtained from a ray-finned fish (*Cichlasoma citrinellum*), a lung fish (*Lepidosiren paradoxa*) and a lobe-finned fish, the coelacanth (*Latimeria chalumnae*) and compared with published sequence data from a tetrapod (*Xenopus laevis* - Rowe *et al* 1985). These data were used to demonstrate that (with statistical confidence) the lung fish mtDNA is more closely related to that of the tetrapod than is the mtDNA of the coelacanth (Meyer *et al* 1990). A result that appears to preclude the possibility that the coelacanth lineage gave rise to land vertebrates: an area of contention amongst palaeontologists and comparative morphologists (Romer 1966; Løvtrup 1977; Lagios 1982; Fritzsch 1987; Forey 1988).

The 12S rRNA gene has been extensively used for elucidating inter-ordial differences, particularly for clarifying various artiodactyl relationships (Miyamoto *et al* 1990; Arnason & Johnsson, 1992). Sequence data from this gene has also been used to describe interfamilial differences between

African mole rats, in which the sequence divergence between families was high (average 16·8%) compared to artiodactyls and primates (Allard & Honeycutt, 1992).

The rate of evolution of mtDNA has been calibrated in primates (Brown *et al* 1982). Comparison of sequence variation in the cytochrome b and cytochrome oxidase I gene in thirteen species of sharks found that the nucleotide substitution rates in these region was seven to eight times slower than in the same region in primates or ungulates (Martin *et al* 1992). These results are generally supported by fossil evidence and it is unlikely that they can be explained by nucleotide composition bias, codon-usage bias or selection. The low divergence rates seen in sharks may be a result of many factors; these animals may be in a period of evolutionary stasis (see Gillespie 1984), they may have an efficient mtDNA repair system or, the low metabolic rate may be related to lower levels of oxygen radical damage (Martin *et al* 1992). The number of amino acid substitutions in cottid fish in Lake Baikal was found to be lower in the cytochrome b gene than in ATPases (Grachev *et al* 1992). This has been thought to be a consequence of an early divergence event due to habitat selection in a very stable and slowly changing environment (Grachev *et al* 1992).

In addition to assessing phylogenies between distantly related taxa, sequence data from the cytochrome b gene has been used to elucidate relationships between closely related species. Sequence data from a 307bp region of the cytochrome b gene has revealed inter- and intrapopulation variation within four species of *Thunnus* (tuna) (Bartlett & Davidson 1991). A similar study of Atlantic cod (*Gadus morhua*) discovered 24 variable nucleotide region in a 369bp region of the cytochrome b gene (Carr & Marshal 1991).
The most variable region of the mitochondrial genome appears to be the control region (section 4·1b). In contrast to the other regions of the mitochondrial genome the control region contains sequences that appear to be non-essential (Brown 1985). The evolution of this region has been found to vary in inter- and intraspecific comparisons (Saccone *et al* 1991). Within the mammalian mitochondrial genome the non-coding region is unique in that, except for the central region of circa 200bp, it can not be aligned interspecifically. Unambiguous alignments are possible

throughout the remainder of the mitochondrial genome (Anderson *et al* 1981; Greenberg *et al* 1983; Saccone *et al* 1991). Bibb *et al* (1981) discovered that the most striking homology was to be found at the 3' and 5' ends of the region. The rate of substitution in the control region varies widely in comparison with the rate of substitution found in protein-coding genes. Brown (1985) found the substitution rate in the control region to be 1·4 to 2·8 times that found in the protein genes, but the higher figures of 5 times and 10 times have also been reported (Greenberg *et al* 1983; Wilson *et al* 1985; Cann *et al* 1987).

The termination associated sequence (TAS) was found to be conserved relative to the D-loop region of the control region. This is consistent with a functional role being assigned to the TAS in termination of the D-loop (Brown *et al* 1986). However an alternative explanation may be that there is a higher mutational load for the triple-stranded region (Olivo *et al* 1983). Based on the assumption that only the D-loop region is evolving at a high rate, the 225bp sequence derived from this study was estimated to be diverging by 3·5% per million years (Thomas *et al* 1990).

Populations of African cichlid fish (*Tropheus* sp.) from Lake Tanganyika are very similar morphologically, but were found to carry a *'surprisingly large genetic divergence'* in terms of genetic variation in the D-loop and cytochrome b genes (Sturmbauer & Meyer 1992). The lineage of six species from Lake Tanganyika was found to contain twice as much variation as the very diverse cichlid assembly of Lake Malawi and six times more variation than the Lake Victoria species group. *Tropheus* are known to be confined to rocky habitats and have a limited capacity for open water dispersal. The pattern of variation was geographically partitioned, thus suggesting that major fluctuations in the water level of the lake have imposed constraints on the distribution and consequent speculation of this lineage (Sturmbauer & Meyer 1992).

Conserved sequence blocks have also been found in the control region of other species - suggesting a degree of functional conservation. For example, in man (Anderson *et al* 1981; Greenberg *et al* 1983), dolphin (Southern *et al* 1988), cod (Johanssen *et al* 1990), fin whale (Arnason *et al* 1991) and harbour seal (Arnason & Johnsson 1992). Some of these conserved blocks in the mitochondrial genome are species specific and may represent more specialised functions (Southern *et al* 1988).

Analysis of sequence data from the control region in human populations has given strong support to the 'Garden of Eden' hypothesis (Wainscoat 1987) for the origins of modern man. The limited genetic variability seen in human mtDNA has been accounted for in terms of a recent common ancestry - thus implying that all modern mtDNA genotypes originated from a single woman who lived in Africa some 200,000 years ago (Cann *et al* 1987). An analysis and alignment of the entire D-loop region of 14 human races along with 2 chimpanzee species: common (*Pan troglodytes*) and pygmy (*P. pygmaeus, P. paniscus*) (Pesole *et al* 1992) has indicated the age of 'our common mitochondrial mother' or 'Eve" to be 0·8 ± 0·2 million years. This age is derived by fixing the time of divergence between human and chimpanzee at 5 million years (Wilson & Sarich 1967). The average substitution rate was calculated at 12 ± 5 substitutions per 100 variable sites per million years. (Pesole *et al* 1992). These data for the origin of modern man imply that our ancestors would have been in the *H. erectus* population and thus the appearance of *H. sapiens* need not have been a speciation event. However, the age of 'Eve' is open to different interpretations depending on (for example) the estimate of the rate of evolution and the human chimpanzee divergence time.

Molecular data derived from the mitochondrial genome have provided much evidence and statistical support in understanding relationships within and between groups of animals, particularly so in groups which have been areas of contention among palaeontologists. Restriction analysis of the complete mitochondrial genome has been used on many occasions, as have sequence data from various regions, such as the control region, cytochrome b gene, ND genes, and, to a lesser extent, the tRNA genes. The investigations that have produced sequence data have, of necessity, made use of the polymerase chain reaction to amplify minute quantities of DNA. Without the advent of PCR data for even a short region of DNA would be difficult and tedious to obtain due to the length of time needed for constructing and screening libraries for each sample. Use of the polymerase chain reaction has thus greatly enhanced phylogenetic work and genetic population studies of the types mentioned in this chapter.

The advantages of PCR amplification and sequencing techniques include:

- sequence analysis of 'ancient' DNA extracted from museum and archaeological samples (Pääbo *et al* 1988)
- Sensitivity great enough to permit analysis of DNA from a single cell (Zhang *et al* 1992), single hairs (Vigilant *et al* 1989) and single sperm (Liu *et al* 1992).
- Elimination of the need to purify mtDNA preparations from contamination by nuclear DNA.
- Ease of which a relatively large number of samples can be processed in a comparatively short time.

In the remainder of this chapter I shall report the use of the polymerase chain reaction to amplify the D-loop region of muntjac deer prior to digestion with various restriction enzymes to produce population diagnostic patterns of restriction fragment polymorphism.

## 4·2 MATERIALS AND METHODS

### 4·2a Isolation of mtDNA

There are several methods of obtaining intact mitochondrial DNA from animal cells. The most reliable method is that of caesium chloride density gradient centrifugation (Brown *et al* 1979; Lansman *et al* 1981), in which a concentration gradient is created and macromolecules congregate at a particular level in the gradient dependent upon their buoyant density. In the presence of ethidium bromide, supercoiled DNA can be separated from non-supercoiled. Ethidium bromide intercalates between the relaxed strands of the double helix of DNA thereby causing a decrease in the buoyant density. Super coiled DNA (with no 'free' ends) has very little freedom to unwind and can therefore only bind a limited amount of ethidium bromide, thus the decrease in buoyant density is less than in linear or open circular DNA. Consequently mitochondrial DNA can be separated from nuclear DNA, as well as RNA and proteins. This method is very reliable and produces contamination free mitochondrial DNA but it is very time consuming and not appropriate for processing the 394 samples that formed the basis of this study. There was also a problem with degradation in many of the muntjac samples collected before the hunters were supplied with vials of DMSO solution. Degradation may produce DNA molecules of equivalent buoyant density to the mtDNA. Therefore this method was rejected and a number of 'rapid' methods tried and considered.

These methods included:-

1) The rapid mtDNA isolation method described by Chapman and Powers (1984). This method describes extraction of mtDNA from mature oocytes from fish. In this (muntjac) study I only had muscle tissue from tongue, and being a robust tissue, hand held homogenisation is difficult and tends to break up nuclear DNA, resulting in nuclear DNA contamination. For nearly 400 samples this method would be very labour intensive and, given the poor initial results, it was abandoned.
2) The method of Welter *et al* (1989) was also considered, but as this method also relied on homogenisation it was deemed to be inappropriate (as above) for the muntjac samples.
3) For similar reasons as above the methods of Palva and Palva (1985), and Tamura and Aotsuka (1988) were also considered to be inappropriate.

Of critical importance in all these techniques are (1) contamination: the ratio of nuclear DNA to mtDNA, small quantities of nuclear DNA contamination will reduce and may even eliminate resolution of mtDNA fragments. (2) Tissue volume: the relatively low abundance of mitochondria in muscle demands amounts of tissue in the order of 100gm (Chapman and Powers, 1984), and the very large dilution factors (1:9) necessary to separate muscle fibrils from mitochondria require centrifugation of volumes in the order of 1L. (3) Ease of homogenisation: mitochondrial membranes are less likely to be disrupted and the nuclear DNA less likely to be broken up if the tissue itself is easily homogenised.

**4·2b Alternative Approaches**

As none of the 'rapid' isolation techniques seemed appropriate, a different approach was considered - one that would be unresponsive to nuclear DNA contamination and yet sensitive to very small quantities of mtDNA. This technique is described by Hynes *et al* (1989) and requires the purification and cloning of mtDNA from one individual, which can then be used as a hybridisation probe for screening restriction digests from a large number of samples. The probe will hybridise specifically to mtDNA even in the presence of nuclear DNA contamination and will easily detect pico-gramme quantities of mtDNA restriction fragments. Although this approach seemed attractive it was abandoned in favour of a much quicker, albeit less definitive procedure.

Many studies of mtDNA variation have involved digestion of the complete mitochondrial genome with a number of restriction enzymes in order to display variation in terms of RFLPs, as has been discussed previously. Other studies have shown that the control region is highly variable and can be used to detect genetic variation among populations (Wainscoat 1987; Thomas *et al* 1990; Sturmbauer & Meyer 1992). In some cases the ND genes, and in particular $ND_6$ have been found to carry deletions (Garesse 1988; Arnason & Johnsson 1992). However Bernatchez *et al* (1992), found the control region to be almost invariant in *Salmo trutta* populations. This may be due to a slower rate of evolution in this region in salmon compared to other vertebrates, or there may be some, as yet, unknown functional constraint(s).

### 4·2c Use of The Polymerase Chain Reaction

The development and subsequent refinements of the polymerase chain reaction now permits very specific amplification of minute quantities of DNA from tissue, single hairs and museum specimens (Thomas *et al* 1990). To investigate the muntjac populations that form this study it was decided to amplify the complete control region and $ND_1$ plus $ND_6$ genes and subject them to restriction enzyme digest to produce a small number of restriction fragments which, if the regions proved variable, would then show sub-population differences.

The polymerase chain reaction utilises highly specific short oligonucleotide sequences, commonly of 15 - 35 base pairs, known as primers. These primers are designed from known sequences that flank the region of interest. Alternatively 'universal' primers may be used which are designed to amplify a sequence between two regions that are highly conserved in a wide variety of species (Kocher *et al* 1989). The primers are added in excess to the template DNA in the presence of an appropriate buffer, free nucleotides and a thermostable DNA polymerase. The double stranded template DNA is denatured by heating to about 93°C then cooled to 40 - 60°C (empirically determined in relation to the degeneracy of the primers). At the appropriate temperature the primers optimally anneal to the template, thus permitting the DNA polymerase to start synthesis of antisense copies of the region between the primers. The temperature of the reaction is raised to 70°C, which is the optimal temperature for the DNA polymerase. The DNA polymerase is disassociated, and the template and synthesised DNAs are denatured by heating the reaction to 93°C, and the whole cycle repeated. The cycle is repeated 30 to 40 times and as each copy becomes a template in proceeding cycles, the rate of accumulation of product is exponential (Hoelzel & Green 1992).

In this study the PCR was used for two areas of investigation: (1) amplifying the control region and (2) amplifying two of the open reading frame (ORF) regions, the 2kb (ND1+16S rRNA) and 2·5kb (ND5+ND6) regions of mtDNA. The reactions were performed in either a Perkin-Elmer Cetus 9600, a Hybaid OmniGene TR3 CM220 or a Hybaid Thermal Reactor TR1.

### 4·2d Amplification of control region and ORF Regions.

Initially DNA was taken from 17 of the 24 Woburn Abbey samples. The original extractions were used and diluted 1:10 with $ddH_2O$. This gave a range of concentrations across all samples of between 7ng/µl and 300ng/µl. It was found that isolation of mtDNA from the samples was not necessary as sufficient for PCR amplification was present in the original total genomic purifications, and the homology of the primers used was such that (having derived the appropriate stringency requirements) only the specific targets were amplified. To check for contamination, a negative control substituting $ddH_2O$ for DNA was carried out for every series of PCR reactions.

#### 4·2di Amplification of the mitochondrial ORF region.

Primers were provided by Dr. Heather Hall from the Institute of Zoology, London. The primer sequences were designed by Dr. N. Georgiatis and L.G.L. Patton of Washington University.

Primers for 2kb region 16S + $ND_1$ (ND = genes of the NADH complex):

Forward primer (16S tRNA)

5′- ACC CCG CCT GTT TAC CAA AAA CAT - 3′ (24 mer)

Reverse primer (met tRNA)

5′- GGT ATG AGC CCG ATA GCT TA - 3′ (20 mer)

Primers for 2·5kb region ($ND_5$ + $ND_6$):

Forward primer (leu tRNA)

5′ - AAT AGT TTA TCC AGT TGG TCT TAG G - 3′ (25 mer)

Reverse primer (glu tRNA)

5′ - TTA CAA CGA TGG TTT TTC ATA GTC A - 3′ (25 mer)

Initially 1µl of DNA was taken at whatever concentration the original extraction produced. If, after optimisation, an amplification failed, the concentration of DNA was reduced 10 fold, and 10 fold again if necessary. If the amplification continued to fail the concentration was increased two fold and two fold, etc. It was found that a few extractions required dilution whilst others required concentration - regardless of the actual amount of DNA present. Generally 50ng to 500ng amplified successfully.

Using filter pipette tips and sterile PCR tubes the following reaction was set up, the ingredients being added in this order:

| | |
|---|---|
| autoclaved milli-Q water | 37·9μl |
| total genomic DNA | 1·0μl |
| *10x *Taq* XL reaction buffer (NBL) | 5·0μl |
| **dNTPs | 4·0μl |
| primers (25pM) | 1μ + 1μl |
| *Taq* XL DNA polymerase (NBL) | 0·1μl |
| | 50μl |

*1x buffer = 45 mM Tris-HCl *p*H 8.8; 11 mM $(NH_4)_2SO_4$; 4·5 mM $MgCl_2$; 4·5 μM EDTA; 6·7 mM ß-mercaptoethanol + 110 μg/ml bovine serum albumin.

**2′-Deoxy (nucleosides) 5′-triphosphate (Pharmacia) at a concentration of 1·25mM each.

If the reaction was to be performed in either of the Hybaid machines then it was overlaid with 20μl of mineral oil (Sigma). A mineral oil overlay is not necessary when using the Cetus machine. The cycle conditions were as follows:

In the Hybaid thermal cyclers;

94°C/3min, (94°C/1min - 55°C/1min - 72°C/1min) x 30 cycles, 72°C/10min - ambient.

In the Cetus thermal cycler;

94°C/3min, (94°C/30sec - 55°C/30 - 72°C/30) x 30 cycles, 72°C/10min - cool to 4°C

**4·2dii Amplification of the mitochondrial control region.**

Five primers were used, four from the laboratory stocks; thr-l and phe-h, which flank the complete control region; dl-h and its complement dl-l, which are located in the middle of the D-loop at the origin of replication and H34 (supplied by Dr Helen Stanley of the Institute of Zoology, London), which is located approximately 63 bases 3' of dl-l. See Figure 4·1

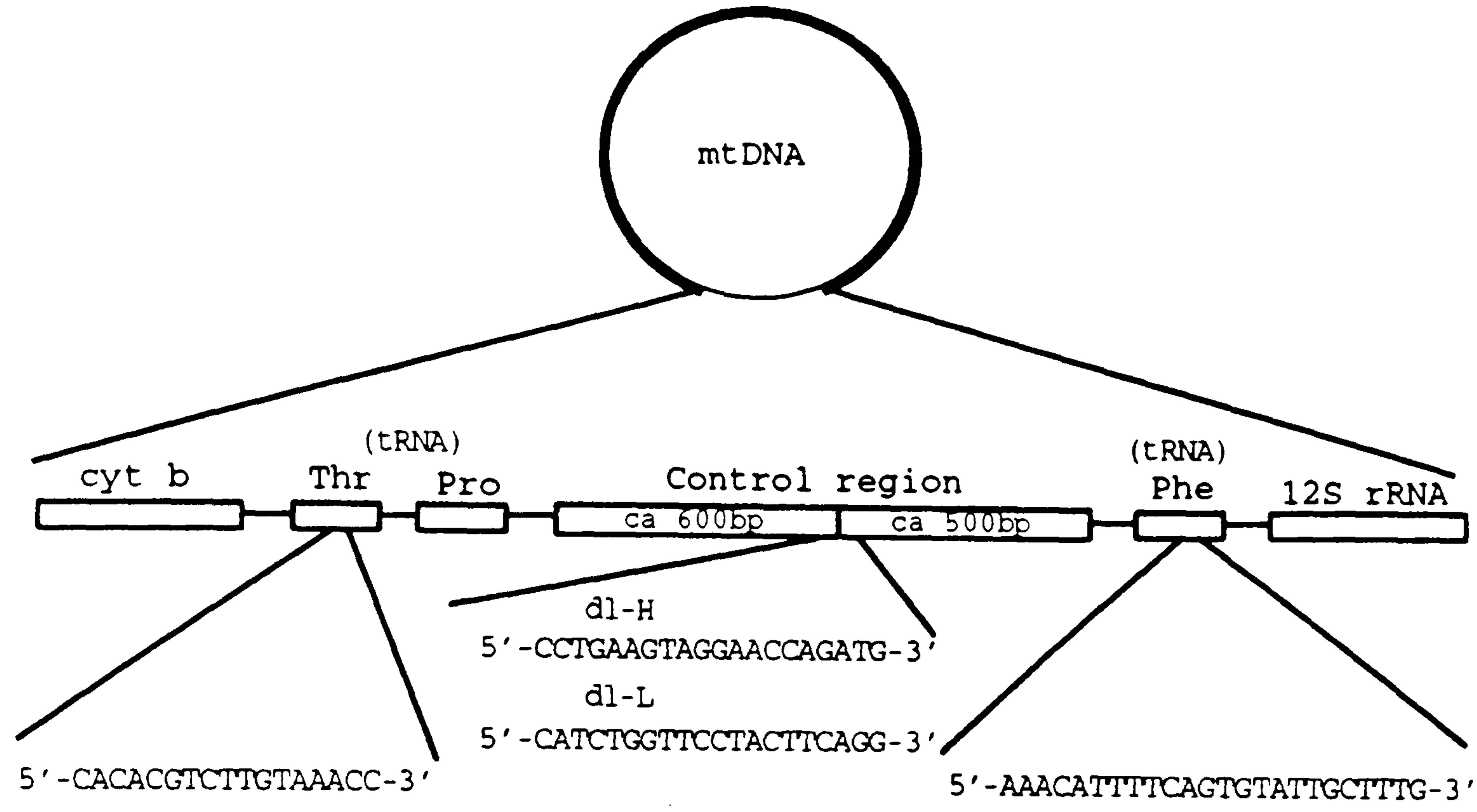


Figure 4·1 The control region and flanking regions of muntjac mitochondrial DNA

| | | |
|---|---|---|
| Thr-l (thr) | 5′- CAC CAG TCT TGT AAA CC - 3′ | (17 mer) |
| Dl-h | 5′- CCT GAA GTA GGA ACC AGA TG - 3′ | (20 mer) |
| Dl-l | 5′ - CAT CTG GTT CCT ACT TCA GG - 3′ | (17 mer) |
| H34 | 5′ - CCA AAT GCA TGA CAC CAC AG - 3′ | (20 mer) |
| Phe-h (phe) | 5′ - AAA CAT TTT CAG TGT ATT GCT TTG - 3′ | (24 mer) |

The primer H34 was taken from the mtDNA sequence of the harbour seal (*Phoca vitulina*) (Arnason & Johnsson, 1992).

The polymerase chain reaction applied to the control region was essentially the same as that applied to the ORF region with the following changes:

| | |
|---|---|
| dd$H_2O$ | 41·5μl |
| DNA (1:10 dilution) | 1μl |
| 10x *Taq* buffer | 5μl |
| dNTPs (1 mM each) | 2μl |
| primers (25 pm) | 0·2μl + 0·2μl |
| *Taq* XL polymerase | 0·1μl |
| | 50μl |

Cycle conditions:

94°C/3min, (93°C/30sec - 51°C/90sec - 72°C/90sec) x 40 cycles, 72°C/10min -ambient

To assess the quantity and success of the PCR reaction 10% of the PCR product was taken and loaded into a 14cm 1% agarose gel with a 123bp ladder molecular weight marker in an adjacent well and run at 5v/cm for 30 minutes. The PCR product then being visualised by staining with ethidium bromide.

### 4·2e Digestion with restriction endonucleases.

Restriction fragment lengths of the various amplified regions of the mitochondria were assessed for variation between individuals. Those enzymes which produced Restriction Fragment Length Polymorphisms (RFLP) were used to screen all the populations.

Eight microliters of the reaction was taken and digested according to the manufacturer's recommendations with 3 units of enzyme for three hours in a final volume of 20µl. The total product was then run out on a 2% agarose gel and visualised as above.

A suite of 10 four, five or six base-pair recognition site restriction endonucleases were selected.

| | | | |
|---|---|---|---|
| *Alu* I | AG$^{\dagger}$CT | *Ava* II | G$^{\dagger}$C(AT)CC |
| *Dde* I | C$^{\dagger}$TNAG | *Hae* III | GG$^{\dagger}$CC |
| *Hin* fI | G$^{\dagger}$NATC | *Mbo* I | $^{\dagger}$GATC |
| *Msp* I | C$^{\dagger}$CGG | *Rsa* I | GT$^{\dagger}$AC |
| *Taq* I | T$^{\dagger}$CGA | *Xba* I | T$^{\dagger}$CTAGA |

Of the ten restriction enzymes used, three were found to produce polymorphic restriction fragments. These were *Hin* fI, *Mbo* I and *Rsa* I - the remainder were thus eliminated.

### 4·2f Localisation of control region restriction sites

Samples representative of each mitochondrial restriction fragment pattern were amplified by PCR. One half of the control region was amplified using the primer pair dl-l/phe-h and digested with the three informative restriction enzymes, as described above, and compared with the digests of the total control region (as amplified using the primer pair thr-l/phe-h). The product was visualised by running into a 2% agarose gel, followed by staining with ethidium bromide, as previously described. By comparing the various restriction fragments produced, and in association with sequence data from the central region of the control region it was possible to determine the location of the restriction sites (Brown 1990).

### 4·2g Selection and Purification of PCR Product

The total product of a 50µl PCR reaction was concentrated in a rotary vacuum dryer until the volume had been reduced to approximately 20µl. Five microliters of loading dye were added and the product run into a 20cm, 1% low-melting-point agarose gel for 30 minutes at 5v/cm, and visualised as above. The band of interest was excised, weighed, and transferred to a 1·5 ml Eppendorf tube containing 3x the weight by volume of NaI. The agarose and NaI were heated at 45°C until the agarose was fully dissolved. Ten microliters of glassmilk (powdered flint glass 325 mesh, prepared by boiling in 50% nitric acid, centrifuging out the glass, washing in water until neutral, and stored at 4°C in 50% water) was added, the tube was vortexed and placed on ice for 10 minutes, vortexing or flicking every 2 minutes or so. The eppendorf was centrifuged at 13000 rpm for 30 seconds and the supernatant removed. The pellet was washed three times by gently resuspension in 400µl NEW WASH (50% ethanol, 0·1 M NaCl; 10 mM Tris-HCl pH7·5; 1 mM EDTA, stored at -20°C), then repelleted by centrifuging for 10 seconds at 13000 rpm, and aspirating off the supernatant. After the final wash, the pellet was resuspended in 15µl $ddH_2O$, and the DNA eluted by incubation at 45°C for 5 minutes followed by centrifugation at 13000 rpm for 30 seconds, and careful removal of the supernatant containing the purified DNA. The elution procedure was repeated with a further 10µl of $ddH_2O$, resulting in final volume of 25µl containing the DNA. To check the purity and estimate the concentration of the recovery, 10% of the elutate was taken and run into a 1% agarose gel and visualised as above. The remainder was stored at -20°C.

### 4·2h Sequencing

Double stranded direct sequencing reactions were performed using the dideoxy chain-termination method of Sanger *et al* (1977), as modified after Green *et al* (1989) and Winship (1989). Immediately prior to the sequencing reaction the following mixtures were prepared and kept on ice. The volumes are sufficient for eight sets of sequencing reactions. A control of single stranded M13mp18 along with the -40 universal primer (both from the Sequenase kit - USB) was included with every set of reactions.

1) Primer mixture

| | |
|---|---|
| 2·5µl | primer at 25pM (either thr-l; dl-h; dl-l; phe-h) |
| 8·0µl | 5x Sequenase reaction buffer (United States Biochemical) |
| 3·0µl | Dimethylsulphoxide (DMSO) |
| 5·0µl | MilliQ $H_2O$ |

2) Labelling/Enzyme mixture

| | |
|---|---|
| 8·0µl | 0·1 M Dithiothreitol (DTT) |
| 8·0µl | labelling mix at 1:20 dilution (USB) |
| 11·0µl | MilliQ $H_2O$ |
| 3·0µl | DMSO |
| 4·0µl | $\alpha^{35}$S-dATP (1000Ci/mMol - Amersham) |
| 1·0µl | Sequenase enzyme. Version 2·0 (USB) |

3) 'Chase' mixture

| | |
|---|---|
| 16·0µl | MilliQ $H_2O$ |
| 5·0µl | dNTPs (at 1 mM each) |
| 2·5µl | DMSO |
| 1·5µl | 5x Sequenase reaction buffer |

4) Termination mixes

2·5µl of each termination mix (USB) were transferred to 500µl eppendorf tubes, sealed and kept at 37°C in a heating block.

The Reaction

1) Annealing. In a 500µl eppendorf tube, 3µl of purified PCR product was added to 4µl of primer mix and heated to 100°C for five minutes in a boiling water bath, followed by plunging into an ethanol/dry ice bath for two minutes. The tube was then spun briefly and kept on ice.

2) Extension. 4µl of labelling/enzyme mix was added, gently mixed by pipetting up and down once and incubated at room temperature for five minutes.

3) Termination. 2µl of the reaction from (2) was added to each of the termination mix tubes (at 37°C), mixed gently by pipetting once, and incubated at 37°C for five minutes. 2µl of 'chase' was added and the incubation continued for a further five minutes. 4µl of stop solution (95% formamide: 20 mM EDTA) was added and the reactions stored at -20°C.

A 5% Sequagel (15ml Sequagel concentrate, 52·5ml Sequagel diluent, 7·5ml Sequagel buffer + 600μl 10% ammonium persulphate and 30μl TEMED = NNN´N´ tetramethylethylenediamine) was prepared and pre-run for about 45 minutes. The samples were heated to 95°C for 2 minutes and 2.5μl loaded and run for circa 3 hours at 1·5kV. After the gel run was complete, the gel was transferred to a sheet of 3mm Whatman filter paper, covered with Saranwrap, and dried on a heated vacuum drier at 80°C. Exposure was on Kodak X-omat AR scientific imaging film, or Fuji RX x-ray film for a length of time determined by the activity on the membrane as measured with a Minimonitor 6000 (Mini-Instruments). The sequences were recorded by eye.

**4·2i Methods of Analysis**

Restriction maps of the control region were constructed from the restriction fragment data and a small stretch of sequence data. A composite restriction map was then produced and the sequence of restriction sites used to define composite genotypes for phylogenetic analysis using the program PAUP (Phylogenetic Analysis Using Parsimony - Swofford 1990).

An assessment of spatial association between sub-populations was investigated graphically, including regression analysis. Investigations of the relationships between sub-populations was approached from three directions. First, sub-populations were clustered by parsimony analysis using the branch-and-bound algorithm in PAUP - with weighting applied to each genotype to reflect its frequency of occurrence within localities.

The second approach was to obtain estimates of the migration rates between sub-populations from the relation $G_{ST} = 1/(1 + 4Nm)$, where N is the population size and *m* is the migration rate. These data were then plotted against distance for each sub-population. Cavalli-Sforza's arc distance was calculated and interpreted diagrammatically by a distance-Wagner dendrogram.

## 4·3 RESULTS

The amplification and restriction enzyme digestion of the 2kb (ND1+16S rRNA) and 2·5kb (ND5+ND6) regions of mtDNA found no restriction site variation in the representative individuals tested from the Woburn population. Further investigation of this region was therefore abandoned.

### 4·3a Restriction Profiles

The mitochondrial control region however proved to be variable enough to be informative and further work concentrated only on this region. The restriction fragment patterns produced by each informative restriction enzyme were recognised in relation to the molecular weight marker - and to each other per enzyme. Only the variable bands were used to identify a restriction profile. The variable restriction profiles for each of the three restriction enzymes used are shown diagrammatically in Figure 4·2. The complete data set may be found in Appendix C.

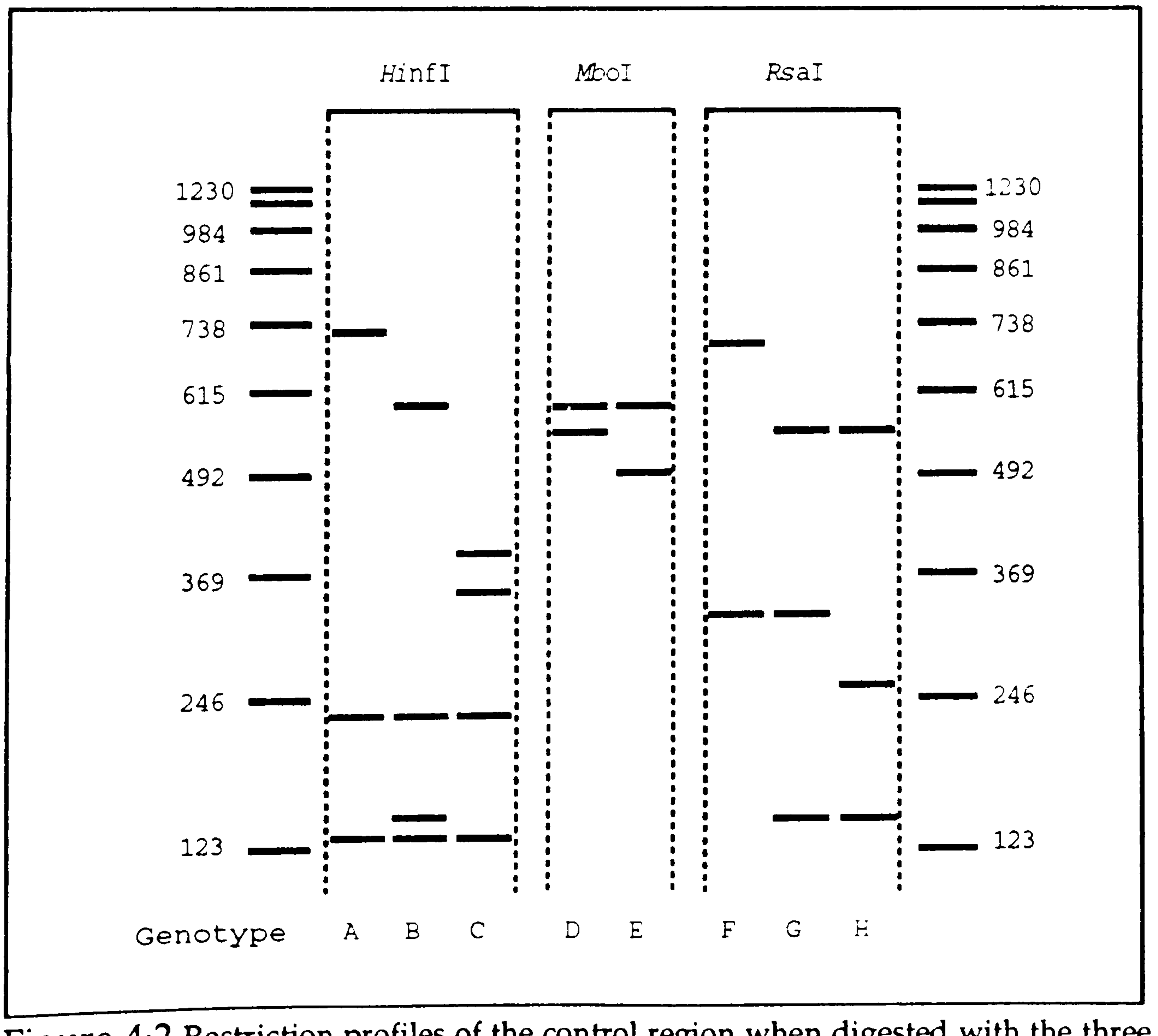


Figure 4·2 Restriction profiles of the control region when digested with the three informative restriction enzymes

The approximate molecular weights of the major fragments (in base pairs) are listed in Table 4·1. Minor fragments are those which are so small that they have either run off the gel or are too small to be identified by ethidium bromide staining, The figures in the 'minor fragments' column of Table 4·1 must therefore be taken as cumulative totals and may represent more than one fragment.

| Rest. Enz. | Genotype | Fragment Size in bp | | |
|---|---|---|---|---|
| | | variable | common | minor fragments |
| *Hin* fI | A | 720 | 230/125 | 55 |
| *Hin* fI | B | 590/130 | 230/125 | 55 |
| *Hin* fI | C | 390/330 | 230/125 | 55 |
| *Mbo* I | D | 590/540 | — | — |
| *Mbo* I | E | 590/490 | — | 50 |
| *Rsa* I | F | 670/300 | — | 160 |
| *Rsa* I | G | 540/300/130 | — | 160 |
| *Rsa* I | H | 540/250/130 | — | 210 |

Table 4·1 Fragment size of the various control region genotypes, in bp.

**4·3b Sequence Data from Two Regions of the control region**

Two regions were successfully sequenced, the region 3' to the threonine gene - using the primer thr-l and the central region (the origin of replication of the heavy strand) using the primers dl-l and dl-h. The sequence data are listed below, with the primer sites in bold and the restriction sites in bold and underlined.

←thr...5′-

CCCGGGCCTCGCTCTTGTTCCTAGGGACGGTTCGCCCAACGACCAAAG

TGCGCC**GTAC**CACT**GATT**CGAGCACCA**GATC**ACCTGCCCTATGC**GTAC**

AACTGTTCCTAACTAAACTGAGTTTACACGATACATGCTAATTCGTT

CCC**GTAC**ATGATACATGACATTTATTATCTTCATGCACGA.........CTAT

GACATCACAGTTATGTGTGAGCATGGGCTGATTAGTCATTAGTCCAT

CGAGATGTCTTATTTAAGAGGAAAGAGTGGGC**GATTC**TAGGTGAGA

TGGC**CCTGAAGAAAGAACCAGAT**GTCTGATAAAGTTCATTAAGTAG

CTACCCCCACA**GTAC**ATGGGCCCGGAGCGAGAACAAGGTCCCTGCCA

AGCGGGTTGCTGGTTTCACGCATGGGCCCGGAGCGAGAACAAGGTCC

TGCCAAGCGGGTTGCTGGTTTCACGCTGACTTAAATTGCTATTGTAC

A... 3′→phe

*Hin* fI - GANTC    *Mbo* I - GATC    *Rsa* I - GTAC

**4·3c Restriction Mapping of the control region**

From these data in association with the restriction profiles produced, as mentioned in section 4·2c, it was possible to produce a restriction map for the control region. There are some areas of doubt, but insufficient to affect the final result. In the following all the fragment sizes are approximate and derived mainly through comparison with a flanking 123bp ladder.

For *Hin* fI: See Figure 4·3. In digests of the complete control region compared with digests of half of the control region (dl-l → phe-h), fragments of 230bp and 125bp were found to be common to both the complete control region and the half control region. Therefore these fragments must originate in the region dl-l → phe-h. The remaining control region forms the 720bp fragment, 'A'. Genotype 'B', a 590bp fragment plus a 130bp fragment, but the later is difficult to identify given that the 125bp common fragment is only approximately this size - the two fragments could be similar enough to be indistinguishable. The restriction site that creates genotype 'B' is taken from the sequence data. Genotype 'C' (390bp + 330bp) is derived from a division of the 720bp fragment in 'A'.

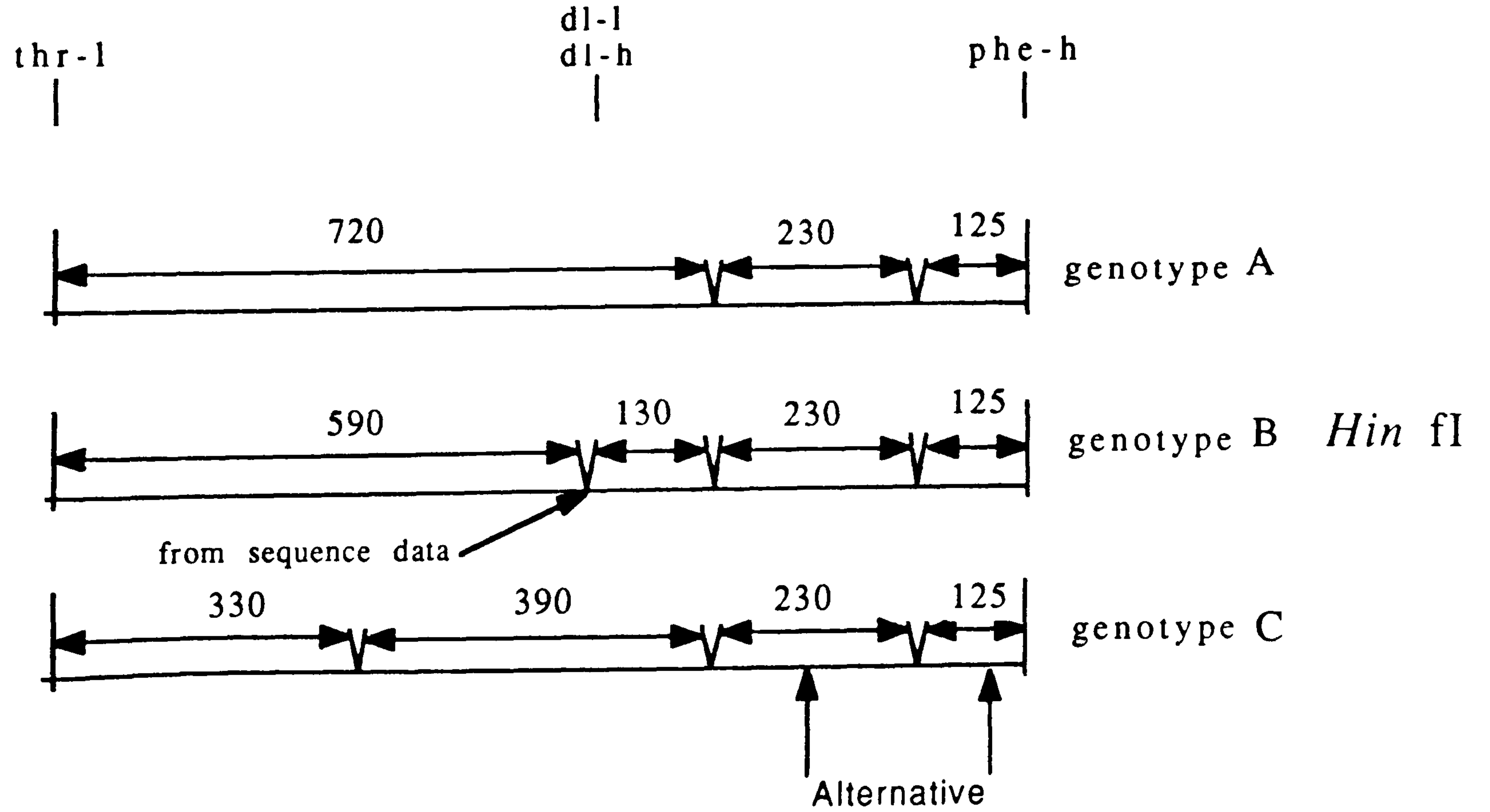


Figure 4·3 Restriction map of control region digested with *Hin* fI. 'V' refers to the location of restriction sites and the numbers are the sizes, in bases.

For *Mbo* I: See Figure 4·4. There were no restriction sites within the dl-l → phe-h region, and as this region is circa 500bp, the restriction sites for the 590bp and 520bp fragments must reside in the thr-l → dl-h region. The 590bp fragment restriction site being proximal to the dl-h priming site is confirmed by the reduction of the 520bp fragment to 480bp and the presence of a restriction site proximal to the thr-l end of the control region - from the sequence data.

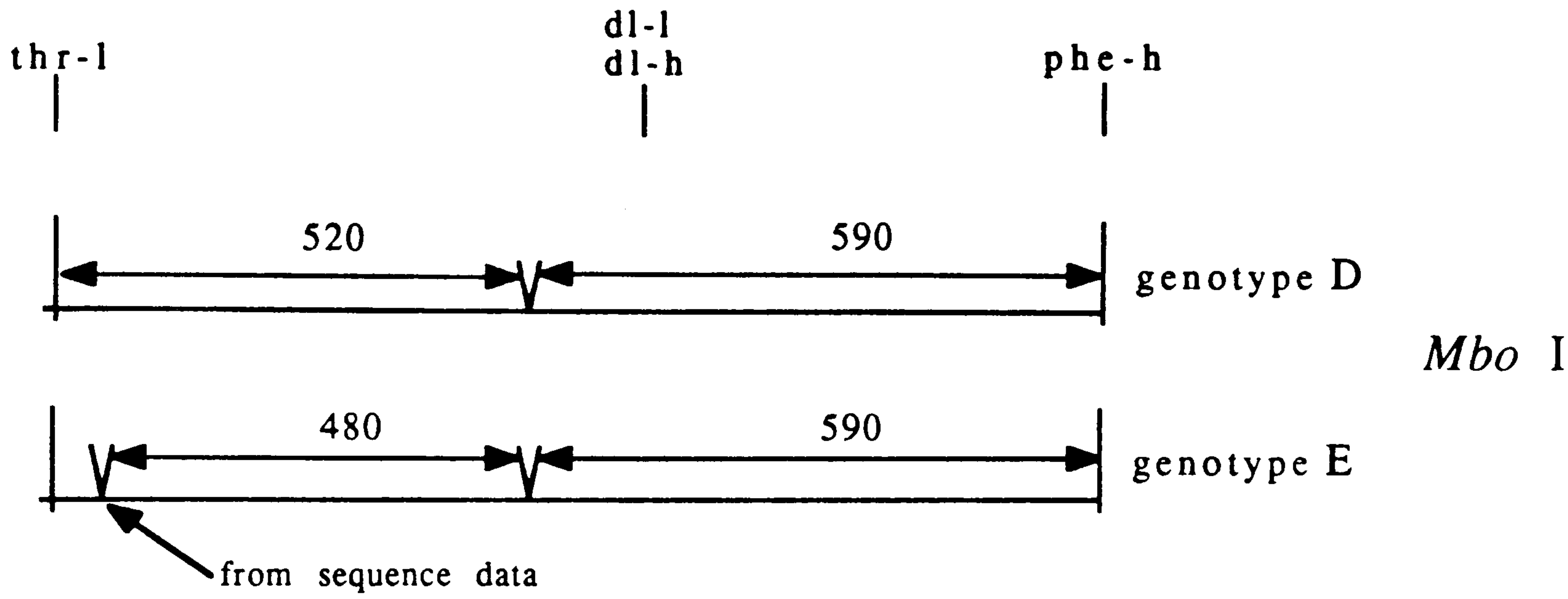


Figure 4·4 Restriction map of control region digested with *Mbo* I. 'V' refers to the location of restriction sites and the numbers are the sizes, in bases.

For *Rsa* I: see Figure 4·5. For genotype 'F' (670bp + 300bp), there were no restriction sites in the dl-l → phe-h region, therefore the restriction site for the 670bp fragment must be in the thr-l → dl-h region, proximal to the dl-h priming site. From the sequence data there are several restriction sites at the thr-l end of the control region, producing fragments ranging from 17bp to 80bp. The 300bp fragment is thus probably contiguous with the 670bp fragment. For genotype 'G' the restriction sites are known from the sequence data. There is some uncertainty about the location of the restriction site for genotype 'H', but not enough to have any significant effect.

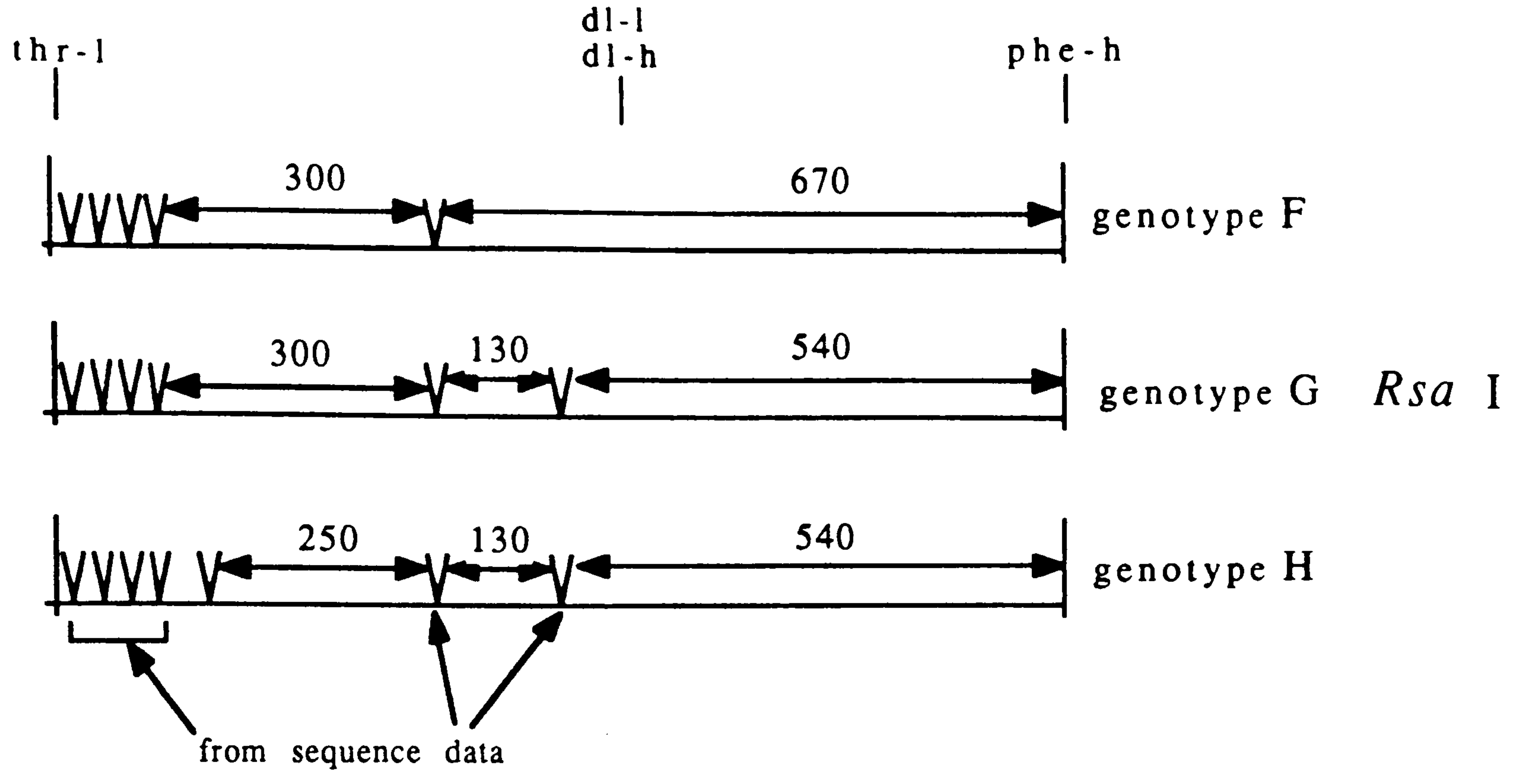


Figure 4·5 Restriction map of control region digested with *Rsa* I. 'V' refers to the location of restriction sites and the numbers are the sizes, in bases.

If all these data are combined, a composite restriction map of the control region can be compiled, as per Figure 4·6; the letters refer to the genotype and the numbers to the approximate fragment size in base pairs. The numbers 1 to 9 on the bottom line are arbitrarily assigned to the restriction site locations.

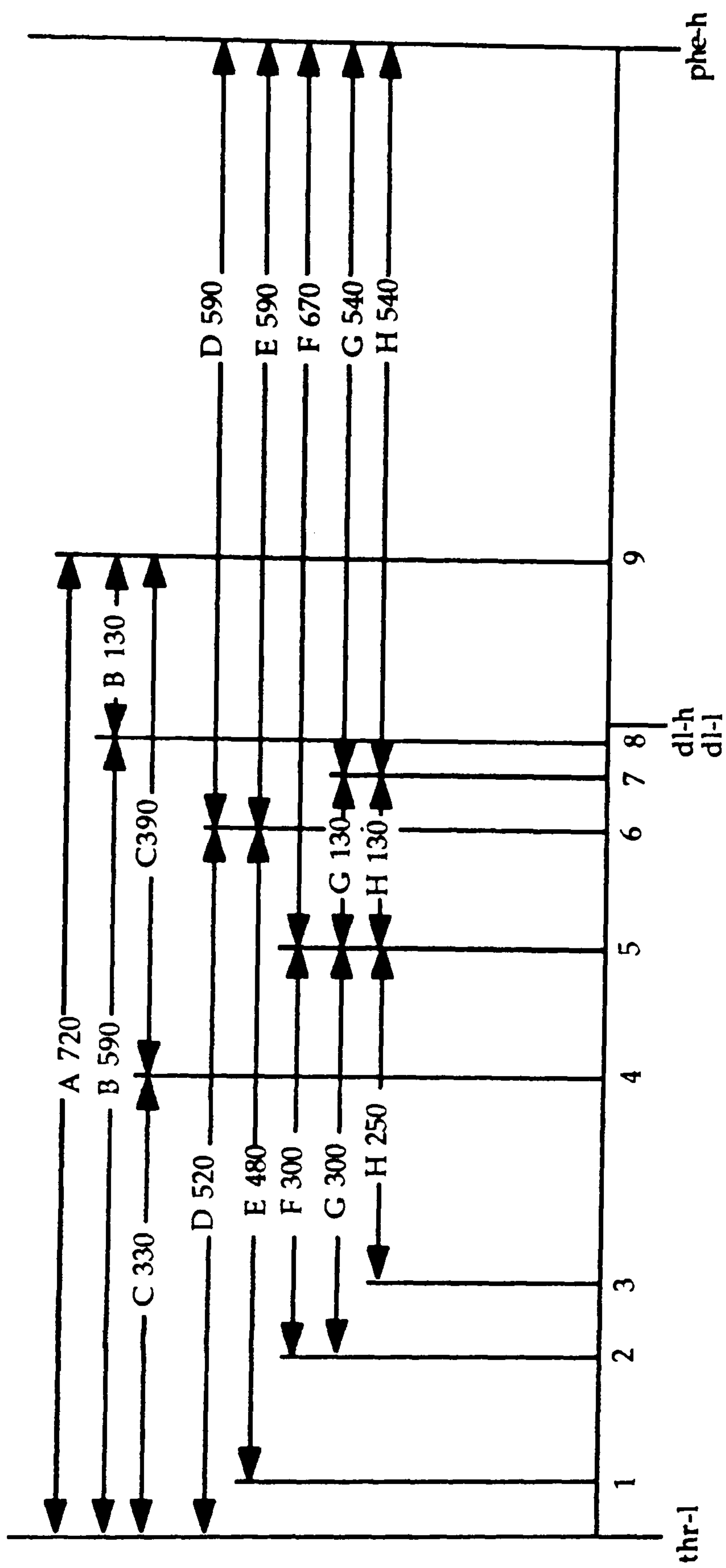


Figure 4·6 Composite restriction map of the control region based on the data from the three restriction digests (*Hin* fI, *Mbo* I and *Rsa* I) and the sequence data, as displayed in Figures 4.3, 4.4 and 4.5. The vertical lines refer to the locations of restriction sites and the numbers are the sizes, in bases.

A presence or absence matrix was constructed from the above restriction map. See Table 4·2 The rightmost (1) or (0) referring to the presence or absence of restriction site '9' from Figure 4·6, the next (1) or (0) referring to restriction site '8' and so on.

| Genotype | Restriction site locations per genotype | Presence (1) or Absence(0) |
|---|---|---|
| ADF | 9 - 6 - 5 - 2 | 0 1 0 0 1 1 0 0 1 |
| ADG | 9 - 6 - 7 - 5 - 2 | 0 1 0 0 1 1 1 0 1 |
| AEG | 9 - 6 - 7 - 5 - 2 - 1 | 1 1 0 0 1 1 1 0 1 |
| BDG | 9 - 8 - 6 - 7 - 5 - 2 | 0 1 0 0 1 1 1 1 1 |
| BDH | 9 - 8 - 6 - 7 - 5 - 3 | 0 0 1 0 1 1 1 1 1 |
| BEG | 9 - 8 - 7 - 6 - 5 - 2 - 1 | 1 1 0 0 1 1 1 1 1 |
| CDF | 9 - 6 - 5 - 4 - 2 | 0 1 0 1 1 1 0 0 1 |
| CEF | 9 - 6 - 5 - 4 - 2 - 1 | 1 1 0 1 1 1 0 0 1 |

Table 4·2 Presence or absence table for each genotype, by restriction sites.

The 'binary' coding for the genotypes was used for entering the data into the program PAUP (see later).

### 4·3d Derivation of Genotypes.

Due to having a small founder population, not all theoretically possible composite genotypes were evident in the samples tested. Eight were uncovered out of a possible eighteen - given that there were no further restriction sites to be found.

Taking the most parsimonious approach, the relationships between the restriction sites that form the composite genotypes may be displayed thus

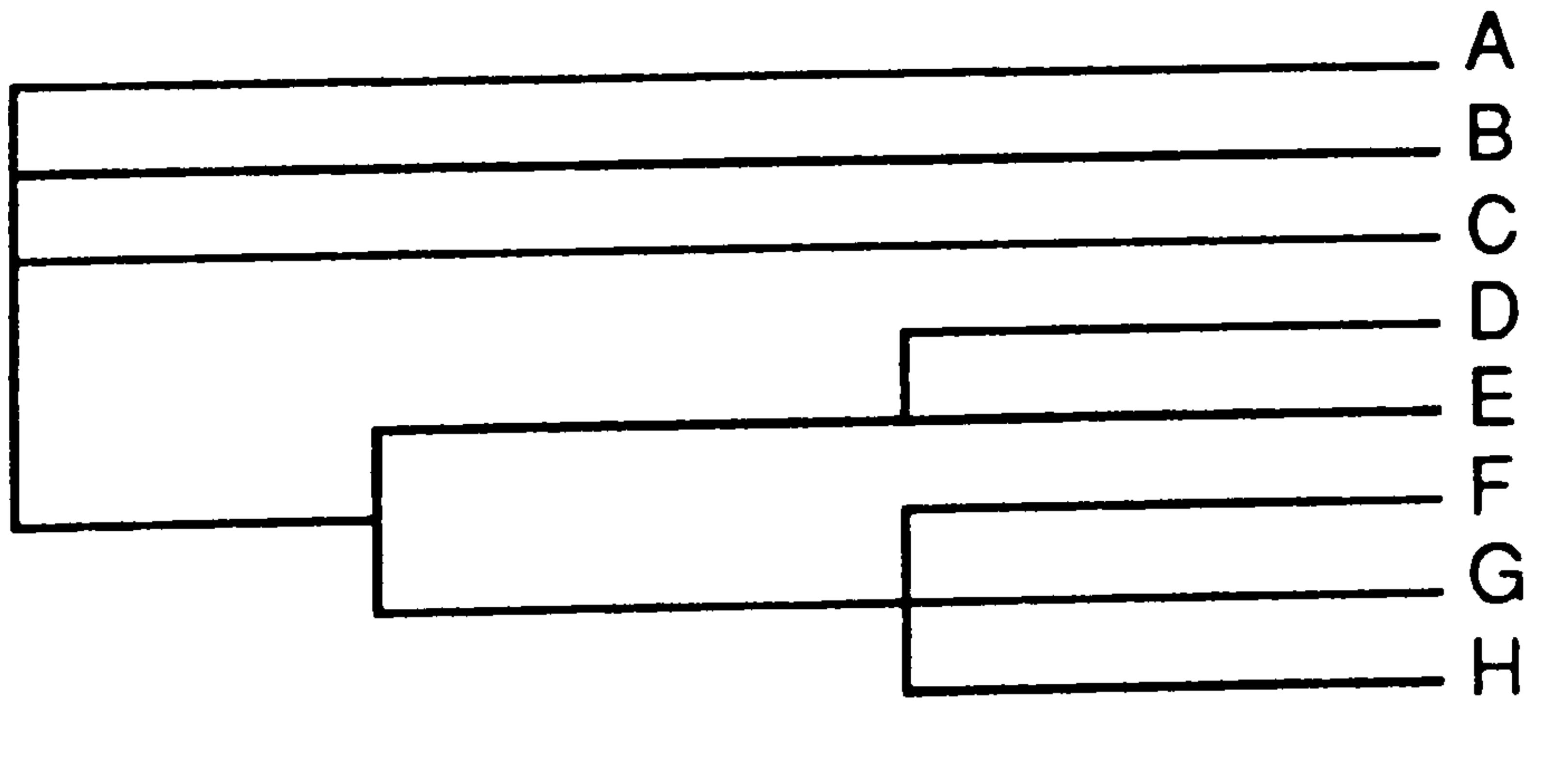


Figure 4·7 Most Parsimonious Tree of Restriction Sites.

B and C can - by one base change each - be derived from A (and vice versa). E can be derived from D by one base change (and vice versa). H can be derived from G by two base changes, and G can be derived from F, by one base change (and vice versa).

To demonstrate the relationships between all possible composite genotypes (hereinafter referred to as 'genotypes'), a tree was constructed from the total number of possible genotypes by using the 'branch-and-bound' algorithm included in the programme PAUP (Phylogenetic Analysis Using Parsimony) version 3·0 (Swofford 1990). The genotypes evident from the samples investigated are in bold.

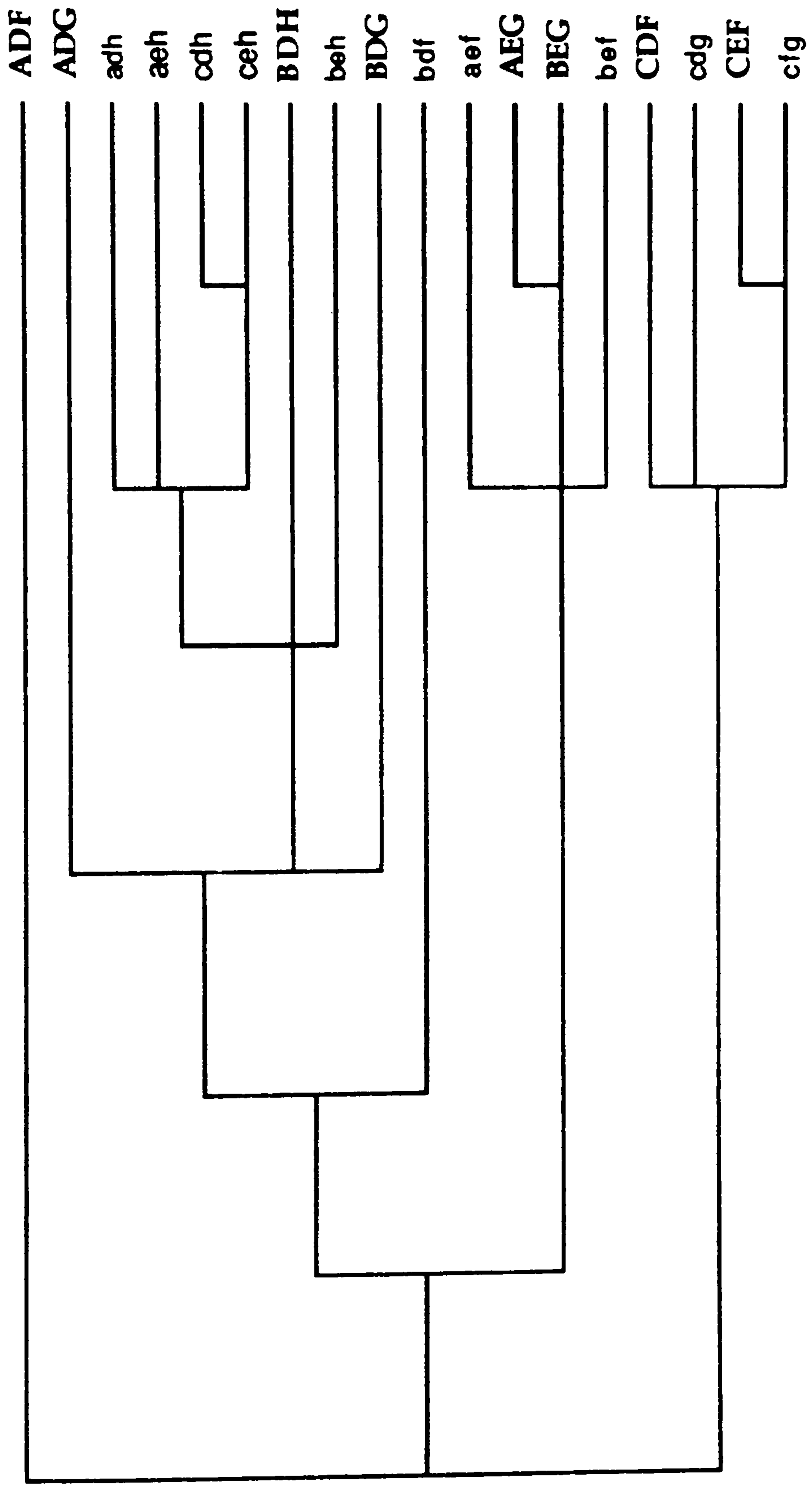


Figure 4·8 Dendrogram of all possible genotypes. Actual genotypes detected are in uppercase.

### 4·3e Genotypes

A total of eight genotypes were defined - as composites of the three digests. These were ADF (frequency of occurrence in the total population ($f_{tp}$) = 0·177); ADG ($f_{tp}$ = 0·005); AEG ($f_{tp}$ = 0·49); BDG ($f_{tp}$ = 0·101); BDH ($f_{tp}$ = 0·041); BEG ($f_{tp}$ = 0·018); CDF ($f_{tp}$ = 0·165); CEF ($f_{tp}$ = 0·002). The distribution of the genetic variation between populations from the eighteen locations is described in Table 4·3, with frequencies of occurrences at the sub-population level in parenthesis. The number of different genotypes in each locality ranged from one in the Bowden Park and Minsbury localities, to five at Knebworth. The commonest genotype (AEG) was found in 17 of the 18 sub-populations at a frequency ranging from 0·07 to 1·0. Three rare genotypes were observed, two of these rare genotypes were unique to two localities, CEF at Eling (frequency of occurrence in sub-population ($f_{sp}$) = 0·03), BEG at Southblock ($f_{sp}$ = 0·28), whilst the third rare genotype, ADG, was observed at Knebworth ($f_{sp}$ = 0·04) and King's Forest ($f_{sp}$ = 0·02).

| POP. | ADF | ADG | AEG | BDG | BDH | BEG | CDF | CEF | total |
|---|---|---|---|---|---|---|---|---|---|
| Wav | | | 2(0.4) | 3(0.6) | | | | | 5 |
| Tus | 14(0.77) | | 2(0.11) | 2(0.11) | | | | | 18 |
| Sbl | 13(0.52) | | 4(0.16) | 1(0.04) | | 7(0.28) | | | 25 |
| Sal | 2(0.14) | | 5(0.36) | 7(0.50) | | | | | 14 |
| Wob | | | 7(0.30) | 4(0.16) | | | 13(0.54) | | 24 |
| Hpk | 2(0.29) | | 5(0.71) | | | | | | 7 |
| Bpd | | | 13(1.0) | | | | | | 13 |
| Elg | 2(0.06) | | 24(0.80) | | | | 3(0.10) | 1(0.03) | 30 |
| Wst | 3(0.21) | | 10(0.71) | | | | 1(0.07) | | 14 |
| Owd | | | 19(0.95) | 1(0.05) | | | | | 20 |
| Nth | | | 24(0.96) | 1(0.04) | | | | | 25 |
| Min | | | 17(1.0) | | | | | | 17 |
| Bkp | 2(0.09) | | 12(0.54) | | 3(0.14) | | 5(0.23) | | 22 |
| Knb | 4(0.14) | 1(0.04) | 2(0.07) | 15(0.54) | 6(0.21) | | | | 28 |
| Hex | 3(0.10) | | 25(0.86) | | 1(0.03) | | | | 29 |
| Btf | | | 12(0.67) | | 6(0.33) | | | | 18 |
| Kf | | 1(0.02) | | | | | 34(0.98) | | 35 |
| Eus | | | 10(0.40) | 6(0.24) | | | 9(0.36) | | 25 |
| Twn | 25(1.0) | | | | | | | | 25 |
| totals | 70 | 2 | 193 | 40 | 16 | 7 | 65 | 1 | 394 |

Table 4·3 Distribution of genotypes throughout the populations, with frequencies in parenthesis.

## 4·3f Assessment of Spatial Associations

The number of genotypes detected in each population may be related to the sample size. Figure 4·9 describes the number of composite genotypes detected with increasing sample size from all 18 localities. Overlaid on this histogram is the regression slope of number of genotypes against sample size which, being positive, confirms a positive correlation between sample size and the number of genotypes detected (regression coefficient 0·45, slope 0·06, P=0·06 t-test). The two commonest genotypes were observed from sample sizes as low as seven, whilst unique genotypes were only detected from sample sizes greater than 25.

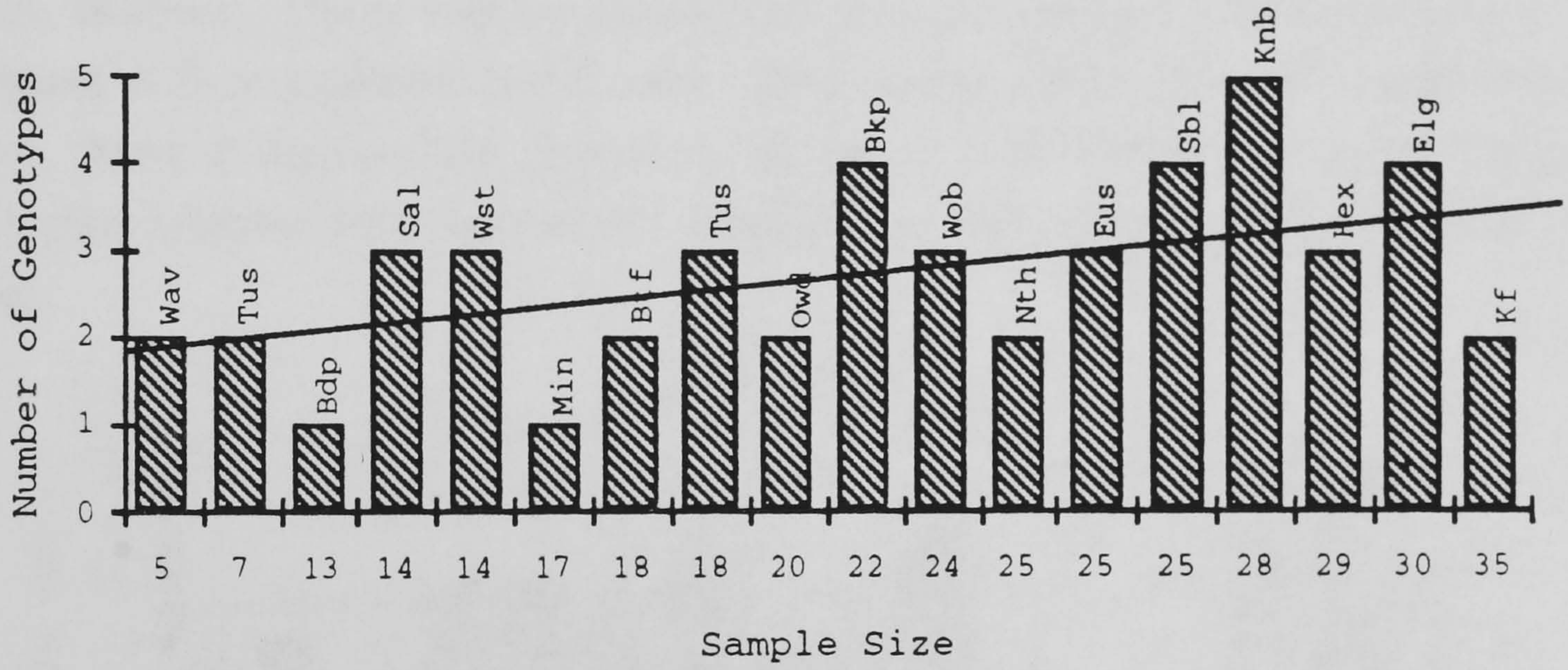


Figure 4·9 Relationship between number of genotypes and sample size

To attempt to further clarify the distribution of genotypes, the localities were divided into three groups, defined in radiating circles, dependent on their distance from Woburn. The groups were labelled near, medium and far with reference to Figure 2·1.

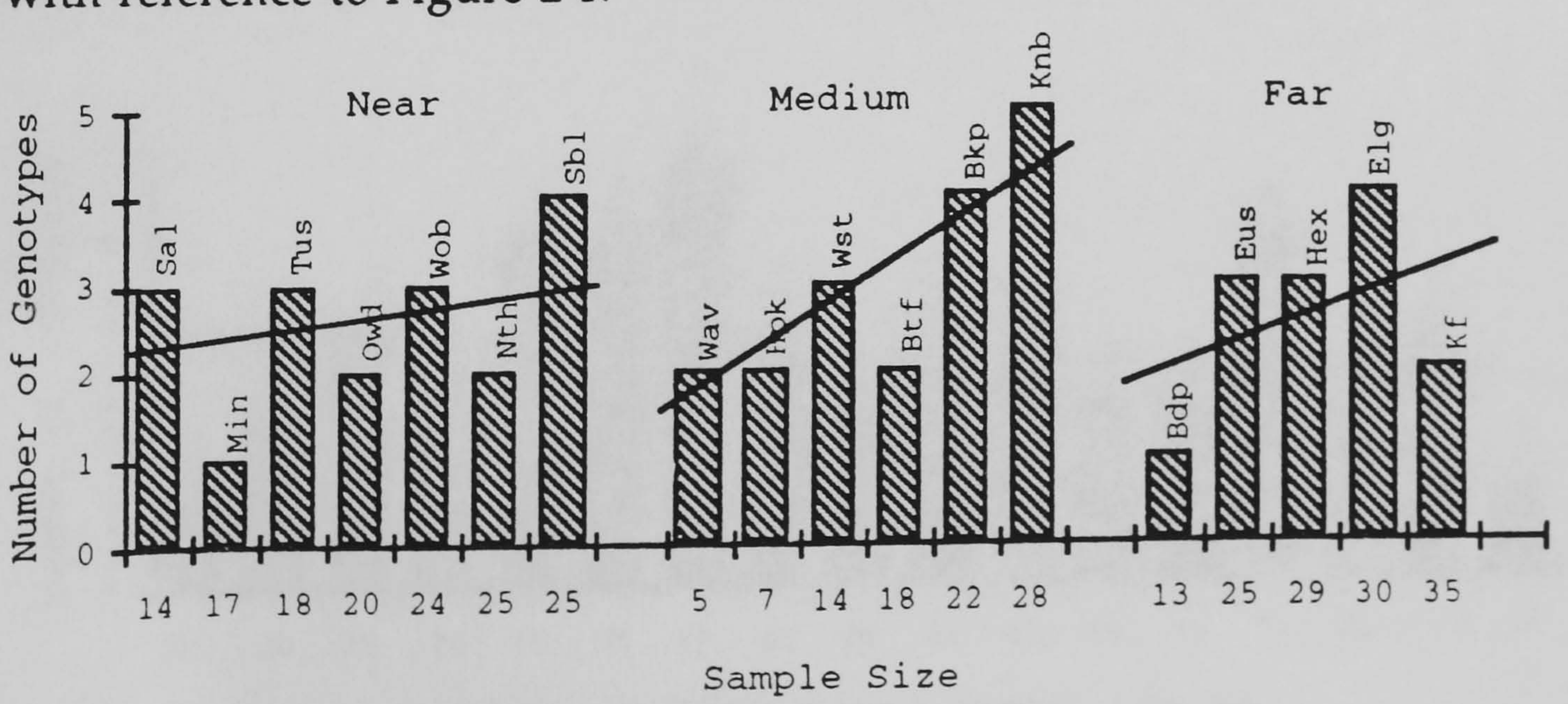


Figure 4·10 Relationship between number of genotypes and sample size for groupings of localities.

The histograms were again overlaid with the regression slope which in all three cases is positive (regression coefficient near = 0·28, medium = 0·86, far = 0·08; slope near = 0·06, medium = 0·12, far = 0·08; P near = 0·66, medium = 0·02, far = 0·29 t-test). As can be seen, all slopes are positive but there is only a significant correlation between number of genotypes uncovered and sample size within the confines of these distance groupings in the 'medium' group.

Regression analysis was also employed to determine if there was any association between sample size and distance from Woburn as well as to test for association between the number of genotypes detected and distance from Woburn. There was no correlation between sample size and distance (Figure 4·11, regression coefficient = 0·05, slope = 0·01, P = 0·85 t-test), nor was there a correlation between distance and number of genotypes detected (Figure 4·12, regression coefficient = 0·24, slope -0·008, P = 0·35 t-test).

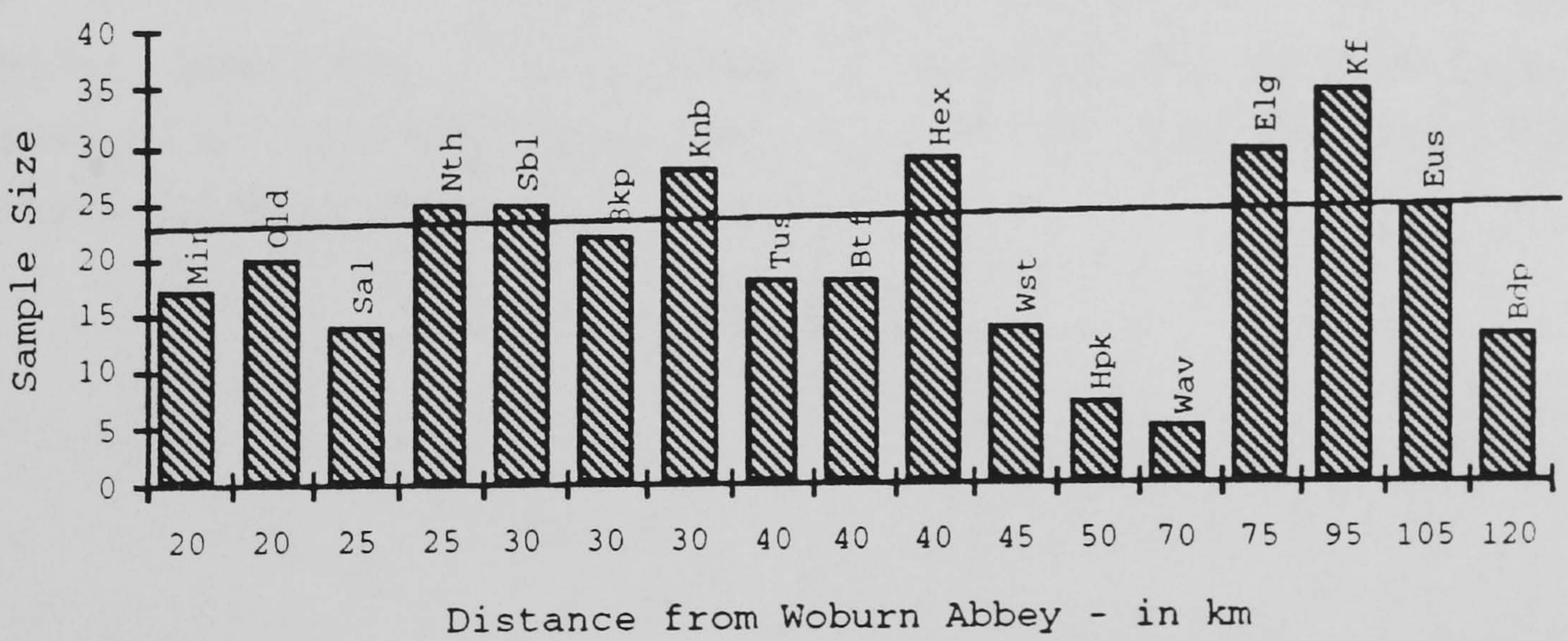


Figure 4·11 Relationship between distance from Woburn and sample size.

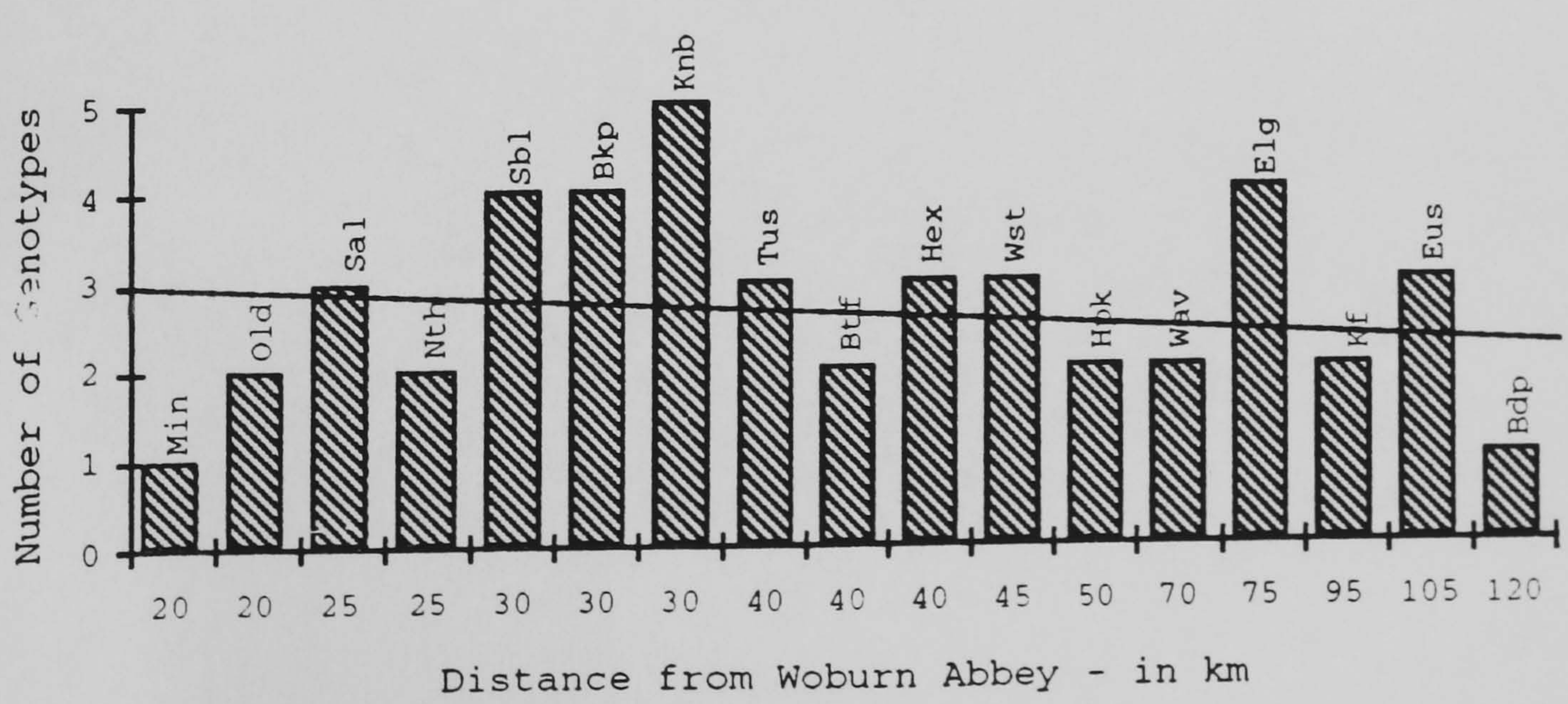


Figure 4·12 Relationship between distance from Woburn and number of genotypes.

### 4·3g Relationships Between Localities

Localities were clustered by parsimony analysis of genotypes per population, using the 'branch-and-bound' algorithm included in the programme PAUP (Phylogenetic Analysis Using Parsimony) version 3·0 (Swofford 1990). The frequency of occurrence of each genotype per population was calculated and the data entered in the 'binary' form from Table 4·2. Weighting was applied to each population in direct relationship to the frequency of occurrence of each genotype within the individual locality. All genotypes were originally (assumed) to be found at Woburn - the centre of origin. Therefore, for the purpose of constructing this tree I have assumed that all genotypes actually exist at Woburn, and have included them here in equal amounts. An implicit assumption is that there has been no mutational change since the original group of animals were imported. An assumption that may be justified by referring to section 4·1e. The tree search produced two trees that differed only in the alternative positioning of two localities. A consensus tree was constructed from these two trees, and is presented in Figure 4·13. The PAUP input file may be found in appendix $C_1$.

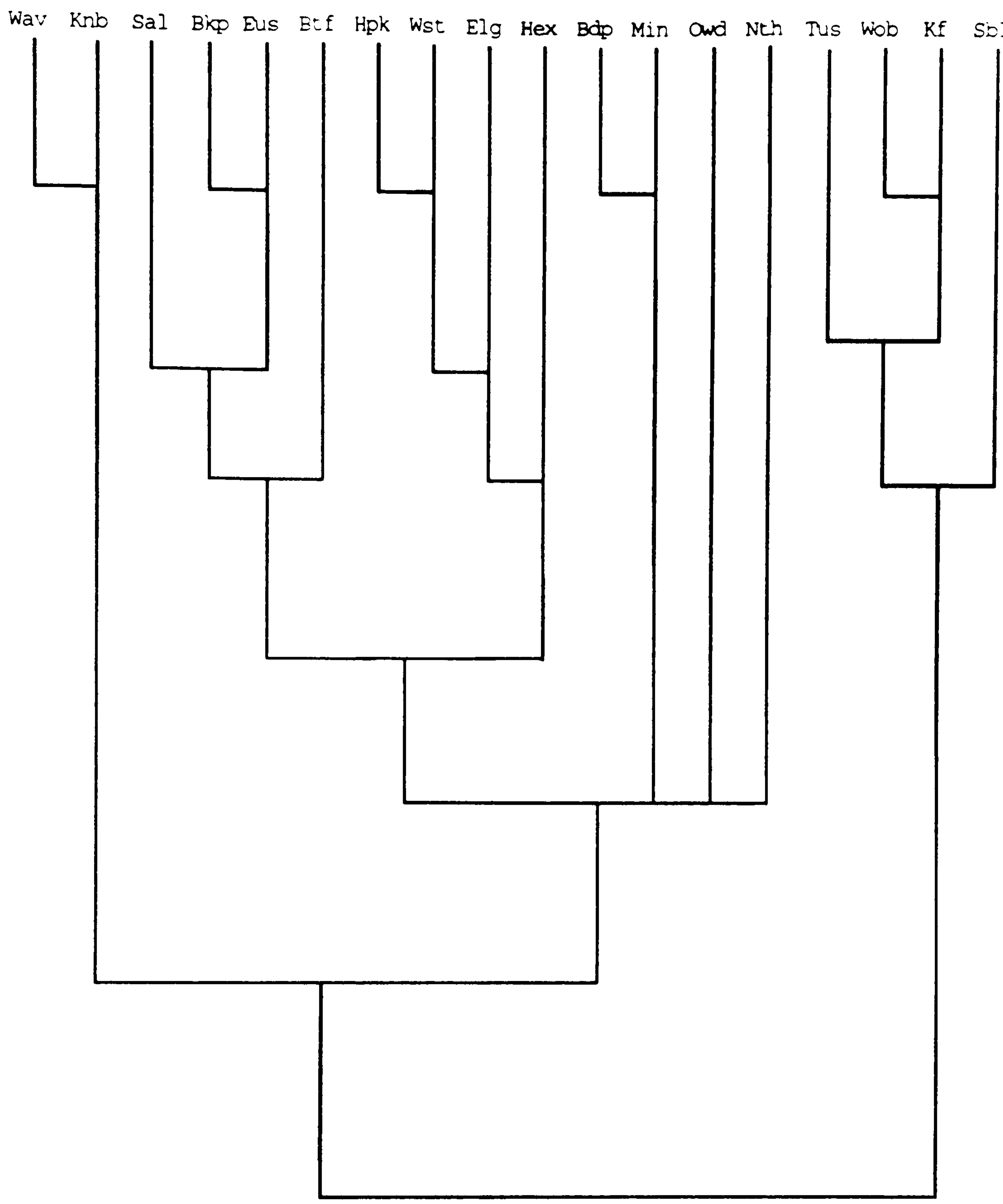


Figure 4·13 Consensus Tree from mtDNA control region PAUP Output.

From Figure 4·13 it can be seen that the localities fall into five clusters. These cluster are, in order of genetic similarity [Sal-Bkp-Eus-Btf], [Hpk-Wst-Elg-Hex], [Bdp-Min-Owd-Nth], [Wav-Knb], [Tus-Wob-Kf-Sbl]. These groupings are interpreted in a geographical context in Figure 4·14.

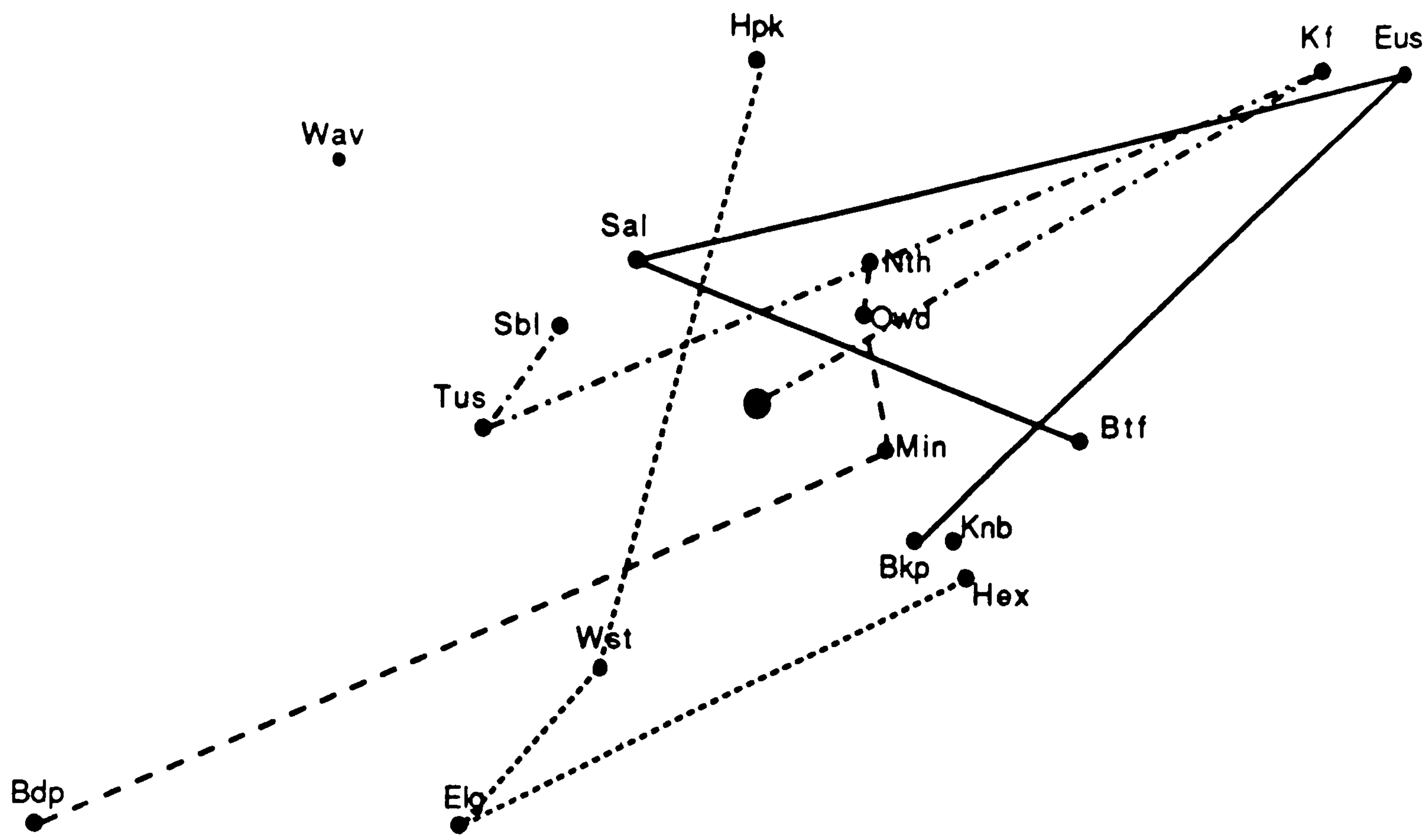


Figure 4·14 Diagrammatic representation of the four population clusters taken from the dendrogram in Figure 4·13. Different style lines are only used for clarity.

Localities in close geographical proximity are found to be genetically more distant than would intuitively be expected and locations far apart are genetically closer than would be expected. For example the groupings [Bkp-Knb-Hex] and [Kf-Eus] are geographically close (within 15km of each other, and within 10km respectively), but genetically distant. The groupings that are genetically close are generally geographically distant. Exceptions are [Tus-Sbl-Wob] and [Nth-Owd-Min] but each of these clusters contains a geographically distant element (i.e. Kf and Bkp respectively).
This approach may not be very satisfactory as undue weight may be given to rare genotypes. Unfortunately the rare genotypes only occurred in a few populations as can be seen from Table 4·3. CEF occurs once in Elg, ADG occurs once in Knb and Kf, and BEG occurs seven times but only in Sbl. Because of the infrequency of these rare genotypes they are of no value in defining inter-locality relationships, except perhaps for noting a connection between Knb and Kf - a connection not apparent from the consensus tree of Figure 4·13.

### 4·3h Sub-Population Differentiation

It may be expected that as sub-populations become geographically further separated a greater level of genetic differentiation would become apparent, assuming more or less constant rates of migration. This principal, known as 'isolation by distance' can be demonstrated to occur if there is a positive correlation between some measure of genetic distance and geographic distance.

Such a measure of genetic differentiation may be derived from Nei's (1987) coefficient of gene differentiation. This coefficient is given as $G_{ST} = D_{ST}/H_T$; where $H_T$ is the gene diversity in the total population and $D_{ST}$ is the inter-population gene diversity. From this equality can be derived the following:

$$G_{ST} = 1/[1+4N(s/s\text{-}1)(m + n)] \qquad 4{\cdot}1$$

where $m$ = migration rate, $s$ = number of sub-populations, $n$ = mutation rate, N = effective size of sub-population

If $s$ is large, and $m$ is very much greater than $n$, then $G_{ST}$ becomes approximately equal to $1/(4Nm + 1)$ (Nei 1987). An estimator of the migration rate between sub-populations can now be obtained as:

$$Nm = [1/G_{ST} - 1\,]/4. \qquad 4{\cdot}2$$

Pairwise values of $G_{ST}$ for all sub-populations were calculated using BIOSYS (Swofford & Selander 1981) and the data transformed to values of N$m$. The results are presented in Table 4·4 with values of $G_{ST}$ above the diagonal and values of Nm below the diagonal. Scatter diagrams were constructed to investigate the correlation between N$m$ and geographic distance from each sub-population to all other sub-populations, and for all sub-populations together using the program STATVIEW (Feldman *et al* 1987). Regression coefficients ($r^2$) were calculated for each scatter plot and the regression line included in the plot. The scatter plots are displayed in Figure 4·15, a-s.

$G_{ST}$ above the diagonal, N*m* below the diagonal

| | Wav | Tus | Sbl | Sal | Wob | Hpk | Bdp | Elg | Wst | Owd | Nth | Min | Bkp | Knb | Hex | Btf | Kf | Eus |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav | | 0.063 | 0.024 | 0.002 | 0.037 | 0.066 | 0.053 | 0.037 | 0.035 | 0.109 | 0.046 | 0.053 | 0.023 | 0.016 | 0.039 | 0.017 | 0.122 | 0.017 |
| Tus | 3.693 | | 0.015 | 0.042 | 0.028 | 0.028 | 0.099 | 0.06 | 0.043 | 0.055 | 0.094 | 0.099 | 0.028 | 0.05 | 0.07 | 0.055 | 0.065 | 0.03 |
| Sbl | 10.26 | 16.42 | | 0.012 | 0.017 | 0.015 | 0.049 | 0.023 | 0.015 | 0.04 | 0.044 | 0.049 | 0.006 | 0.029 | 0.029 | 0.017 | 0.077 | 0.009 |
| Sal | 124.8 | 5.702 | 20.58 | | 0.026 | 0.048 | 0.049 | 0.03 | 0.025 | 0.088 | 0.043 | 0.049 | 0.014 | 0.013 | 0.033 | 0.014 | 0.103 | 0.01 |
| Wob | 6.507 | 8.679 | 14.46 | 9.365 | | 0.041 | 0.082 | 0.045 | 0.038 | 0.069 | 0.076 | 0.082 | 0.014 | 0.032 | 0.059 | 0.039 | 0.029 | 0.005 |
| Hpk | 3.554 | 8.679 | 16.42 | 4.958 | 5.848 | | 0.044 | 0.021 | 0.014 | 0.011 | 0.042 | 0.044 | 0.017 | 0.081 | 0.027 | 0.034 | 0.105 | 0.033 |
| Bdp | 4.467 | 2.275 | 4.852 | 4.852 | 2.799 | 5.432 | | 0.007 | 0.013 | 0.076 | 0 | 0 | 0.033 | 0.105 | 0.003 | 0.024 | 0.193 | 0.053 |
| Elg | 6.507 | 3.917 | 10.62 | 8.083 | 5.306 | 11.65 | 35.46 | | 0.002 | 0.049 | 0.005 | 0.007 | 0.012 | 0.073 | 0.002 | 0.014 | 0.133 | 0.025 |
| Wst | 6.893 | 5.564 | 16.42 | 9.75 | 6.329 | 17.61 | 18.98 | 124.8 | | 0.042 | 0.011 | 0.013 | 0.008 | 0.064 | 0.004 | 0.013 | 0.121 | 0.021 |
| Owd | 2.044 | 4.295 | 6 | 2.591 | 3.373 | 22.48 | 3.039 | 4.852 | 5.702 | | 0.075 | 0.076 | 0.044 | 0.123 | 0.058 | 0.065 | 0.128 | 0.064 |
| Nth | 5.185 | 2.41 | 5.432 | 5.564 | 3.039 | 5.702 | 0 | 49.75 | 22.48 | 3.083 | | 0 | 0.029 | 0.095 | 0.002 | 0.021 | 0.184 | 0.04 |
| Min | 4.467 | 2.275 | 4.852 | 4.852 | 2.799 | 5.432 | 0 | 35.46 | 18.98 | 3.039 | 0 | | 0.033 | 0.105 | 0.003 | 0.024 | 0.193 | 0.053 |
| Bkp | 10.62 | 8.679 | 41.42 | 17.61 | 17.61 | 14.46 | 7.326 | 20.58 | 31 | 5.432 | 8.371 | 7.326 | | 0.035 | 0.018 | 0.009 | 0.076 | 0.005 |
| Knb | 15.38 | 4.75 | 8.371 | 18.98 | 7.563 | 2.836 | 2.131 | 3.175 | 3.656 | 1.783 | 2.382 | 2.131 | 6.893 | | 0.08 | 0.037 | 0.092 | 0.024 |
| Hex | 6.16 | 3.321 | 8.371 | 7.326 | 3.987 | 9.009 | 83.08 | 124.8 | 62.25 | 4.06 | 124.8 | 83.08 | 13.64 | 2.875 | | 0.014 | 0.158 | 0.035 |
| Btf | 14.46 | 4.295 | 14.46 | 17.61 | 6.16 | 7.103 | 10.17 | 17.61 | 18.98 | 3.596 | 11.65 | 10.17 | 27.53 | 6.507 | 17.61 | | 0.124 | 0.019 |
| Kf | 1.799 | 3.596 | 2.997 | 2.177 | 8.371 | 2.131 | 1.045 | 1.63 | 1.816 | 1.703 | 1.109 | 1.045 | 3.039 | 2.467 | 1.332 | 1.766 | | 0.056 |
| Eus | 14.46 | 8.083 | 27.53 | 24.75 | 49.75 | 7.326 | 4.467 | 9.75 | 11.65 | 3.656 | 6 | 4.467 | 49.75 | 10.17 | 6.893 | 12.91 | 4.214 | |

Table 4·4 Pairwise values for $G_{ST}$ and N*m*

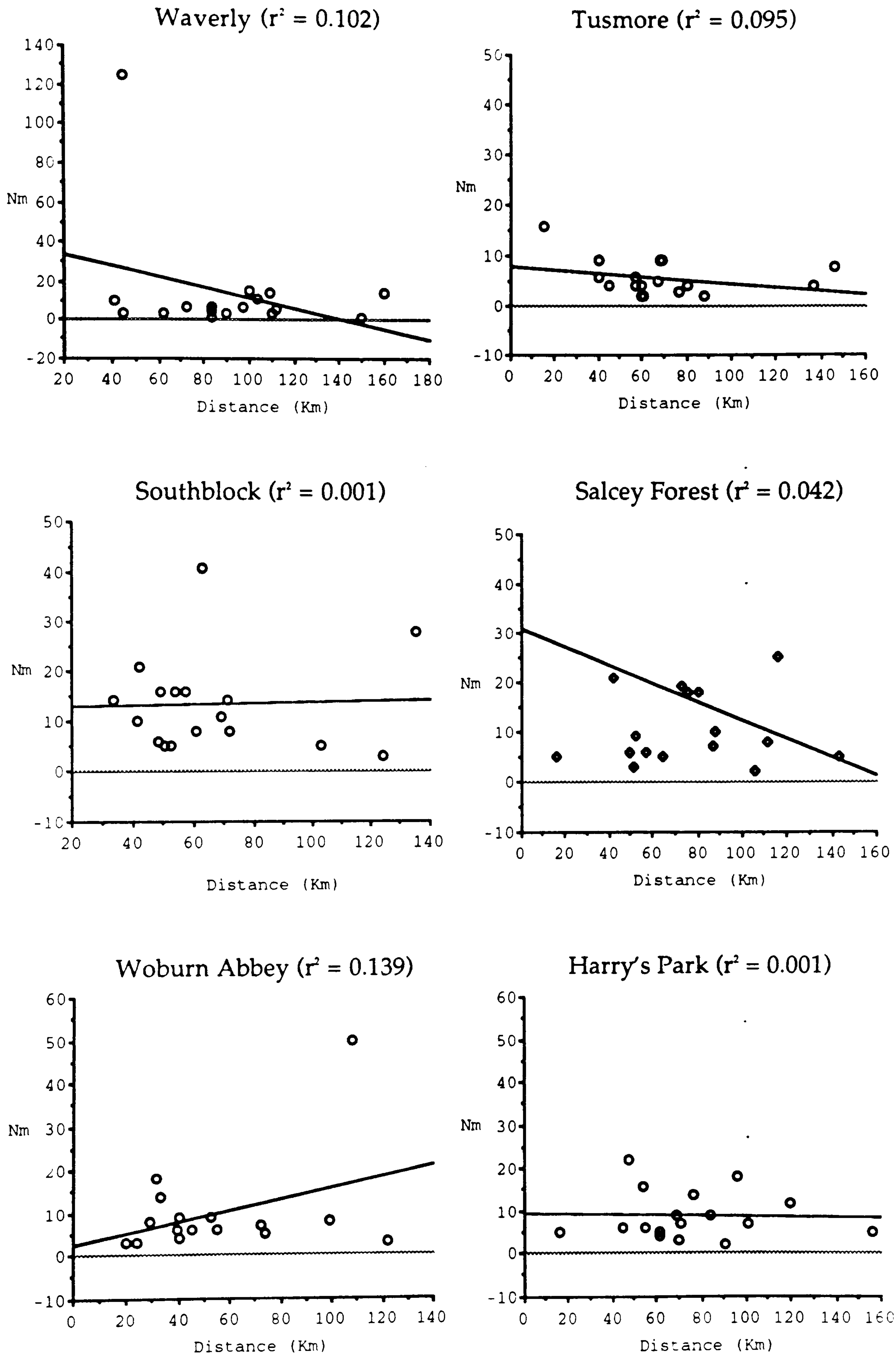


Figure 4·15a - f

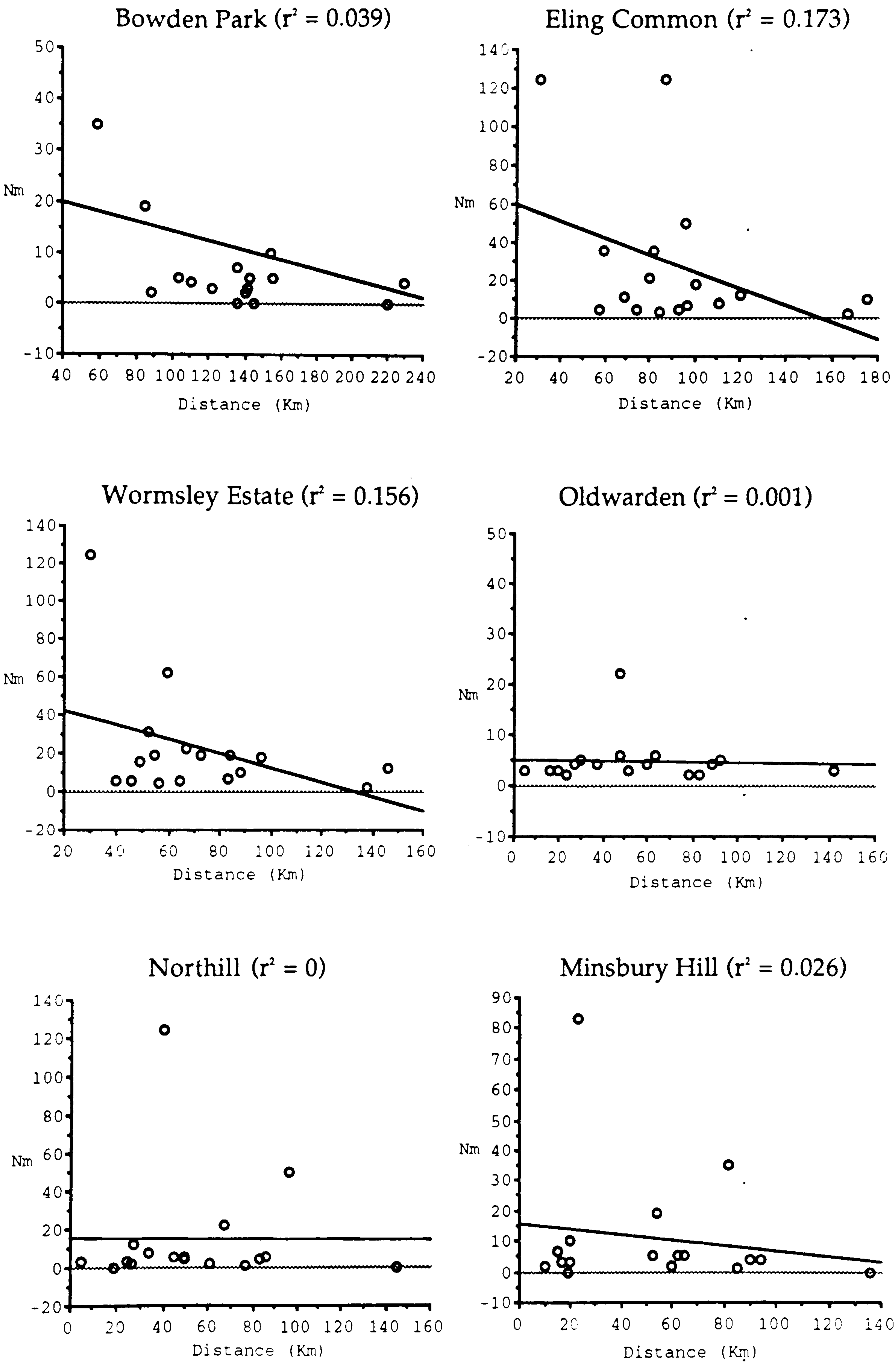


Figure 4·15g - l

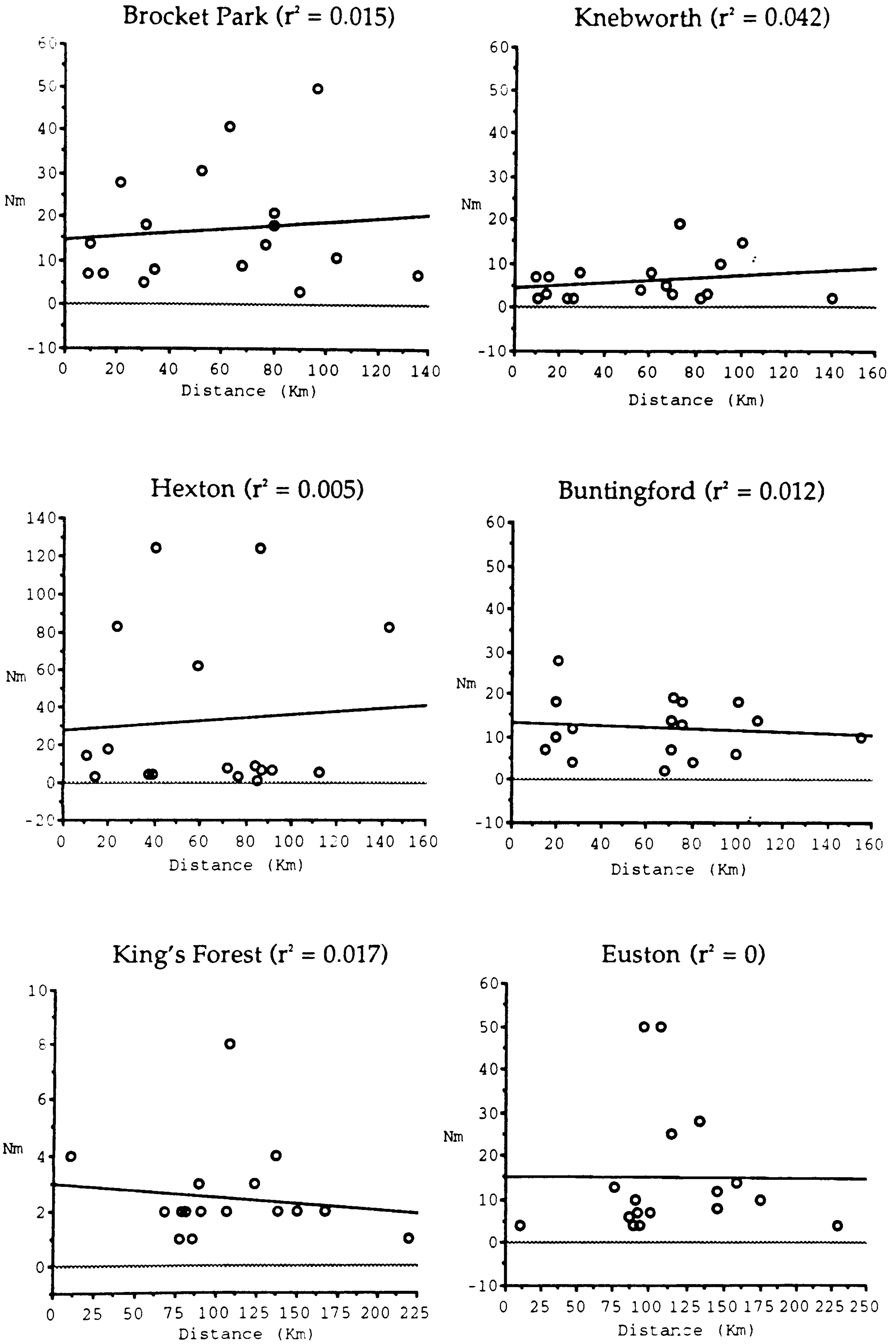


Figure 4·15m - r

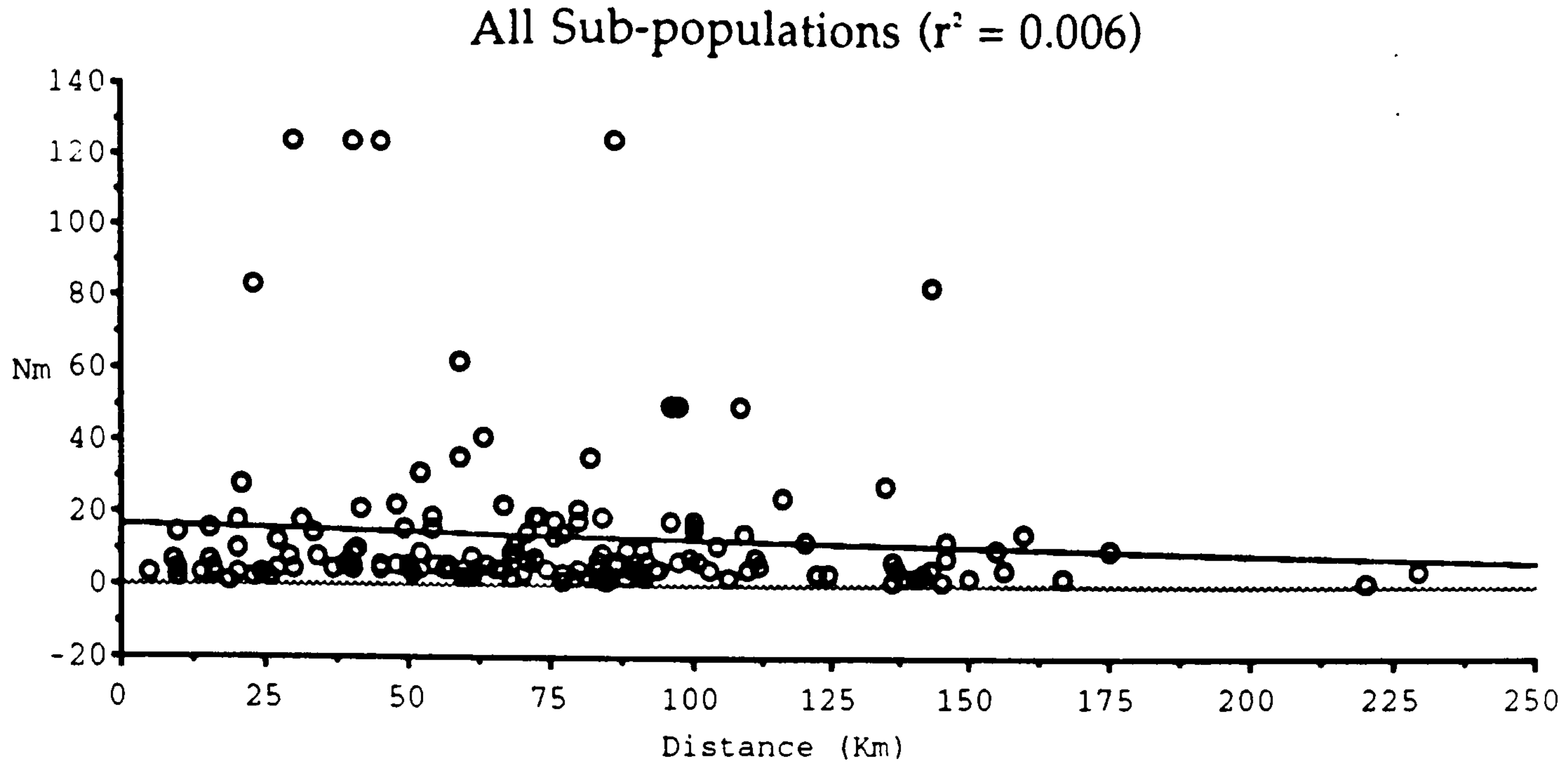


Figure 4·15s

Although the two sub-populations Eling Common (Elg) and Wormsley Estate (Wst) approach the significance level, none of the values of $r^2$ are significantly different from 0 at the 95% level. This result suggests that differentiation by distance is not evident.

According to Allendorf (1986); an average exchange of one individual per generation ($Nm \approx 1$) between populations, irrespective of deme size, is marginally sufficient to prevent genetic differentiation occurring by drift alone. The value of $Nm = 1$ corresponds to a mean $G_{ST}$ of 0·2 (Eq. 4·2). Thus the small values of $G_{ST}$ reported here (range 0 - 0·193, table 4·4) indicate a high rate of migration ($Nm$ range 1·045 - 124·8, table 4·4). The $G_{ST}$ values of 0 indicate contiguity ($Nm = \infty$).

Another approach to assessing the relationships between sub-populations is to construct a dendrogram based on Cavalli-Sforza's arc distance (Cavalli-Sforza & Edwards 1967). To calculate this metric the mitochondrial genotypes were designated AA (ADF), BB (ADG), CC (AEG), DD (BDG), EE (BDH), FF (BEG), GG (CDF), HH (CEF), and entered into BIOSYS program. The BIOSYS program was used to construct the distance-Wagner tree, Figure 4·16.

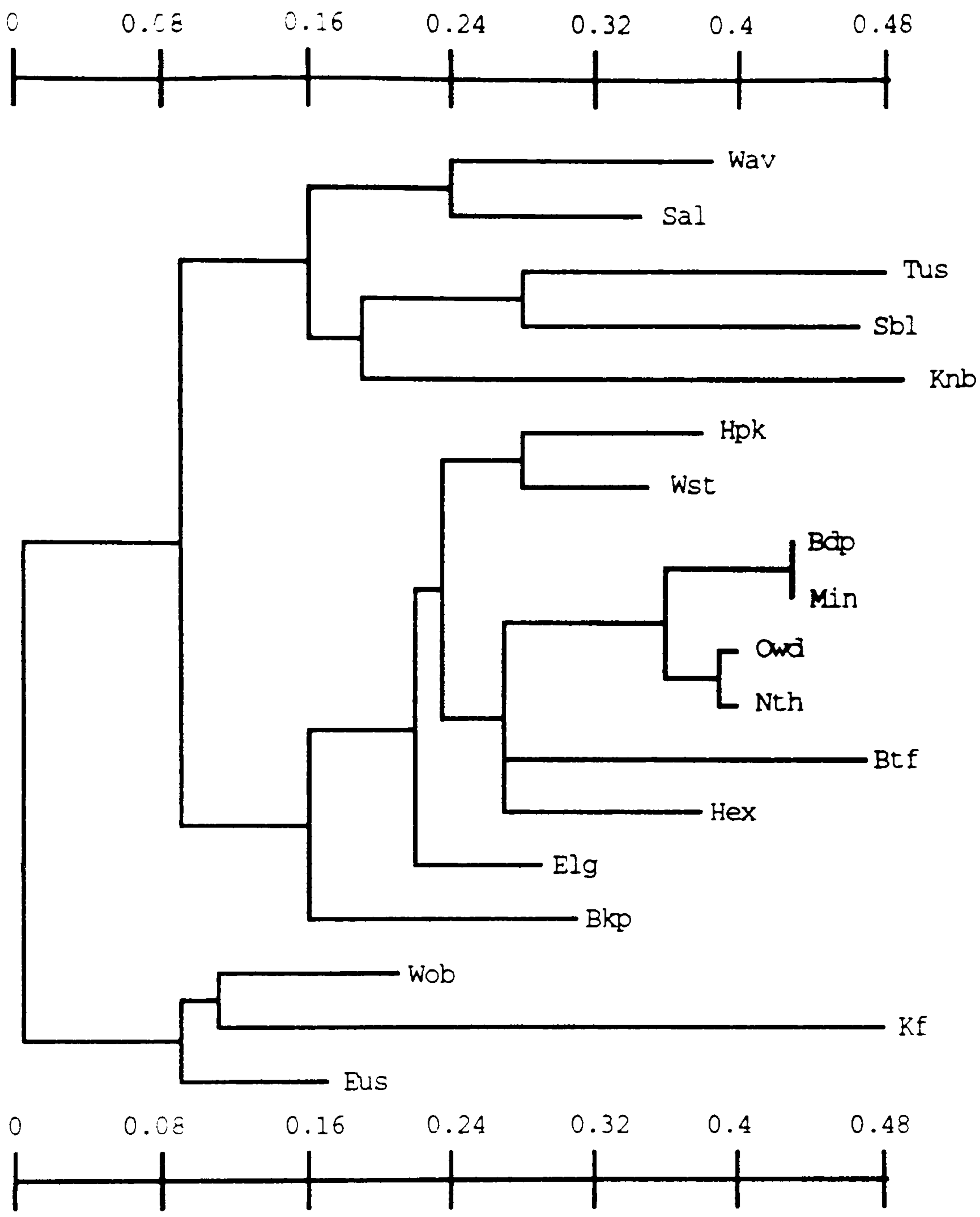


Figure 4·16 Distance-Wagner tree based on Cavalli-Sforza arc distance.

## 4·4 DISCUSSION

The purpose of this chapter is to demonstrate the use of mtDNA variation in an analysis of the genetic variation and relatedness between feral muntjac deer from different locations. The RFLP analysis of total mtDNA was rejected because of the problems of amassing suitable quantities of uncontaminated and undegraded mtDNA. A PCR based approach was then utilised, in which three regions ($ND_1$ + 16S rRNA, $ND_6$ and control region) were amplified and digested with various enzymes. A region of the control region was also directly sequenced.

Previous works of this nature have either used the entire mitochondrial genome and produced large numbers of restriction fragments by digesting with many restriction enzymes, or used sequence data from many individuals.

For example, Lansman *et al* (1983) in an investigation into genetic variation among populations of the deer mouse used eight restriction endonucleases on total mtDNA and uncovered a total of 61 genotypes (or 'restriction morphs') from 135 individuals. Wayne *et al* (1989) used 17 restriction enzymes to produce restriction profiles from total mtDNA from 68 individuals in a study of genetic divergence in sympatric canids. Thomas *et al* (1990), in an investigation of spatial and temporal continuity of extant kangaroo rat populations compared with museum specimens from the same localities, used the PCR to amplify the control region. Direct sequencing was then performed on 225 base pairs of the control region of 112 individuals. Lehman *et al* (1992) in a study of the genetic relationships within and among wolf packs from North America produced 32 composite mtDNA genotypes by digesting the total mtDNA molecule with 21 restriction enzymes.

In the work described in this chapter ten restriction enzymes were originally chosen, but only three of these were found to produce informative restriction fragment lengths. Nevertheless these three enzymes, applied to the control region were found to produce enough polymorphisms to identify eight genotypes.

Although a low degree of nucleotide and sequence conservation has been reported in the 2kb ($ND_1$ + 16S rRNA) and the 2·5kb ($ND_5$ + $ND_6$) regions (Anderson *et al* 1981, Attardi 1985, Johanssen *et al* 1990), the lack of restriction fragment variation found here was not surprising given the small founder event that led to the present U.K. muntjac population. However, some variation was expected and a possible explanation for this lack may be that even small changes in the DNA sequence are not tolerated - a situation which may be expected as these regions code for sub units of rRNA and sub units of the NADH complex. The ten restriction enzymes employed represent 46 base pairs, which is 2·3% of the 2kb region and 1·84% of the 2·5kb region. So alternatively, some variation may exist but not have been visualised by this limited investigation. However there is some evidence to suggest that mtDNA is effectively neutral (Clayton 1982, Forbes & Allendorf 1991b). If this is the case here the lack of variation must be due to some other cause. The U.K. muntjac population was effectively bottlenecked at the beginning of this century in that only a few individuals were imported to this country, and from these the present population has arisen. If the rate of evolution in these ORF regions is low and there was no variation among the individuals that formed the founding population then limited, or indeed no, variation would be expected. The rate of mitochondrial evolution is not constant across all species and has shown low rates of variation in some species although, in general, vertebrate mtDNA is thought to evolve at the rate of 1-2%/$10^6$ years (Brown 1985, Brown *et al* 1979, Wilson *et al* 1985)). In a comparison of other cervids Cronin (1991) discovered sufficient variation in total mitochondrial DNA to re-assess generic designations. A comparison of these ORF regions between muntjac from their native range as well as with those of closely related species, e.g. other members of the Muntiacinae and possibly extending to other members of the Cervidae may prove informative.

The amplification and restriction enzyme digestion of the control region uncovered sufficient variation to identify eight female lineages. In terms of the total number of individuals sampled (394), these genotypes were found to occur at frequencies ranging from 49% for the most common, to 0·25% for the most rare. In principal, patterns of migration and historical branching events can be reconstructed from a good set of gene frequency data. Chapman *et al* (1993) report that muntjac, both male and female, are territorial all year round although adults have been known to disperse from an area in which they have been established for several years (Chapman 1977). The doe's range is stable and there is commonly some overlap between adjacent doe ranges whilst the buck's range is larger and commonly encompasses several doe ranges. Harding (1986) found similar patterns in his study of muntjac. Young males tend to move away from their mother's range before maturity (i.e. before their pedicles have grown) but some are known not to disperse until much later (Chapman *et al* 1993), presumably when they are driven off by the dominant male. Females tend to remain in their mother's area and commonly establish overlapping ranges. Males are thought to be polygynous and some father-daughter mating may occur although there is as yet no evidence for this. From the above it can be seen that bucks are likely to disperse more readily than does. Assuming mtDNA to be only maternally inherited, then the matrilineal transmission will be parallel to the process of dispersal, but only of the females. On a fine scale the maternal birth site of one generation may be traced back to the birth site of the previous generation, and thus provide a model of the pattern of dispersal. However, as the dispersal of female offspring is limited in this species, and as the animals were taken from 'areas' rather than known home ranges only a general picture of the genetic relationships between sub-populations can be suggested

A common approach to determining the relationships between geographical data on gene frequencies, and the approach adopted here, is to apply clustering algorithms to a matrix of genetic distances. These algorithms usually result in dendrograms which give the appearance of confirming that the populations are hierarchically subdivided. These dendrograms must though be interpreted with caution as it may be that the genetic distances reflect isolation by distance, or some unsuspectedly complex pattern of gene flow, rather than a sequence of historical branching events. In the best possible situation, even the most complex of algorithms will produce no more than a crude approximation of the relationships between populations (Felsenstein 1982).

The two dendrograms in this chapter (Figures 4·13 and 4·16), produced by two different methods of interpretation, both demonstrate the great geographical distance between genetically closely related groups of muntjac. For this reason neither of the dendrograms can be interpreted as depicting a hierarchical branching mode of dispersal, but rather they may display a series of complicated founding events. One group of localities that are consistently associated are Northill-Old Warden-Minsbury Hill-Bowden Park. Northill and Old Warden are geographically very close and share the same two genotypes at the same frequency. Therefore they may be viewed as contiguous in terms of gene flow. Minsbury Hill is genetically almost identical to the previous two localities, differing only in that the rarer genotype (BDG) is not found there. The population at Minsbury Hill may therefore be seen as being very close to the populations at Northill and Old Warden. An anomaly exists here in that Minsbury Hill is geographically much closer to Knebworth (10km) than to Northill/Old Warden (20km), and the same distance from Brocket Park (20km), Woburn Abbey (20km) and Buntingford (20km). According to both investigative methods used here Minsbury Hill is genetically distant from Knebworth, Woburn Abbey, Brocket Park and Buntingford. This may be an artefact of the investigative techniques used as an inspection of Table 4·3 reveals that the samples taken from Minsbury Hill consist of only one genotype - the most common. This genotype is found, along with other genotypes in the near localities of Knebworth, Woburn Abbey and Brocket Park. If migration is occurring between Minsbury Hill and these other localities then it is probable that it is only occurring in one direction, i.e. away from Minsbury Hill.

Grouping these three localities, Owd - Min - Nth, can therefore be seen as reasonable as they are geographically close and could well be related by migration. However there is no reason to include Bowden Park (135km SW of Minsbury Hill) in this group other that this sub-population consists only of the genotype AEG (the commonest genotype, which is found in high frequency at the three localities Owd, Min and Nth). Bowden Park could well have been founded from Eling, which is geographically much closer, or indeed from any of the 16 other localities that possess this genotype, if we allow for human-mediated dispersal.

The geographically close localities Brocket Park, Knebworth and Hexton are all within 10km of each other and yet, according to two of the three investigative methods, are genetically distinct. An inspection of Table 4·3 shows that these sites have many genotypes in common - although at dissimilar frequencies.

Another intuitively unexpected result is that of the genetic distance between King's Forest and Euston, two localities only 12km apart. Although in the distance-Wagner tree (Figure 4·16) King's Forest is grouped with Euston - there is still a large (genetic) distance between them. Both these localities share a genotype (CDF) but Euston has two other genotypes (AEG, BDG) which do not appear at King's Forest, again leading to the supposition of directional dispersal.

The data resulting from this investigation, as summarised in Table 4·3 are difficult to interpret in terms of a dispersal from a centre of origin, especially when bearing in mind the limited time since the introduction of these animals to this country. However, it must be born in mind that these results could be the product of sampling variance from small populations. An alternative view of the data is that each occurrence of a genotype in a feral sub-population represents a founding event as a result of either a deliberate release or an escape from a local collection. This is probably the real situation in outlying localities such as King's Forest (which shares a rare genotype only with Knebworth) and localities in which only one genotype is evident (Minsbury Hill and Bowden Park).

In conclusion; the RFLP and sequence analysis of a small variable region of mitochondrial DNA, the control region, has provided enough information to determine the presence of genetically distinct groups of muntjac within a limited geographical region. The common approach to interpreting such data is that of agglomerative clustering. Two such techniques have been used here and although some differences are seen in the results, overall they are both very similar. Interpretation of these dendrograms should proceed with caution, and in this instance it is felt appropriate only to consider the initial clusters and pay little attention to the lower hierarchical levels.

Chapter 5

# THE ANALYSIS OF MICROSATELLITE DNA AS A MEANS OF ELUCIDATING POPULATION STRUCTURE IN FERAL MUNTJAC DEER

## 5·1 Introduction

Microsatellite DNA consists of relatively short runs of a tandemly repeated sequence of bases. The number of bases in a repeat unit is usually within the range of six bases or less (Beckman & Weber 1992). Also known as simple sequences, they have been found to be highly polymorphic and widely dispersed throughout the genomes of all eukaryotes so far investigated (Tautz & Renz 1984, Jarman and Wells 1989, Epplen *et al* 1991, Stallings *et al* 1991), but are under represented in the centromeric and telomeric regions (Wong *et al* 1990, Wintero *et al* 1992). The microsatellite region is generally less than 300 base pairs long and consists of four general types of repeat: mono nucleotides such as $(A)_n$*, dinucleotide repeats e.g. $(GT)_n$ , trinucleotide repeats e.g. $(GAC)_n$, or cryptically simple microsatellites (Tautz *et al* 1986) containing scrambled arrangements of the three former classes. Microsatellites are useful in a number of areas of genetic analysis, for example, as 'index markers' on chromosomes (Todd *et al* 1991), for the construction of genetic maps (Huang 1992, Kwaitkowski *et al* 1992, Weissenbach 1992) and for application to mapping and diagnosis of inherited diseases by linkage (Brook *et al* 1992). To date, variation at microsatellite loci has only been sparsely used in population studies and with a few exceptions most of these have been investigations into racial and ethnic differences in human populations (for example, Edwards *et al* 1992, Roewer *et al* 1993).

From a population genetics point of view the usefulness of microsatellites is twofold. Firstly because of the high degree of polymorphism found, which is a result of variation in the number of repeated sequences within a specific run, and secondly because these variations can be efficiently analysed using the polymerase chain reaction (PCR) and then visualised by a number of standard methods (Litt, 1991, Weber 1990a, b).

---

* For clarity and simplicity, repeat sequences will be written as the sequence in the 5′ to 3′ direction, in this manner $(A)_n$, $(GT)_n$, etc.

The existence of short repetitive sequences of DNA in eukaryotic genomes has been reported since the 1970s (Bruford & Wayne 1993), but it was not until the early 1980s that the extent and distribution of these elements was brought to prominence by Miesfeld *et al* (1981) and Hamada *et al* (1982). Miesfeld *et al* (1981), in an analysis of mouse ribosomal gene non-transcribed spacer (rDNA NTS) showed that homologous sequences are iterated and interspersed throughout the mouse genome. They went on to uncover a new class of interspersed repeat sequence in a human genome library by hybridising with the mouse rDNA NTS. Sequence analysis showed the reiterated region to be $(GT)_{17}$ and, after sub-cloning into a plasmid vector, found that this vector hybridised to DNA isolated from pigeon, *Xenopus*, slime mould and yeast as well as mouse and human. Hamada *et al* (1982), by Southern blotting and hybridisation analysis, uncovered a *'huge number of stretches of (TG) alternating sequence'*. The sequence was found to occur in all the eukaryote genomes they investigated (yeast, salmon, mouse and human); the number of copies of regions of the (TG) repeat ranging from about 100 in yeast to an estimated 200,000 in salmon. In an earlier paper of the same year Hamada and Kakunaga (1982) reported the presence of a sequence of 50 alternating (TG) residues within the intron of the human cardiac muscle actin gene, and estimated the number of $(TG)_n$ copies at $10^5$ per human genome (assuming an average size of $n = 25$). Continuing investigations by Tautz and Renz (1984) showed that (with the exception of (GC) and (AT), which were not shown to exist by the hybridisation technique, but were assumed to exist) all possible mono- and dinucleotides are present in eukaryotic genomes. They also demonstrated the same for the trinucleotide repeat (CAG).

### 5·1a *In Vivo* Functions of Simple Sequence Repetitive DNA

There have been several suggestions for the function of simple sequence DNA over the past ten years or so, but to date the true function or functions are still a matter of conjecture. Johnson and Morgan (1978) suggested that as $(\text{pyrimidine})_n \bullet (\text{purine})_n$ DNAs are known to form triple stranded polynucleotides, then the repeating unit $(TC)_{26}$ - uncovered between histone genes - may have a role in chromosome condensation. This tentative suggestion was based on structural features of the repeat region and modelled using known bond distances and angles. Two methods of folding and cross-bonding were proposed, (1) antifolding in which the chromosome folds simply back on itself and (2) synfolding in which the chromosome forms a loose spiral.

Kedes (1979) reported a (CT) repeat unit in the sea urchin *Strongylocentrotus purpuratus* and *Drosophila* within one of the five spacer regions within the histone gene sets. There are five major histone proteins, the genes for which are tandemly repeated and interspersed with one another. Located between each of the coding regions is a region of 'spacer' DNA. Kedes suggested that if recombinational events depend on base pairing then this simple sequence region will anneal at a more rapid rate than any other segment. In this situation, re-annealing between (CT) and (GA) regions would almost always be the first nucleation event of duplex formation. This (CT/GA) region could therefore be seen as being responsible for minimising out-of-register unequal crossover, while promoting in-register unequal crossing over as a gene correction-amplification mechanism. In the same paper Kedes also suggested that internally located repetitive sequences may be a common feature of tandem multigene families.

Hamada and Kakunaga (1982) suggested that *if* Z-DNA (a left-handed double-helical conformation of DNA first observed in molecules having alternating guanine/cytosine bases) occurs naturally, as in $(TG)_n$ then this Z-DNA may have differential reactivities with molecules such as proteins, and may be important in regulating DNA functions - for example in mutagenesis, recombination and gene expression. Russell *et al* (1983) found that a mutant form of one of the three loci in yeast (*Saccharomyces cerevisiae*) that encode different alcohol dehydrogenase isozymes (ADR2, which encodes the glucose-repressed isozyme ADHII) exhibited enhanced expression. This mutant form was shown to have an increased length in the $(A)_n$ tract 222 base pairs up-stream of the ADR2 gene. The increase was from $(A)_{20}$ to $(A)_{54}$ and $(A)_{55}$. Russell *et al* (1983) suggested that this increase in the $(A)_n$ region may have led to the constitutive expression of the gene by causing disruption of chromatin structure such that the entry of RNA polymerase enzyme was facilitated - and thus gene expression enhanced.

In an investigation of $(CA)_n$ repeat sequences Rogers (1983) suggested that these repetitive regions were insertions because many of them (but not all) are flanked by terminal repeats of six to eight base pairs. For example, from the mouse immunoglobin CγH-Cγ3 intron CCTAACTA - $(TTCCT)_{29}$ $(CA)_{30}$ $(GA)_4G$ - CCTAACTA. Terminal repeats are taken as *prima facie* evidence of insertion at a staggered break. Further evidence that some $(CA)_n$ blocks are insertions may be found in Gebhard and Zachau (1983). In man, substantial $(CA)_n$ blocks may mark the boundaries of gene conversion events, for example, between globin genes (Slightom *et al* 1980) and between immunoglobin Cα genes (Rogers 1983). Stringer found that an insertion of simian virus 40 into the rat genome involved recombination between $(CA)_n$ blocks. The above suggests (plus other examples - see Rogers 1983) that these blocks of $(CA)_n$ repeats are sites or potential sites for strand exchange between chromosomes, or regions susceptable to insertion. Comparison of sequence similarities and differences between allelic $^{A}\gamma$ genes and non-allelic $^{G}\gamma$ and $^{A}\gamma$ genes led Slightom *et al* (1980) to suggest that the simple sequence region was a preferred site for exchange leading to gene conversion-like events. Their data indicated that the $^{G}\gamma$ and $^{A}\gamma$ globin genes on a given chromosome can exchange DNA sequences by a recombinational event between chromosomes as well as within a single chromosome.

Arrays of simple sequences have also been associated with the evolution of new genes. Ohno and Epplen (1983) suggested this function for a 1256bp region of cloned mouse transcript containing many copies of the tetrameric repeat (TATC). Southern (1972) has also suggested a model for the evolutionary fate of repetitious base sequences.

### 5·1b Development and Evolution of Simple Sequence Repetitive DNA

Simple sequence repetitive DNAs are characterised by several features such as (1) they typically consist of a range of repeated motifs (e.g. Tautz and Renz, 1984). (2) Long repeat units often include shorter, internal repeat units (e.g. Rogers 1983), (3) Long polypyrimidines and poly CA tracts are often found (e.g. Jeang and Heyward 1983) and (4) tandem arrangements of closely related motifs are often found (Epplen *et al* 1983).

Although the occurrence and identification of simple repetitive DNA has been widely reported, its function, origins, and origins of diversity remain speculative. Of considerable interest is the range of diversity of these sequences, regions of simple sequence DNA may differ markedly in their organisation, length and base composition. A variety of simple motifs such as $(CA)_n$ and polypyrimidine tracts occur repeatedly in simple sequence DNA in diverse contexts. (Levinson and Gutman 1987a). The slightly more complex motifs $(GATA)_n$ and $(GACA)_n$ are thought to have evolved independently (Levinson *et al* 1985) on the grounds that they are both found in the diverse genomes of *Drosophila*, snakes (*Bungarus* and *Elaphe*) and mice (*Mus*).

The two most commonly suggested mechanisms by which polymorphism arises in simple sequence DNA are those of the random occurrence of unequal cross over and slippage during replication. Both these mechanisms, if they acted randomly, would lead to the constant formation and deletion of simple sequences and thus, in the absence of any selective pressure, the net effects would range from very little change from the original state to 'runaway' increases in number of repeat units in a region, possibly due to additions resulting from unequal crossing over. However there is evidence for selective pressure on some of the very large departures from 'normal' (i.e. non-deleterious) length of simple sequence DNA. The Fragile X Mental Retardation Syndrome - the most common form of inherited mental retardation and second most frequent to Down's Syndrome - has been strongly linked to increased length mutation in a $(CGG)_n$ repeat found within the coding sequence of the FMR-1 gene (Fu *et al* 1992). Similarly a $(CTG)_n$ repeat, when increased substantially beyond the population norm of n = 5 - 27, has been strongly implicated in the onset of Myotonic Dystrophy (DM) - an autosomal dominant neuro-muscular disease and the most common muscular dystrophy in adults (Brook *et al* 1992). DM patients with in excess of 80 sequence repeats were found to have a strong correlation between a greater number of repeats and the severity of the disease. Duplication or triplication with subsequent slippage during DNA replication has been postulated to account for the generation of excessively large alleles (Brook *et al* 1992).

The occurrence of simple sequence DNA in eukaryotes is probably not a matter of evolutionary conservation but probably depends on several factors such as (1) the availability of possible simple sequence DNA formation sites, i.e. regions of DNA with no coding function, (2) the frequency of random or accidental amplifications, (3) the degree to which such repetitive sequences are tolerated within the genome (e.g. Brook *et al* 1992) and (4) the extent to which the mechanism of amplification and deletion cause the simple sequence DNA to spread amongst homologous chromosomes (Dover 1982).

Generally, there are two mechanisms postulated for the generation and development of simple sequence DNA. These are slipped strand mispairing and unequal cross-over. A third mechanism, suggested by Nikaido *et al* (1981), is that $(A)_n$ may arise by reverse transcription of poly A tails of mRNA followed by re-integration into the genome.

### 5·1c Evidence for Slipped Strand Mispairing - The Genesis of Simple Sequences

In a random sequence of bases there is a probability of 1.56% of obtaining a four base run of ACAC. Simple repeats that occur by chance may provide the origins from which longer runs can be developed by expansion due to slipped strand mispairing during replication of a DNA duplex. Using the ACAC example, a two base slippage in the 3′ - 5′ direction may result in the insertion of an AC repeat unit. The insertion resulting from the addition of an AC unit at the 3′ end of the elongating strand. Slippage in the 5′ - 3′ direction may result in the deletion of an AC unit. The deletion resulting from excision of the unpaired repeat at the 3′ end of the elongating strand. The mismatched regions could form single-stranded loops, which would then be targets for excision and repair. Once expanded, the short region of tandem repeats would provide an even more efficient substrate for slipped strand mispairing.

Observations by Streisinger and Owen (1985) support the above as they found that the frequency of spontaneous insertions and deletions in sequences of $(A)_4$ and $(A)_5$ increased when the length of each these runs was increased by single base. Schlötterer and Tautz (1992) have also shown that (*in vitro*) slippage is an important factor in the genesis of simple sequence DNA. *In vitro* synthesis of simple sequences from synthesised templates proceeds via slippage reactions that produce continuously unpaired free ends which can be filled by a DNA polymerase. In this investigation Schlötterer and Tautz (1992) made the assumption that DNA polymerase 1 does not form a stable complex with DNA ends which overlap by only one base pair. Furthermore, they demonstrated that the rate of growth is largely independent of the length of the DNA fragment being synthesised. The experimental systems of Streisinger and Owen (1985) and Schlötterer and Tautz (1992) are somewhat different so their results may not be in fact contradictory. Economou *et al* (1990) and Zuliana and Hobbs (1990) both report that the apparent slippage that occurs during PCR amplification of dinucleotide repeat sequences is significantly diminished as the repeat length increases.

Length dependence would indicate that the addition of bases due to slippage would have to be transferred throughout the region of synthesis before further nucleotides could be incorporated at the elongating ends (Schlötterer and Tautz 1992). Length independence would indicate that unpaired loops must develop in simple sequence regions. *In vivo,* simple sequences are flanked by non-repetitive regions that are unlikely to slip during replication, therefore the addition of bases to the simple sequence regions would cause transient single stranded loops to form - which should become the substrate for repair enzymes. Small single stranded loops in simple sequence regions seem to be structurally accommodated in a DNA helix (Fresco and Alberts, 1960). Frequent initiation of a repair process in simple sequence regions may therefore be thought to contribute to the variability of these regions. However, no somatic instability has been reported in the literature to date (other than in abnormal situations such as tumour growth see Gonzalez-Zulueta *et al* 1993), and so it seems unlikely that a repair process contributes to variation. Levinson and Gutman (1987b) have shown that in *E. coli,* the presence of repair enzymes in fact reduces the mutation rate.

Morral et al (1993a), as a result of their investigations into microsatellite markers for cystic fibrosis, suggested that microsatellite allele variability associated with three microsatellites within the CFTR (cystic fibrosis transmembrane regulator) gene was due to slippage and consequent mis-pairing. Amongst other confirmatory reports of microsatellite allelic differences as a result of slippage and mispairing are Smeets *et al* (1989), Luty *et al* (1990), Fuentes *et al* (1993) and Saha *et al* (1993).

Tandem repeat regions that contain motifs differing by a single change are often associated and may be contiguous (Epplen *et al* 1983, Levinson *et al* 1985, Beckman & Weber 1992, Hamaguchi *et al* 1993, Lagercrantz *et* *al* 1993). A transition (e.g. C substituted by T), a transversion (i.e. A or G substituted by C or T), an insertion or deletion could create new repeat units from existing ones. For example, 5′ - CCCCC - 3′ could, as a result of a transition, become 5′ - CCTCC - 3′, and subsequent slipped strand mis-pairings, depending on the slippage site, would produce 5′ - CCTCTCT... - 3′. Many long repetitious regions such as alternating polypyrimidine tracts (Straus & Birnbiom 1976, Dodd & Straus 1982, Heilig *et al* 1982, Sorge & Hughs 1982, Schmidt & Shen 1985, Schlötterer *et al* 1991, Estoup *et al* 1993) and alternating pyrimidine/purine tracts (Hamada *et al* 1984, Beckman & Weber 1992, Callen *et al* 1993, Estoup *et al* 1993, Hamaguchi *et al* 1993) are frequently observed. Interspersed DNA elements in the form of $(CA)_n$ are one of the most abundant human repetitive DNA families (Weber & May 1989). In examination of over 100 human $(CA)_n$ repeats Weber (1990a) found that perfect repeat sequences, i.e. those without interruptions in the runs of CA equalled about 64% of the total. Imperfect repeat sequences (those with one or more interruptions in the run of repeats) equalled 25%, and compound repeated sequences (those with adjacent tandem simple repeats of a different sequence) represented 11%. The existence of the type of repeat tracts mentioned above may be explained by the combination of base substitutions followed by slipped strand mispairing. As transitions are more likely than transversions it follows that it is more probable that existing pyrimidine and purine tracts will be maintained and that new simple polypyrimidine or polypurine tracts will arise by chance as transitions occur. There are reported to be approximately 50,000 copies of the simple sequence $(GT)_n$ dispersed within the human genome with a repeat size of n = 10 to 60 (Litt & Luty 1989). The common occurrence of $(GT)_n$ tracts (Hamada *et al* 1982, Estoup *et al* 1993, Gonzalez-Zulueta *et al* 1993, Hamaguchi *et al* 1993, Lagercrantz *et al* 1993) may be explained in the following way. Eukaryotic genomes contain 5-methylcytosine (5-mC) as a rare base. This base arises by postsynthetic modification of cytosine and occurs in animals predominantly in the dinucleotide CG (Cooper 1983). Methylated C residues are subject to deamination - causing transition from C to T (Coulondre *et al* 1978, Bird 1980, Razin & Riggs 1980, see also Bird & Taggart 1980); although there are some CG rich regions which do not show elevated levels of TG, for example the major histocompatibility complex (Max 1984). Razin and Riggs (1980) also report that ~90% of methylated C residues occur at 5′ - CG - 3′ nucleotides. Deamination would tend to increase the occurrence of 5′ - TG - 3′ (in the absence of repair) which may then enhance the random occurrence of tandem repeats and subsequent expansion by slipped strand mispairing. This scenario could also account for the rarity of the 5-mCG dinucleotide.

Slipped strand mispairing may also play a major role in repetitive DNA sequence evolution by generating large numbers of short frameshift mutations within simple tandem repeats. (Levinson & Gutman 1987b). Levinson and Gutman (1987b) found that the frequency of detectable frameshifts within a 40 base pair tract of (CA) was greater than 1% and changed more than linearly with length, as the frequency of detectable frameshifts within a 22 base pair tract was lower by a factor of four. Of the frameshifts sequenced in their study, 96% resulted from a deletion or insertion of one or two $(CA)_n$ repeat units. The most frequent events being two base pair deletion, two base pair insertion, 4 base pair deletion; in the ration of 18:6:1.

### 5·1d Evidence for Unequal Crossing Over - The Expansion of Simple Sequences ?

There is generally less evidence for unequal crossing over as a mechanism for generating tandem repeat sequence than there is for slipped strand mispairing. Unequal crossing over is widely viewed as a major feature in the generation and maintenance of multigene families as well as satellite DNA (Ohno 1970, Smith 1973, Anderson & Roth 1977, 1981, Kedes 1979, Kurnit 1979). However there are two constraints placed on unequal crossing over and those are (1) it is an interhelical event and can only occur during chromosome or sister chromatid alignment in cell division and (2) unequal crossing over will be limited by the total length of sequence available for pairing. Smith (1976) has, by argument and computer simulation, put the case for the evolution and maintenance of repeated DNA sequences by unequal crossing over. His simulations showed that repeat regions can, in the absence of selection, develop as a result of random accumulation of random mutations and random homology-dependent unequal cross over. Some evidence for the role of unequal cross over at regions of short or simple sequence comes from an investigation of the switch region of immunoglobin Cμ genes. Immunoglobin heavy chain genes comprise a family of variable region genes and constant region genes,

which are arrayed in several classes. During differentiation a specific gene is expressed as part of one class and at a later stage as part of a different class. This phenomenon, know as 'immunoglobin class switch', involves a unique recombination event that occurs in the region 5′ to the constant region genes (Nikaido *et al* 1981). Structural analysis of these 'switch' regions has shown them to be comprised of tandem repetitions of short unit sequence (Dunnick *et al* 1980). Characterisation of the complete switch region of one of the classes of regions (Nikaido *et al* 1981) by sequence determination and restriction enzyme cleavages found it to be composed of a simple tandem repetition of two types of five base pair repeat unit sequences, 5′ - GAGCT - 3′ and 5′ - GGGGT - 3′. These results support the proposal that recombination (in this region) is mediated by repetitive homologous short sequences.
Although there is very little direct evidence that unequal crossing over plays a major role in the development and maintenance of simple sequence DNA, the present consensus in the literature appears to be that large runs of simple repetitive DNA may function as sites of preferred crossing over during recombination (Slightom *et al* 1980, Nikaido *et al* 1981, Stringer 1982). The similarity of repeat regions may permit recombination between otherwise non-homologous chromosome regions - a situation which has been proposed for the switching region of immunoglobin genes (Nikaido *et al* 1981). Unequal crossing over may be dependent on the process of slipped strand mispairing to generate simple repeat sequences of sufficient size to predispose DNA to interhelical events. As a genomic region becomes increasingly simple and repetitive the probability that non-contiguous sequences mis-pair also increases. Such mispairings could lead to larger duplications, deletions, and other rearrangements such as the deletion of sequences between repeats (Livneh 1983, Owen *et al* 1983, Kunkel 1985).

So far, mutation, substitution and unequal cross over events have been considered to be occurring at random. If this was the case then a reasonable expectation would be that the 'forward' events (i.e. additions) would be more or less cancelled by the 'reverse' events (i.e. deletions). This is obviously not the case as there is a great profusion and variety of simple sequence regions in all the eukaryote genomes that have been investigated. Thus there must be some drive(s) tending to allow the progressive accumulation of simple repeat DNA, at least up to the point at which it becomes deleterious.

In summary; simple sequence DNA is thought to arise fortuitously by slippage-like events working on randomly occurring repetitive stretches of DNA (Levinson & Gutman 1987a). All types of microsatellite are thought to be variable as the variability does not seem to be limited to particular types of simple sequence. Microsatellite variability must therefore be the product of a general mechanism that occurs sufficiently frequently to maintain a high degree of polymorphism within a population, but not frequently enough to occur in successive generations (Tautz 1989).

**5·1e Evolutionary Rate of Change and Inheritance**

The mutational dynamics of microsatellites and microsatellite regions are, as yet, poorly understood and there have been only a few systematic studies on the extent of polymorphism within and between populations (Amos *et al* 1993a, Choudhary *et al* 1993, Taylor *et al* 1994, Gottelli *et al* 1994, Roy *et al* 1994). Slipped strand mispairing during DNA replication is thought to be the most likely mutational event at these loci (Levinson & Gutman 1987a) but it is unclear whether the allelic distributions so created are behaving according to the infinite alleles model or the stepwise mutation model. The infinite alleles model of Kimura and Crow (1964) assumes that new mutations are always different from extant alleles in a population whereas the stepwise mutation model of Kimura and Ohta (1978) implies the possibility of length convergence among alleles which are not identical by descent, a situation which may complicate comparisons between populations (Valdes *et al* 1993).

The mutation rate at microsatellite loci has been variously estimated to be between $10^{-3}$ and $10^{-6}$ (Dietrich *et al* 1992, Edwards *et al* 1992, Kwaitkowski *et al* 1992, Weissenbach *et al* 1992). Edwards *et al* (1992) in a study of four human population groups obtained indirect evidence of mutation rates between $2{\cdot}3 \times 10^{-5}$ and $1{\cdot}6 \times 10^{-6}$ at five loci. They also noted that the higher mutations rates appeared to be associated with greater numbers of tandem repeats in the core motif. Investigations by Morral *et al* (1993a), into cystic fibrosis transmembrane conductance regulator gene (CFTR) microsatellites determined that the evolution of 16 haplotypes associated with the most common cystic fibrosis mutation, coupled with the estimated mutation rate at these loci, suggested that this common mutation originated at least 53,000 years ago. Changes seen in two other less common mutations suggest a date of origin some 35,000 years ago. Although no new alleles were detected at the three CFTR microsatellites in more than 3,000 meiosis analysed, an estimated mutation rate of less than $3.3 \times 10^{-4}$ was calculated.

One of the highest mutation rates reported so far is that of $2{\cdot}7 \times 10^{-3}$ estimated during an analysis of rat fibroblast clones (Slagboom *et al* 1991), although, because of the experimental system, this figure is an average per DNA fragment rather than of microsatellites *per se*. Another mutation rate of the same order has been reported by Lazaro *et al* (1993). As a result of an analysis of an Alu-repeat polymorphic sequence (AAAT) located within an intron of the Neurofibromatasis type 1 gene, new alleles were found in one of the 70 families studied that displayed this disorder and three of the CEPH (Centre d'Etude Polymorphisme Humain) families. This gave an estimated mutation rate of $3{\cdot}6 \times 10^{-3}$. Bowcock *et al* (1993) have reported an observed spontaneous mutation rate of $1{\cdot}2 \times 10^{-4}$ during a study of polymorphic $(CA)_n$ satellites in a human chromosome 13 genetic library. Dallas (1992) estimated microsatellite mutation rates in three mouse microsatellites derived from an inbred strain (brother/sister matings). He arrived at figures of $4{\cdot}7 \times 10^{-4}$ and $1.2 \times 10^{-4}$. Kwaitkowski *et al* (1992) during construction of a GT polymorphism map of human chromosome 9 (9q region) identified two new spontaneous mutations, giving an overall observed spontaneous mutation rate for dinucleotide polymorphism of $4{\cdot}3 \times 10^{-4}$.

The mutation rates of $10^{-6}$ and $10^{-5}$ (Edwards *et al* 1992) were derived from *in vivo* studies whereas the higher rates of $10^{-4}$ (Kwaitkowski *et al* 1992, Morral *et al* 1993a, Bowcock *et al* 1993) and $10^{-3}$ (Slagboom *et al* 1991, Lazaro *et al* 1993) were derived from *in vitro* studies. The work of Dallas involved 20 generation inbred strains of mouse and the microsatellites were chosen for this study on the basis of having above average number of alleles so as to provide a bias towards a high mutation rate (Dallas 1992). Although such high mutation rates may be expected from repetitive DNA (Slagboom *et al* 1991) there may be some proportion of the higher rates that is due to an experimental system, estimates from *in vivo* studies may more accurately represent the somatic mutation frequency.

The high mutation rates noted here, and those reported elsewhere, could occasionally lead to problems for questions such as paternity exclusions but it would not lead to significant errors in statistical estimates of relatedness (Queller & Goodnight 1989).

**5·1f The $(GT)_n$ repeat.**

The most extensively studied dinucleotide to date is the $(GT)_n$ repetitive sequence. Studies of large numbers of $(GT)_n$ polymorphisms have indicated that 80% of these sequences (where $n \geq 20$) will be polymorphic with heterozygosity 0·7 (Weber 1990a). $(GT)_n$ repeats have already been found associated with a few hundred human genes (Saha *et al* 1993) and are becoming widely used for linkage analysis in humans (Weber 1990c). In the work reported here only $(GT)_n$ microsatellites have been used.

$(GT)_n$ repetitive sequences are widely dispersed in the genomes of all eukaryotes so far investigated, including yeast, fish, amphibians, insects and mammals (Nordheim & Rich 1983, Hamada & Kakunaga 1982, Hamada *et al* 1982, Estoup *et al* 1993). The estimated number of copies of $(GT)_n$ repetitive sequences in different eukaryotic species ranges from 100 in yeast, to 50,000 in man to 100,000 in mice to 200,000 in salmon (Hamada *et al* 1982). These repetitive sequences have the potential to form Z-DNA (Nordheim & Rich 1983) and it has been demonstrated that under certain conditions a $(GT)_n$ repetitive sequence can enhance transcriptional activity of genes in plasmid constructs (Hamada *et al* 1984). Nevertheless, any functional significance for $(GT)_n$ repeats remains largely speculative (Stallings *et al* 1991).

However, it has been hypothesised that these repetitive sequences may have a role in genetic recombination (Pardue *et al* 1987). A large amount of crossing over has been postulated to take place within $(GT)_n$ repetitive regions as there is a great amount of variation in the length of these sequences at the same locus in humans (Weber & May 1989). Amongst further evidence for this hypothesis are the two observations that (1) the genome of SV40 virus contains a short $(GT)_7$ repeat and this virus integrates into the rat genome near a GT repeat (Stallings *et al* 1991) and (2) the β-satellite regions of *Drosophila* polytene chromosomes are deficient in GT repeats and do not undergo recombination. Although consistent with this hypothesis, such evidence is of course only conjectural. Stallings *et al* (1991) found no evidence for a correlation between the number of $(GT)_n$ repeats and increased recombination. In general, the mouse genome undergoes less recombination than the human genome, yet contains a greater number of $(GT)_n$ repeats (Stallings *et al* 1991).

It is not known whether the positions of $(GT)_n$ repeats have been extensively conserved along the length of chromosomes. At the gross level Pardue *et al* (1987) found by *in situ* hybridisation that polytene chromosomes from distantly related *Drosophila* species had similar distributions of $(GT)_n$ repetitive sequences. Along the same lines, Braaten *et al* (1988) have found that similar restriction fragments containing rRNA genes from mouse, rat and man contained $(GT)_n$ repeat sequences. At a fine level Stallings *et al* (1991) made a comparative analysis of 31 homologous sequences containing $(GT)_n$ repetitive sequence from several mammals representing four orders. It was found that the positions of these repeats have been conserved between closely related species such as humans and other primates, and conserved to a lesser extent between more distantly related groups such as primates and rodents.

**5·1g Conservation of sequences flanking microsatellite regions.**

There is some evidence to date that some microsatellite flanking sequences are conserved across closely related species. Moore *et al* (1991) applied 48 sets of bovine primer pairs to total genomic DNA from sheep, horses and humans. Specific PCR products were obtained from 27 primer pairs (56%) with ovine DNA, 20 of which were polymorphic. Rather less were successful with equine DNA - 3 primer pairs produced monomorphic products, and no specific products were obtained using human DNA.

In a series of papers Buchanan, Swarbrick and Crawford (1991, 1992), Buchanan and Crawford (1992a-d), Swarbrick and Crawford (1992), Swarbrick Buchanan and Crawford (1991b, c), Swarbrick *et al* (1992) reported that various ovine microsatellite (unassigned autosomal) primer pairs derived from a *Sau* IIIA size selected library could be used to amplify loci in alpaca (*Lama pacos*), red deer (*Cervus elaphus*), goat (*Capra hircus*), cattle (*Bos taurus*) and horse (*Equus caballus*). All loci were polymorphic in sheep and the majority were polymorphic in the other species. Three of the ten primers noted above would not amplify in the other species. From the work of Schlötterer *et al* (1991) it would appear that diverse cetacean species are homologous for certain microsatellite flanking sequences. The PCR primers originally derived from the long-finned pilot whale (*Globicephala melas*, Delphinidae) were found to amplify loci from representatives of all major cetacean radiations. The level of conservation found here is much greater than that found in the similar study by Moore *et al* (1991), or the various reports by Buchanan *et al* (mentioned above). Such a level would indicate that either cetacean DNA evolves more slowly or that the radiation of modern whales occurred more recently than currently thought (Schlötterer *et al* 1991).

**5·1h Non-amplifying Alleles**

A recently discovered problem with PCR amplification of microsatellite loci is the existence of null alleles (see Phillips *et al* 1991a and Callen *et al* 1993). Null alleles have been recognised for some time for protein polymorphisms and more recently for VNTR markers (Chakraborty, *et al* 1992). In PCR amplification of microsatellite loci null alleles may arise when a mutation occurs in the primer site of the flanking region, and of such magnitude to prevent primer binding and thus preclude synthesis. In a paper by Callen *et al* (1993) in which 23 $(AC)_n$ repeat markers from human chromosome 16 were investigated in the 40 CEPH (Centre d'Etude du Polymorphisme Humain) families, seven markers were found in which null alleles were demonstrated (i.e. 30% of those investigated). One null allele was sequenced and found to be the result of an 8bp deletion within the priming site. Alleles have also been reported where there is inhibition of the PCR reaction (as opposed to prevention) resulting in faint bands (Phillips *et al* 1991b, Weber *et al* 1991). A paper by Koorey *et al* (1993) also reported the existence of a null allele but in this case sequencing of the hidden allele revealed a single base substitution (T replacing C) at the 3′-most position of the binding site for the original CA strand primer.

These null alleles may not be recognised when there is PCR product from the other homologue; quite often it is difficult to determine true homozygotes from radiograph band intensity alone. The result of the existence of a null allele will be to score a false homozygote, or if family studies were being pursued, to record non-paternity. In the absence of family studies the presence of a null allele can only be a matter of conjecture supported (maybe) by a heterozygote deficiency (Callen *et al* 1993). In a large sample a few individuals that were supporting null alleles, and therefore scored as homozygotes, would have very little effect on the overall measure of heterozygosity. For a deviation from the expected level of heterozygosity either there must be a large number of individuals supporting null alleles, or some other cause must be sought. Where null alleles are suspected they can easily be identified by using primers designed from a different area of the flanking region. A possible 'cause' for null alleles may be that the primers were designed from flanking regions that were in effect regions of cryptic simplicity (Tautz *et al* 1986) and thus some variation may be expected between individuals.

**5·1i Previous use of Microsatellites in Population Studies**

To date the most widespread use of microsatellites has been to aid investigations of inherited genetic disorders in humans (e.g. Richards *et al* 1991 - fragile X syndrome; Morral *et al* 1993a,b - cystic fibrosis; Dausse *et al* 1993 - familial hypertrophic cardiomyopathy; Mitchison *et al* 1993 - Batten disease) and to construct linkage maps of animal genomes (see Weissenbach 1992). As yet, microsatellite DNA has only been used infrequently as an adjunct to investigations of social structure/relationships or population structure. The most complete survey to date is that of Edwards *et al* (1992), who examined microsatellite variation of five trimeric and one tetrameric loci in four ethnic groups of humans in the USA. The patterns of allele frequency distribution were found to be quite different at different loci. In four out of the five loci the distribution of allele frequencies was found to be complex and multimodal. At some loci the common allele appeared to be the same for all groups whilst at other loci the predominant allele did not coincide between populations. In some groups genotype proportions were found to deviate from Hardy-Weinberg expectations due to heterozygous deficiency, this may be due to inbreeding or difficulties in differentiating (and thus scoring) between heterozygotes and homozygotes, or possibly a combination of both. The presence of null alleles would also be scored as a heterozygote deficiency.

A similar study but on a smaller scale was conducted by Roewer *et al* (1993) into the familial structure of Yanomama indians settling in villages in southern Venezuela. Between these groups of villages there is a complex biological relationship as a result of village fissioning and changing pattern of alliances associated with inter-village antagonism. Polygyny is common (which tends to reduce effective population size), especially amongst leaders, who take additional wives, frequently from among the sisters of the first wife. Results confirmed the expected allele distribution for a population with a high incidence of consanguineous mating, i.e. a deficiency of observed heterozygotes at autosomal loci and the majority of men carried the same Y chromosome allele. Analysis of allelic distribution confirmed that one particular village was founded by closely related males from a local community and females that had migrated or been captured from other areas.

The results of the very few population studies that have been published on non-humans have shown high levels of microsatellite variation, and thus the usefulness of this technique. Amos *et al* (1993a) uncovered between three and 54 alleles per locus in long-finned pilot whales (*Globicephala melas*). The great variability of the latter locus permitted Amos *et al* to asses male mating behaviour within pods. They found that pilot whales are strongly matrifocal, with both sexes usually staying with their natal pods. Furthermore they found that adult males never fathered offspring from within their home pod - the implication being that pilot whales must mate when two or more pods meet, or when members of one pod temporarily exchange members with another. Gottelli *et al* (1994) surveyed nine microsatellite loci from Ethiopian wolves. The number of alleles varied from two to six, which is less than in other large canids (e.g. Roy *et al* 1994). However the microsatellite primers used were derived from a domestic dog library (Ostrander *et al* 1993) and a higher level of allelic diversity may have been uncovered if species specific primers had been used. This study detected diagnostic dog alleles in phenotypically abnormal wolves, strongly suggesting that Ethiopian wolves had hybridised with domestic dogs. Another interesting and important finding to emerge from this work is a definitive case of multiple paternity (assuming no germ-line mutation), one litter was fathered by both a dog and a wolf. In a similar study Roy *et al* (1994) analysed variation at ten microsatellite loci in order to quantify the extent of genetic variation among populations of wolf-like canids of North America (that is, the grey wolf - *Canis lupus*, the coyote - *Canis latrans* and the red wolf -

*Canis rufus*). Using primers designed from a domestic dog library (Ostrander *et al* 1993) they uncovered between four and twenty alleles per locus, which was sufficient to allow Roy *et al* to support the hypothesis that the intermediate phenotype of the red wolf has been derived from hybridisation between the grey wolf and the coyote.

Genetic variability in one of Australia's rarest animals, the Northern Hairy-Nosed (NHN) Wombat (*Lasiorhinus krefftii*) was investigated using sixteen simple sequence loci (Taylor *et al* 1994). The NHN wombat has suffered a drastic population decline over the past 120 years; consequently the levels of variation were found to be severely limited in comparison to an outbred closely related species, the Southern Hairy-Nosed (SHN) Wombat (*Lasiorhinus latifrons*). From the NHN wombat only one to three alleles per locus were found whereas in the SHN wombat between one and ten alleles were found at the same sixteen loci. The reduction in microsatellite variation suggested a bottleneck population of ten to twenty individuals for the 120 year period of their decline.

Microsatellite loci have been identified in social insects and have been found to be highly polymorphic, and thus of great value. For example, Choudhary *et al* (1993) explored the potential of microsatellite loci as genetic markers for investigating kinship patterns in the social wasp *Parachartergus colobopterus*. The five loci chosen were found to be far more variable than the allozymes previously used. Two trinucleotide repreats $(CAC)_n$ $(AAT)_n$ were found to have up to three alleles and the dinucleotide repeat $(GT)_n$ displayed up to 12 alleles and had a multimodal distribution. Hamaguchi *et al* (1993) found highly polymorphic $(GT)_n$ repeats in the polgynous ant *Leptothorax spinosior* and was able to use this polymorphism to elucidate maternity in multi-queen colonies.

Other reports of microsatellite loci being much more variable than allozymes are those of Evans (1993) and Hughes and Queller (1993).

In summary, the great abundance and variability of simple sequence DNA has, in almost every investigation, proved to be of notable value in providing greater understanding of reproductive success, mating behaviour, levels of population variation and substructuring, social structure and degrees of hybridisation. This level of information content coupled with the ease and rapidity of the polymerase chain reaction and the possibility of non-invasive sample collection makes microsatellite DNA analysis an important technique in many fields of investigation.

### 5·1j Analytical Techniques

There are many methods for estimating the degree of relatedness amongst groupings of animals. Cliff & Ord (1981) summarise many of these methods and Sokal and others (Sokal & Oden 1978a; Sokal 1979; Sokal & Wartenberg 1981) have introduced some of these methods to population biology. In this study the main parameters chosen to investigate population structure are those of Wright's $F_{ST}$, Nei's $G_{ST}$ and Spatial Correlation, as well as analysing heterozygosity, and the incidence of rare alleles.

The genetic variability of a divided population is usually measured by the average heterozygosity per locus. In a study of subdivided populations Wright (1943, 1951) showed that the variation in gene frequency may be analysed by the fixation indices, or F-statistics $F_{IS}$, $F_{IT}$ and $F_{ST}$. In summary, the F-statistics are inbreeding coefficients that differ according to the reference population. $F_{IS}$ measures inbreeding in individuals relative to their sub-population. $F_{IT}$ measures inbreeding in an individual relative to the total population, and $F_{ST}$ is a measure of inbreeding in a sub-population relative to the total population. These parameters are related by $1 - F_{ST} = (1 - F_{IS}) / (1 - F_{IT})$. F-statistics were designed in terms of one locus with two alleles; multi-allelic loci are treated by keeping the most common allele and pooling the other alleles.

In this analysis of the muntjac data, $F_{ST}$ is the parameter of most value and is used to give a measure of the level of sub-structuring between sub-populations. If all sub-populations are in Hardy-Weinberg equilibrium with the same allele frequencies then $F_{ST}$ will equal zero (i.e. no partitioning). If all sub populations are in the opposite state then $F_{ST}$ will equal 1 (i.e. indicating complete isolation of sub-populations).

Nei (1973) has pointed out that in the presence of multiple alleles the relationship between $F_{IT} / F_{IS} / F_{ST}$ is not valid, except in the special case of random differentiation with no selection (Nei 1965). Nei proposed an alternative method of measuring the degree of (gene) differentiation between pairs of populations. Nei's method is based on

the identities of two randomly chosen genes within and between sub-populations, and is independent of the number of alleles (Nei 1972). In a later paper Nei (1973) extended his method to include hierarchically structured populations and showed that the gene frequency variation in a sub-structured population can be analysed directly in terms of heterozygosity. Nei's measure of gene differentiation in a sub-divided population relative to the total population is given by $G_{ST}$ - the coefficient of gene differentiation (Nei 1973).
In an analogous way to Wright's $F_{ST}$, $G_{ST}$ is defined as $1 - G_{ST} = (1 - J_S) / (1 - J_T)$, where $J_S$ and $J_T$ are the average expected heterozygosity within sub-populations and the expected homozygosity in the total population respectively.

There are several different approaches to spatial correlation; the specific method I shall follow in this work is that of spatial autocorrelation as described in papers by Sokal and Oden (1978a, 1978b) and Oden and Sokal (1986).
If animals (or plants) migrate in a sequential fashion then there will be a genetic relationship between geographically close neighbours. As the groups of animals separate further the strength of this relationship will tend to diminish under the influence of many factors. But nevertheless a genetic relationship should still exist. More formally, if the data values collected at a particular geographic locality depend on values at neighbouring localities then the data are 'spatially autocorrelated' (Oden 1984). Spatial autocorrelation analysis therefore tests whether an observed value of a variable (in this case an allele frequency) at one locality is significantly dependent on values of the same variable at neighbouring localities. In summary, spatial autocorrelation can be broken down into four stages; (1) erection of a system of geographical (or other) interconnections between the localities sampled. The usual method, and the one employed here, is the Gabriel connected graph (Gabriel & Sokal 1969). In such a graph, any two locations, A and B, are considered to be contiguous if no other locality lies on or within a circle whose diameter is the line AB. This arrangement is just one of many possible patterns, the choice of which will depend on the origins of the data and the desired method of interpretation. For example, localities may be connected according to time-of-invasion or some defined cline rather than geographical adjacency. (2) Creation of an adjacency matrix,

taken from the Gabriel-connected graph, that describes the relatedness of localities in a pair-wise manner. From this matrix, the data are arranged in distance classes dependent upon the number of lines (or 'edges') connecting any two localities. (3) Calculation of Moran's statistic 'I' (Moran 1950), the autocorrelation coefficient for interval and ordinal data (Sokal & Oden 1978a). See also Cliff and Ord (1973) for a more detailed account. (4) Construction of a 'correlogram' - a graphical interpretation of the plots of Moran's 'I' as a function of interlocation distance (Kendal 1973). On this type of graph the abscissa shows the distance class and the ordinate, Moran's coefficient, 'I'. Positive autocorrelation indicates similarity amongst neighbouring localities (1 = maximum similarity), negative autocorrelation indicates difference (-1= maximum difference) and a mean of 0 indicates that the localities are no more or less similar than predicted by chance. In other words, if high values of the variate at one locality are associated with high values at neighbouring localities, the autocorrelation is positive. When high and low values alternate, the autocorrelation is negative.

## 5·2 MATERIALS AND METHODS

### 5·2a Initial Screening of the Loci and Choice of Primers

The PCR primers used in this part of the investigation came from a number of sources. Each locus was screened for polymorphism using ten randomly chosen individuals from the King's Forest population. All the following primer pairs were assessed but only a few proved to be informative (i.e. not monomorphic). The primer pairs and their origins are listed below in Table 5·1.

| | |
|---|---|
| MAF 18 | Ovine dinucleotide repeat |
| GT Strand | 5′- TTT TCC TTC AGA GCC AAG AGG ACA AG - 3′ |
| CA strand | 5′- GTA TAC TGC CTC TCT GCG TGA TGG G - 3′ |
| Ref: | Crawford, Buchanan & Swarbrick (1990) |
| MAF 23 | Ovine dinucleotide repeat |
| GT Strand | 5′- GTG GAG GAA TCT TGA CTT GTG ATA G - 3′ |
| CA Strand | 5′- GGC TAT AGT CCA TGG AGT CGC AG - 3′ |
| Ref: | Swarbrick, Buchanan & Crawford (1990). |
| MAF 33 | Ovine dinucleotide repeat |
| GT Strand | 5′- GAT CTT TGT TTC AAT CTA TTC CAA TTT C - 3′ |
| CA Strand | 5′- GAT CAT CTG AGT GTG AGT ATA CAG - 3′ |
| Ref: | Buchanan & Crawford (1992c). |
| MAF 35 | Ovine dinucleotide repeat |
| GT Strand | 5′- AGT TAC AAA TGC AAG CAT CAT ACC TG - 3′ |
| CA Strand | 5′- TCA AGA ATT TTG GAG CAC AAT TCT GG - 3′ |
| Ref: | Swarbrick, Buchanan & Crawford (1991a). |
| MAF 36 | Ovine dinucleotide repeat |
| GT Strand | 5′- CAT ATA CCT GGG AGG AAT GCA TTA CG - 3′ |
| CA Strand | 5′- TTG CAA AAG TTG GAC ACA ATT GAG C - 3′ |
| Ref: | Swarbrick, Buchanan & Crawford (1991b). |
| MAF 50 | Ovine dinucleotide repeat |
| GT Strand | 5′- GGG ACA TGC AGC TAT ACA CTT GAG - 3′ |
| CA Strand | 5′- GTA GAC TAC TCA TGA AAA TCA GGT CTT AGG - 3′ |
| Ref: | Swarbrick, Buchanan & Crawford (1992a). |
| MAF 65 | Ovine dinucleotide repeat |
| GT Strand | 5′- AAA GGC CAG AGT ATG CAA TTA GGA G -3′ |
| CA Strand | 5′- CCA CTC CTC CTG AGA ATA TAA CAT G -3′ |
| Ref: | Buchanan Swarbrick & Crawford (1992). |

| | |
|---|---|
| MAF 70 | Ovine dinucleotide repeat |
| GT Strand | 5′- CAC GGA GTC ACA AAG GAG TCA GAG C - 3′ |
| CA Strand | 5′- GCA GGA CTC TAC GGG GCC TTT GC - 3′ |
| Ref: | Buchanan & Crawford (1992b). |
| MAF 109 | Ovine dinucleotide repeat |
| GT Strand | 5′- TAA TTG AAT TTG AAG TGT ATA TGC CTA AAT GC - 3′ |
| CA Strand | 5′- GGA AGA TTA GAA CTT TCA TAT ATC TTT AAA CTC - 3′ |
| Ref: | Swarbrick & Crawford(1992). |
| MAF 209 | Ovine dinucleotide repeat |
| GT Strand | 5′- TCA TGC ACT TAA GTA TGT AGG ATG CTG C - 3′ |
| CA Strand | 5′- GAT CAC AAA AAG TTG GAT ACA ACC GTG G - 3′ |
| Ref: | Buchanan & Crawford(1992a). |
| MAF 214 | Ovine dinucleotide repeat |
| GT Strand | 5′- AAT GCA GGA GAT CTG AGG CAG GGA CG - 3′ |
| CA Strand | 5′- GGG TGA TCT TAG GGA GGT TTT GGA GG - 3′ |
| Ref: | Buchanan & Crawford (1992). |
| JP 15 | Cervid dinucleotide repeat |
| GT Strand | 5′-GGA TAT TTA ATA TAA GTT AGC AAT GAG AAA GAA GG-3′ |
| AC Strand | 5′- GGA AAT ACC TTA ATC TTT CAT TCT TGA CTG TGG - 3′ |
| Ref: | J. Pemberton pers. comm. |
| JP 23 | Cervid dinucleotide repeat |
| GT Strand | 5′- CCG CAG AAC AAV CTA ACG CCC AAG - 3′ |
| CA Strand | 5′- CCT TTG TCG CTT GGA TTT TC - 3′ |
| Ref: | J. Pemberton pers comm. |
| JP 27 | Cervid dinucleotide repeat |
| GT Strand | 5′- GAT CCC CTC CTT GTG CCA C - 3′ |
| CA Strand | 5′- GTC TGT GGG TCT ATT TCT GAT TTG C -3′ |
| Ref: | J. Pemberton pers comm. |
| JP 38 | Cervid dinucleotide repeat |
| GT Strand | 5′- GGC AAT ACA CTG GAA TAA TCT G - 3′ |
| CA Strand | 5′- CTG CAC AGA GTC GGA CAC AAC - 3′ |
| Ref: | J. Pemberton pers comm. |
| BOVIRBP | Bovid dinucleotide repeat |
| A Strand | 5′- TGT AGT ATC ACC TTC TAT GCT TC - 3′ |
| B Strand | 5′- GTC TTA GGT AAT CAT CAG ATA GC - 3′ |
| Ref: | derived by D. MacHugh - J. Pemberton pers comm. |

| | | |
|---|---|---|
| CYP 21 | | Bovid dinucleotide repeat |
| A Strand | 5′- GGA GGG TTA CAG TCC ATG AGT TTG - 3′ | |
| B Strand | 5′- TCG CGA TCC AAC TCC TCC TGA AG - 3′ | |
| Ref: | Fries, Eggen & Womack (1993). | |
| BOLA-DRBP1 | | Bovid dinucleotide repeat |
| A Strand | 5′- ATG GTG CAG CAG CAA GGT GAG CA - 3′ | |
| B Strand | 5′- GGG ACT CAG TCT CTC TAT CTC TTT G - 3′ | |
| Ref: | Fries, Eggen & Womack (1993). | |
| D1 S6 | | Bovid dinucleotide repeat |
| A Strand | 5′- CGA GTT TCT TTC CTC GTG GTA GGC - 3′ | |
| B Strand | 5′- GCT CGG CAC ATC TTC CTT AGC AAC- 3′ | |
| Ref: | Fries, Eggen & Womack (1993). | |
| FSHB | | Bovid dinucleotide repeat |
| A Strand | 5′- CTT GGG ATA TAG ACT TAG TGG CAT G - 3′ | |
| B Strand | 5′- GCA CAA GTC ACA GTT TCT AAG GCT A - 3′ | |
| Ref: | Fries, Eggen & Womack (1993). | |
| D21 S4 | | Bovid dinucleotide repeat |
| A Strand | 5′- GTG GAC TAT AGA CCA TAA GGT C - 3′ | |
| B Strand | 5′- GCT GTG ATG GTC TAC GAA TGA - 3′ | |
| Ref: | Fries, Eggen & Womack (1993). | |

Table 5·1 Primer Pairs Investigated

With the exception of those primers donated by J. Pemberton (BOVIRBP plus those prefixed 'JP'), all were synthesised on a Beckman DNA synthesiser and purified in accordance with the manufacturer's recommendations.

### 5·2b Choice of Method for Visualising the PCR Product

Two methods of visualising the results of amplification of microsatellite regions were tried, one 'cold' and one 'hot'. The 'cold' method involved visualising the PCR product by running it into an acrylamide gel and staining with ethidium bromide. This method proved too unreliable for a rapid analysis of a large number of samples. Only thirteen samples could be run per gel (plus two molecular weight markers) and with nearly 400 sample to assess it was thought more appropriate to utilise a larger gel - up to 60 samples at one time - and a more sensitive procedure ('hot') - that of end labelling one of the primers with $\gamma^{32}$P-dATP. Weber and May (1989) have demonstrated that by end labelling only one primer, the clarity of the resulting bands is greatly improved. This method is described below. Also there may some difficulty in resolving small differences (i.e. of two base pairs) by the former method.

The polymerase chain reaction was performed essentially as already described in Chapter 4 (4·2a) but with the following major change. Two rounds of amplification were initiated. The first round at a low stringency for seven cycles and the second round at a high stringency, but with a lower denaturing temperature (Yap & McGee 1991), for thirty cycles. The basic structure of this protocol was communicated by J. Pemberton.
DNA was taken from ten samples chosen at random from the King's Forest collection and diluted 1:10 with autoclaved MQ $H_2O$. This gave a range of concentrations across all samples of between 7ng/µl and 300ng/µl. To check for contamination, a negative control substituting autoclaved MQ $H_2O$ for DNA was carried out for every series of PCR reactions. If an amplification failed the concentration of DNA was reduced 10 fold, and 10 fold again if necessary. If the amplification continued to fail the concentration was increased two fold and two fold, etc. It was found that about 60% of the extractions required a 10 fold dilution and 20% of these required a further ten fold dilution whilst very few required concentration - regardless of the actual amount of DNA present. The primer concentrations in all reactions was 25pM.

### 5·2c Preliminary reaction

In all these reactions 1 µl of DNA was used from a 1:10 dilution, as above. The GT strand or 'A' primer was chosen as the primer to be labelled. All PCR reactions were done in a Hybaid Thermal Reactor TR1 and overlaid with two drops of mineral oil.

End-labelling reaction: to a 0·5ml eppendorf were added, in order:

| | |
|---|---|
| Autoclaved MQ $H_2O$ | 6·0µl |
| 10x Poly Nucleotide Kinase Buffer | 1·0 µl |
| Primer (25 pM) | 1·5 µl |
| $\gamma^{32}$P-dATP (10 µCi/µl) | 0·5 µl |
| Poly Nucleotide Kinase (10 units/µl) | 1·0 µl |
| | 10·0 µl |

The reaction was incubated at 37°C for 45 minutes and then the enzyme denatured by heating to 95°C for 5min. After centrifugation at 12,000rpm for 5 seconds the following ingredients were added:

| | |
|---|---|
| Autoclaved MQ $H_2O$ | 52·5µl |
| Dimethylsulphoxide (DMSO) | 5·0µl |
| $MgCl_2$ (50 mM) | 0·5 µl |
| *Taq* XL buffer | 10·0µl |
| dNTPs (1mM each) | 10·0µl |
| Unlabelled Primer (25mM) | 1·0 µl |
| *Taq* XL | 1·0µl |
| | 90·0 µl |

The reaction was mixed gently by pipetting and 9·0 µl added to a 500µl eppendorf containing 1µl of DNA.
Amplification Cycle:
After a preliminary denaturation step of heating to 93°C for 3 minutes the samples were run through seven amplification cycles which consisted of denaturing at 93°C for 30 sec, annealing at 50°C for 60 sec, extension at 72°C for 90 sec. These seven cycles were followed by thirty cycles of 89°C for 30 sec, annealing at 54°C for 60 sec, extension at 72°C for 90 sec. A final extension at 72°C for 10 minutes was added and the reaction allowed to cool to ambient.

**5·2d First Optimisation**

The above protocol was modified to produce a larger quantity of end-labelled primer per reaction, and lesser amounts of the various reagents were used in the PCR.

Primer end-labelling reaction:

| | |
|---|---|
| Primer | 4·0 µl |
| Autoclaved MQ $H_2O$ | 6·5µl |
| 10x PNK Buffer | 1·5 µl |
| $\gamma^{32}$P-dATP (10 µCi/µl) | 1·5 µl |
| Poly Nucleotide Kinase (10 units/µl) | 1·5 µl |
| | 15 µl |

The reaction was incubated at 37°C for 45 minutes and then the enzyme denatured by heating to 95°C for 5min. After centrifugation at 12,000rpm for 5 seconds the reaction was stored at -20°C. This provided enough end-labelled primer for 100 reactions.

Initially the amplification reaction was calculated according to this recipe:-

| | |
|---|---|
| Autoclaved MQ $H_2O$ | 6·215µl |
| (DMSO) | 0·5 µl |
| $MgCl_2$ (50 mM) | 0·05 µl |
| *Taq* XL buffer | 1·0 µl |
| dNTPs (1mM each) | 1·0 µl |
| Unlabelled Primer | 0·05 µl |
| Labelled Primer | 0·135 µl |
| *Taq* XL | 0·05 µl |
| | 9·0 µl |

and then increased to give sufficient mixture for 50 reactions.
The above reaction mixture were added to a 500µl eppendorf containing 1µl of DNA and subjected to the cycles of the PCR as described above.

**5·2e Second Optimisation**

This reaction proved to be successful and generally produced a strong signal (ca 500cps). The protocol was further modified to give a greater concentration of product to be loaded per lane, as well as reducing the concentration of (and thus exposure to) $\gamma^{32}P$.
On completion of the reaction, 2µl of stop solution were added and the reaction stored at -20°C. 2·5µl of each reaction were run out on a 5% Sequagel and visualised as described in Chapter 4 (4·2e). The timing of the electrophoresis run was determined empirically; aimed at locating the polymorphic band in the central region of the gel. This time varied between 1·5 and 2·5 hours.

**5·2f Preparation of Molecular Weight Marker (A-T Ladder)**

A standard sequencing reaction was performed using the -40 universal primer from the USB Sequenase kit (5′-GTTTTCCCAGTCACGAC-3′), single stranded M13mp18 (Sequenase kit) as template and reagents from

Clone-Tech's 'Multi-pol' sequencing kit. The manufacture's protocols were followed except that only the termination mixes 'A' and 'T' were used. Upon completion of the reaction 7 µl of stop solution and 3 µl $H_2O$ was added to each reaction tube. The reactions were pooled, briefly spun and stored at -20°C. This preparation gave a total volume of 60 µl, enough for 24 loadings. The reaction was heated to 90°C for 3 minutes prior to loading. 2·5µl of this 'AT' ladder was loaded in the flanking lanes.
From the published sequence of M13mp18, and in association with the location of the priming site (Sequenase Version 2 protocol book - 3rd Edition) it is possible to calculate the exact molecular weight of the adjacent microsatellite bands.

### 5·2g Scoring the Bands

In the finally optimised reaction each microsatellite produce one strong band per allele, one main band in homozygotes and two bands in heterozygotes. Commonly there were many weaker bands at intervals occasionally above, but generally below the main band. These weaker bands are caused by slippage events during the polymerase chain reaction (Hauge & Litt 1993) and therefore the interval is dependent upon the number of base pairs that form the repeat unit. Each microsatellite primer pair also produced a pattern of low molecular weight sub-bands, but these were remote from the bands of interest and did not affect the scoring. The bands were scored by eye from the autoradiographs by comparison with the molecular weight marker.

### 5·2h Sources of Error in the Polymerase Chain Reaction

A common source of confusion in interpreting the results of a polymerase chain reaction is the presence of amplified contaminating DNA. The PCR technique is extremely sensitive to very small amounts of DNA, theoretically a single molecule is sufficient to produce a scorable result, especially so in a visualising technique as sensitive as end labelling. The sensitivity is great enough to permit analysis of DNA from a single cell (Zhang *et al* 1992), a single hair (Vigilant *et al* 1989) or a single sperm (Liu *et al* 1992). To minimise the risk of contamination dedicated pipettes, sterilised solutions and plastic-ware were used at each stage of the protocol. To minimise the possibility of contamination occurring and remaining undetected, negative controls were always included.

Contamination was not detected in this or other sensitive techniques used.
Another PCR based problem is the possible presence of 'null' alleles. These are alleles which exist but are not amplified due to mutations in the priming site within the flanking regions. See discussion for comments on this problem.

**5·2i Analysis of simple sequence polymorphisms.**
Genetic polymorphism for each population was measured as the mean number of alleles per locus, observed heterozygosity ($H_O$) and that expected from Hardy-Weinberg assumptions ($H_E$); (Nei 1978, 1987). Deviations from Hardy-Weinberg equilibrium were tested using the chi-square test with pooling (Hartl & Clark 1988), because of the presence of rare alleles. The standardised variance in allele frequencies among populations, $F_{ST}$ (Wright 1969), was calculated using modifications described by Nei (1977). Allele frequencies, heterozygosity estimates and $F_{ST}$ were calculated using the computer program BIOSYS-1 (Swofford & Selander 1981).
In comparison to $F_{ST}$, Nei's (1978) $G_{ST}$ was calculated and dendrograms constructed for both indices. The discussion focuses on the $G_{ST}$ results as these are thought to be more reliable as indicators of the level of sub-structuring. $G_{ST}$ is independent of the number of alleles, whereas Wright's $F_{ST}$ is essentially for bi-allelic loci and requires pooling of the minor alleles for multi-allelic loci.
The occurrence of rare alleles (i.e. defined here as > 1% throughout the study area) was correlated with sub-populations and tentative relationships suggested on the basis of sharing rare alleles.
A number of techniques have been developed which can be used to analyse spatial relationships in allele frequencies (Mantel 1967, Sokal and Oden 1978a, Rossi *et al* 1992). Here I follow the methods of Sokal and Oden (1978a, b) and Oden and Sokal (1986). The null hypothesis is that there is no correlation in space or with respect to inter-sub-population distance throughout the study area. Three different approaches were used to define the spatial relationships of the sub-populations. (1) Geographically nearest neighbours, (2) neighbours in terms of expanding radii from Woburn Abbey, and (3) nearest neighbours in terms of an artificial cline derived from the ranked

microsatellite allele frequencies of two ubiquitous alleles from different loci.

## 5·3 RESULTS

### 5·3a Number of alleles and size range per locus

Of all the microsatellite primer pairs screened only the following produced polymorphic banding patterns; JP23 - JP27 - JP38 - BOVIRBP - MAF23 - MAF65 and DRBP1. The remaining primer pairs either produced a monomorphic pattern or no result at all. The number of alleles and fragment sizes are listed in Table 5·2. The complete data set may be found in Appendix D.

| Locus | Actual Number of Alleles | Actual Allele Size Range (bp) | Possible Number of Alleles within of size range | Possible Allele size Range (bp) |
|---|---|---|---|---|
| JP23 | 7 | 144 -156 | 7 | 144-156 |
| JP27 | 5 | 170/180/184/192 /200 | 16 | 170-200 |
| JP38 | 5 | 201/209/213/229/231 | 16 | 201-231 |
| BOVIRBP | 15 | 155-175 179-185 | 16 | 155-185 |
| MAF23 | 16 | 107-111 115-137 141 | 18 | 107-141 |
| MAF65 | 7 | 92/94/100/106/108 112/114 | 12 | 92-114 |
| DRBP1 | 2 | 102 108 | 4 | 102-108 |

Table 5·2 Actual number of alleles and size range along with possible number of alleles (within the confines of the two extremes of each range) and size range

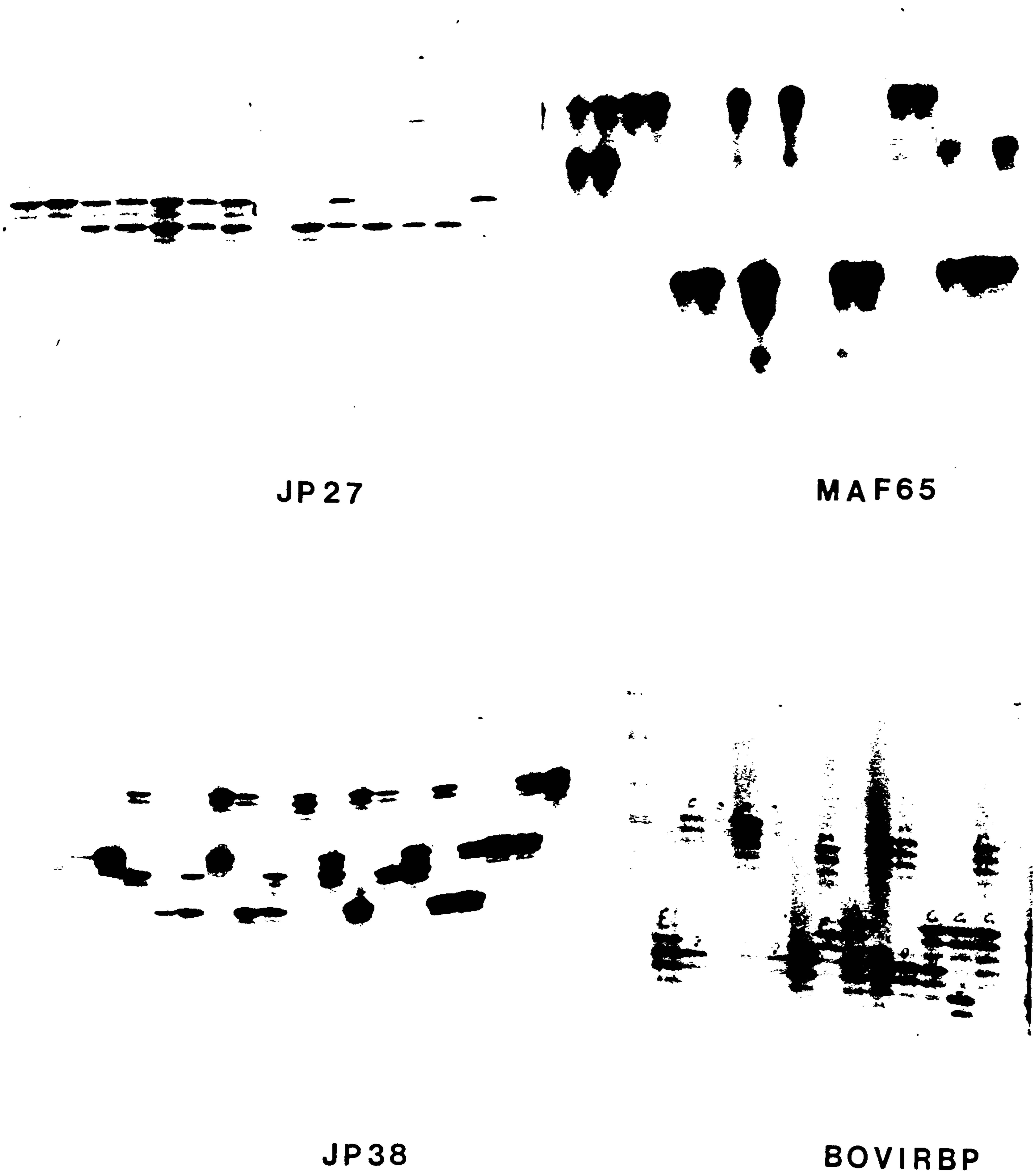


Figure 5·1
Microsatellite variation determined at the loci JP27, JP38, BOVIRBP and MAF65. Each lane contains a sample from an individual deer, with the 'AT' ladder flanking. Homozygotes show a strong single band whilst heterozygotes show two strong bands. Sub-bands can clearly be seen, but do not confound the interpretation.

## 5·3b Allele Frequencies.

Frequency distribution histograms were constructed for all loci at all populations. These are shown in Figures 5.2 - 5.8. For clarity only the allele lengths of detected alleles are included.

Locus JP23

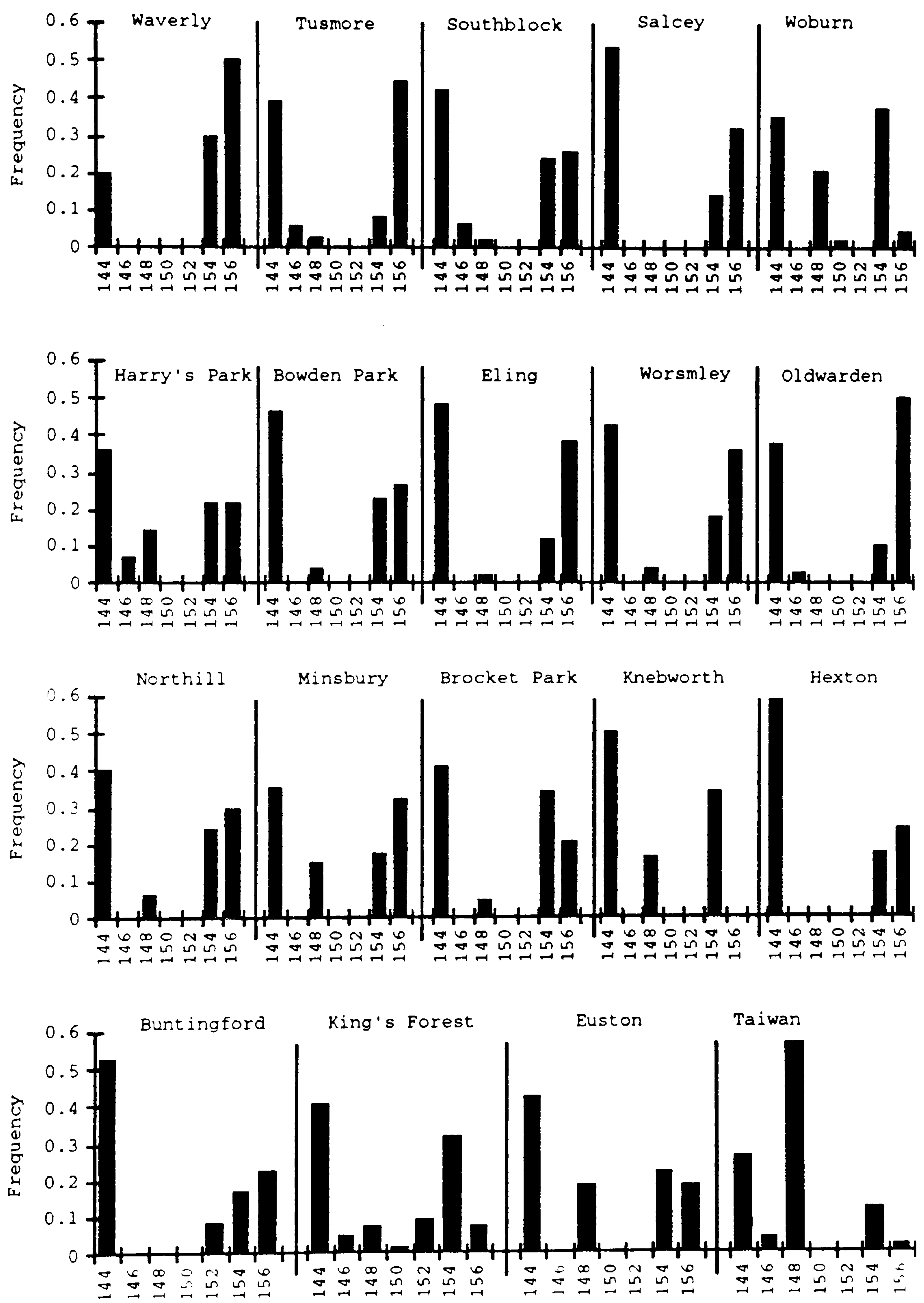


Figure 5·2a - d

Allele length (bp)

Locus JP27

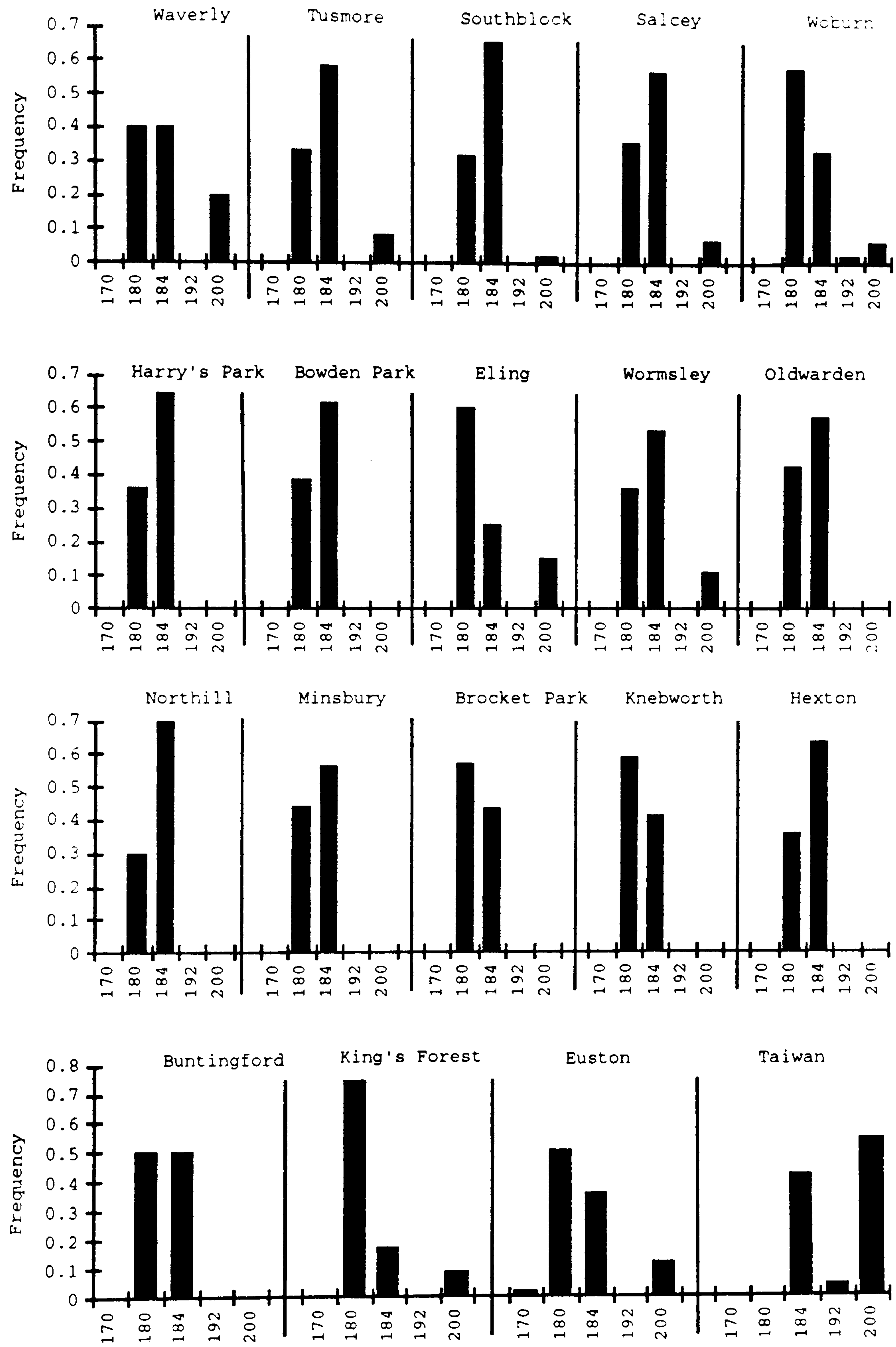


Figure 5·3a - d

Allele length (bp)

LocusJP38

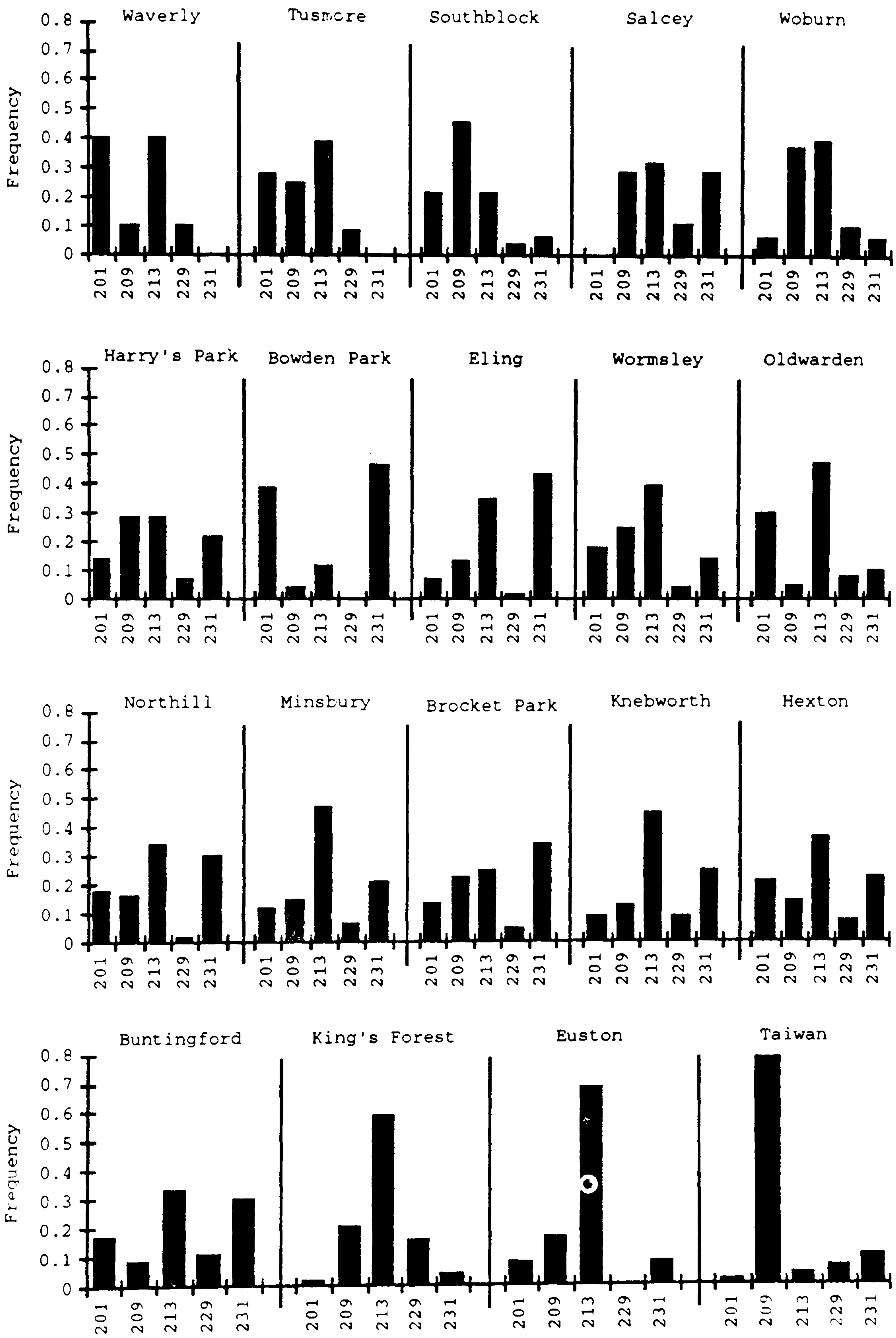


Figure 5·4a - d Allele length (bp)

Locus BOVIRBP

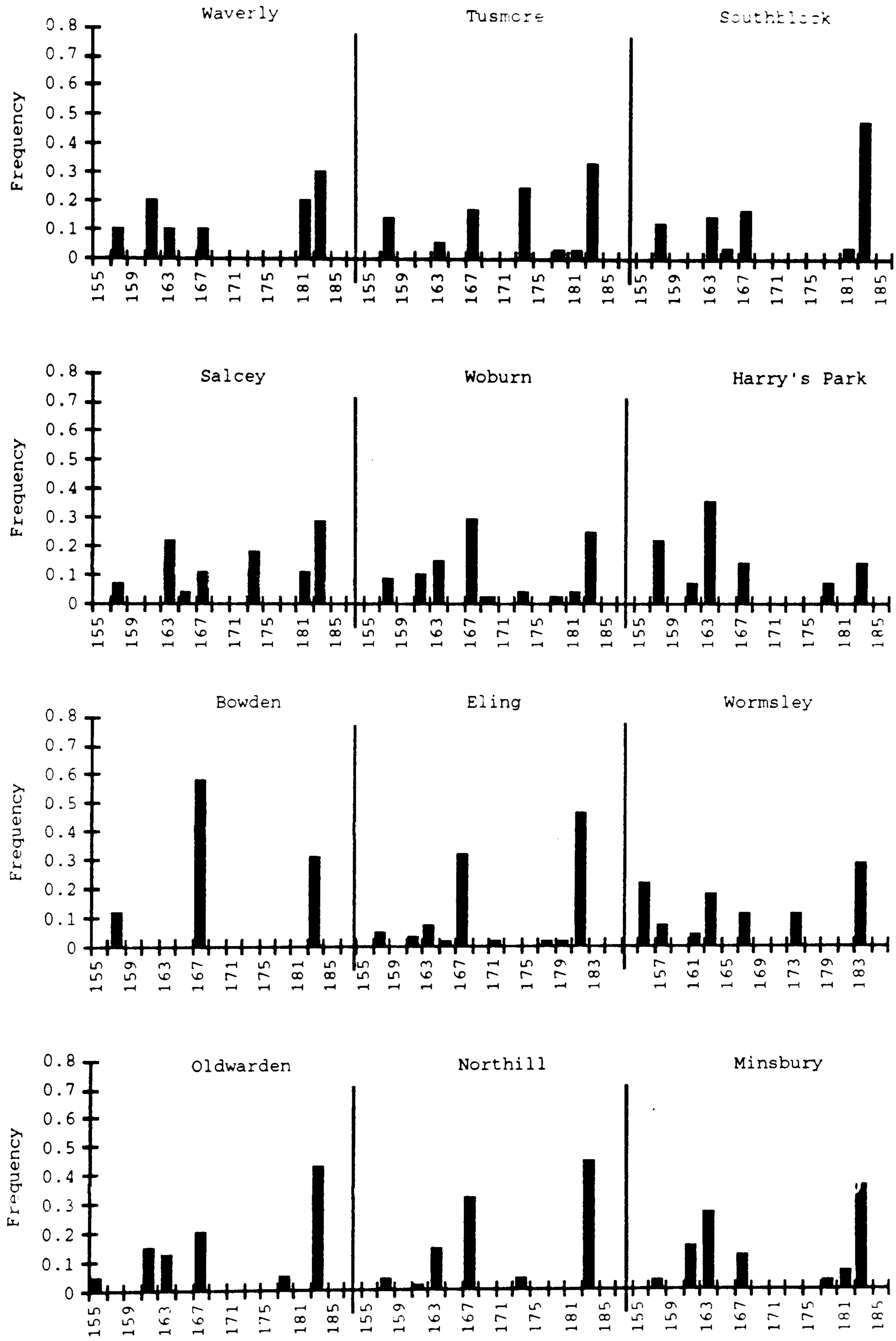


Figure 5·5a - d Allele length (bp)

Locus BOVIRBP

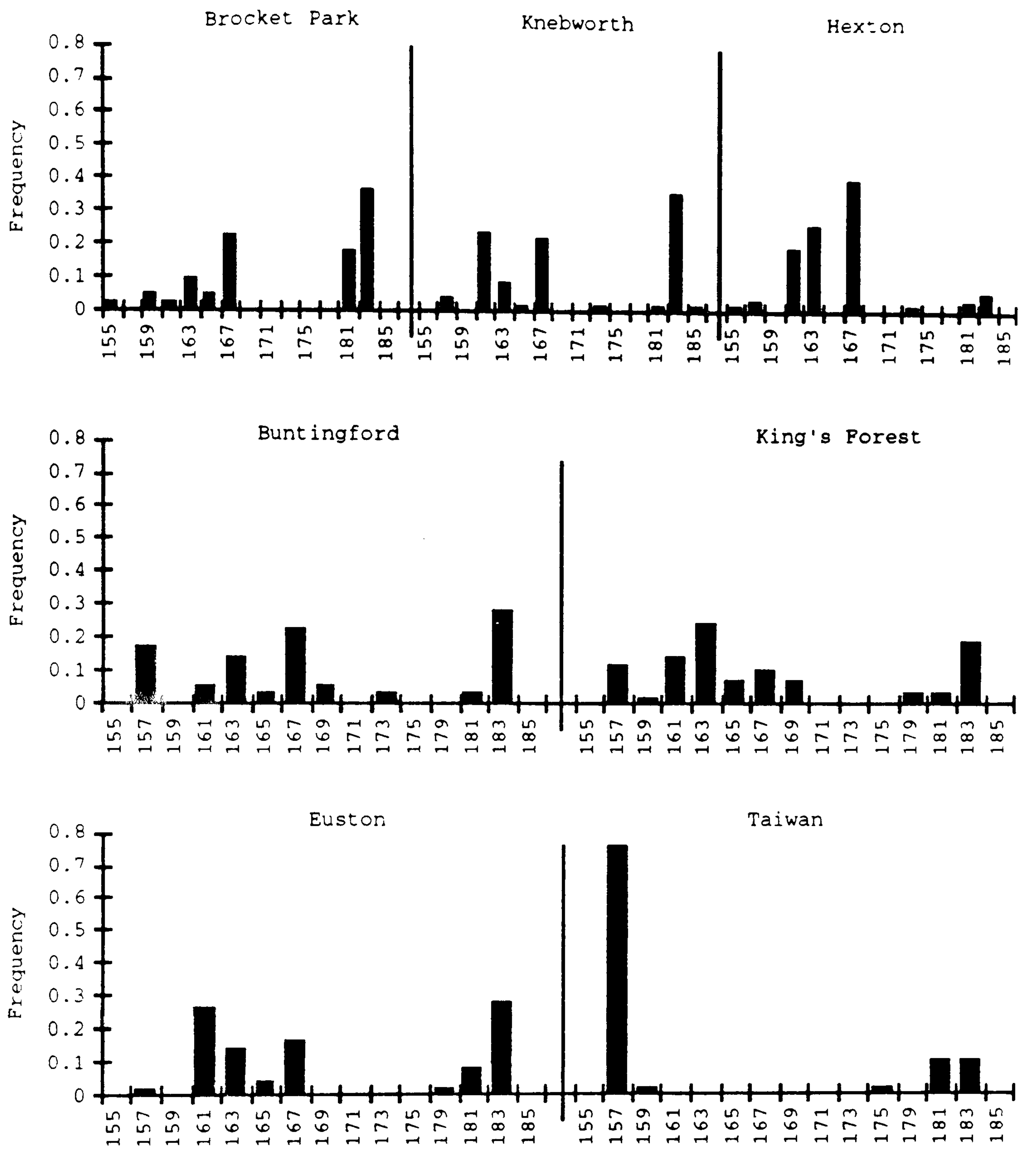


Figure 5·5e - g Allele length (bp)

Locus MAF23

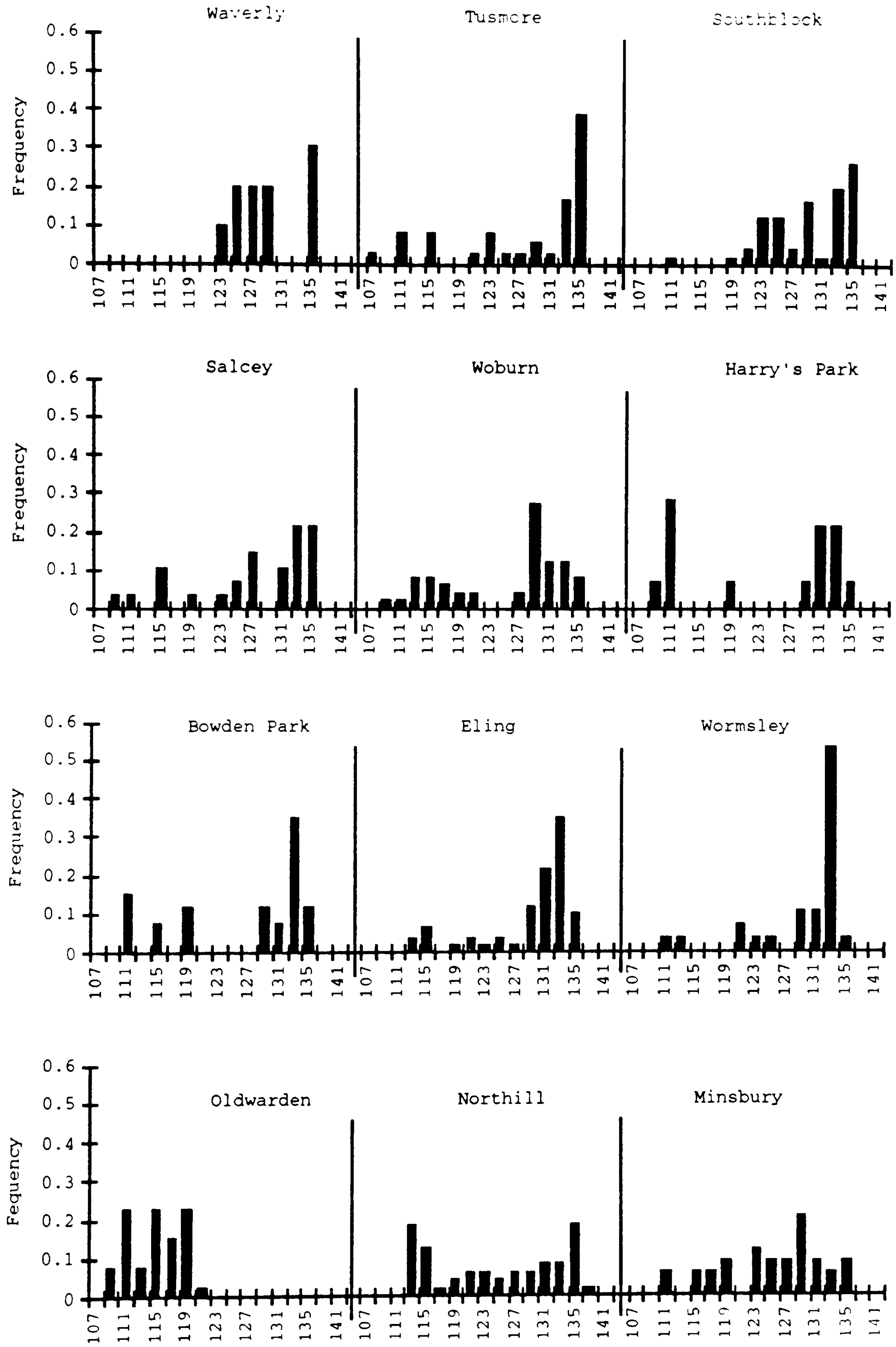


Figure 5·6a - d

Allele length (bp)

Locus MAF23

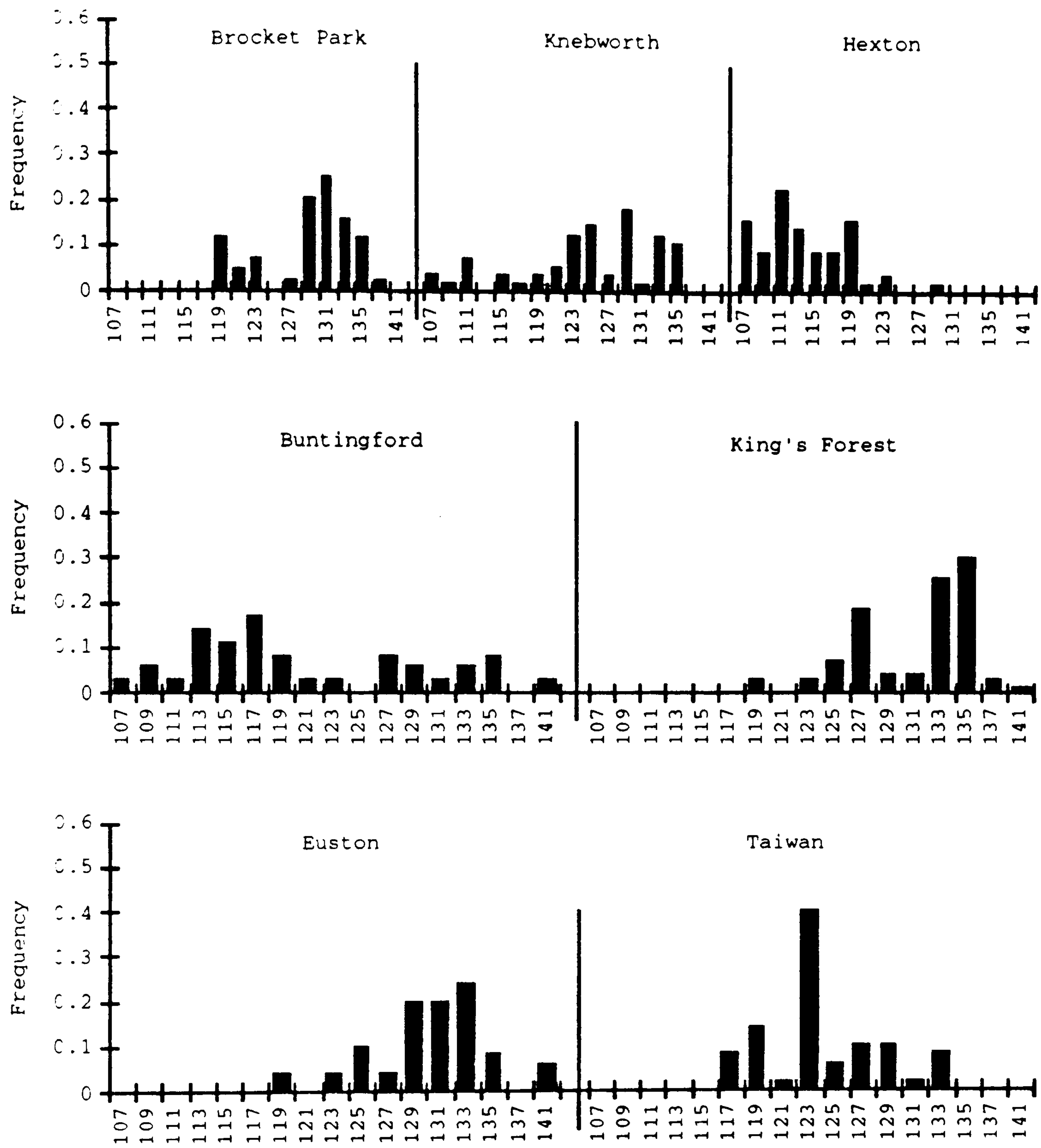


Figure 5·6e - g Allele length (bp)

Locus MAF 65

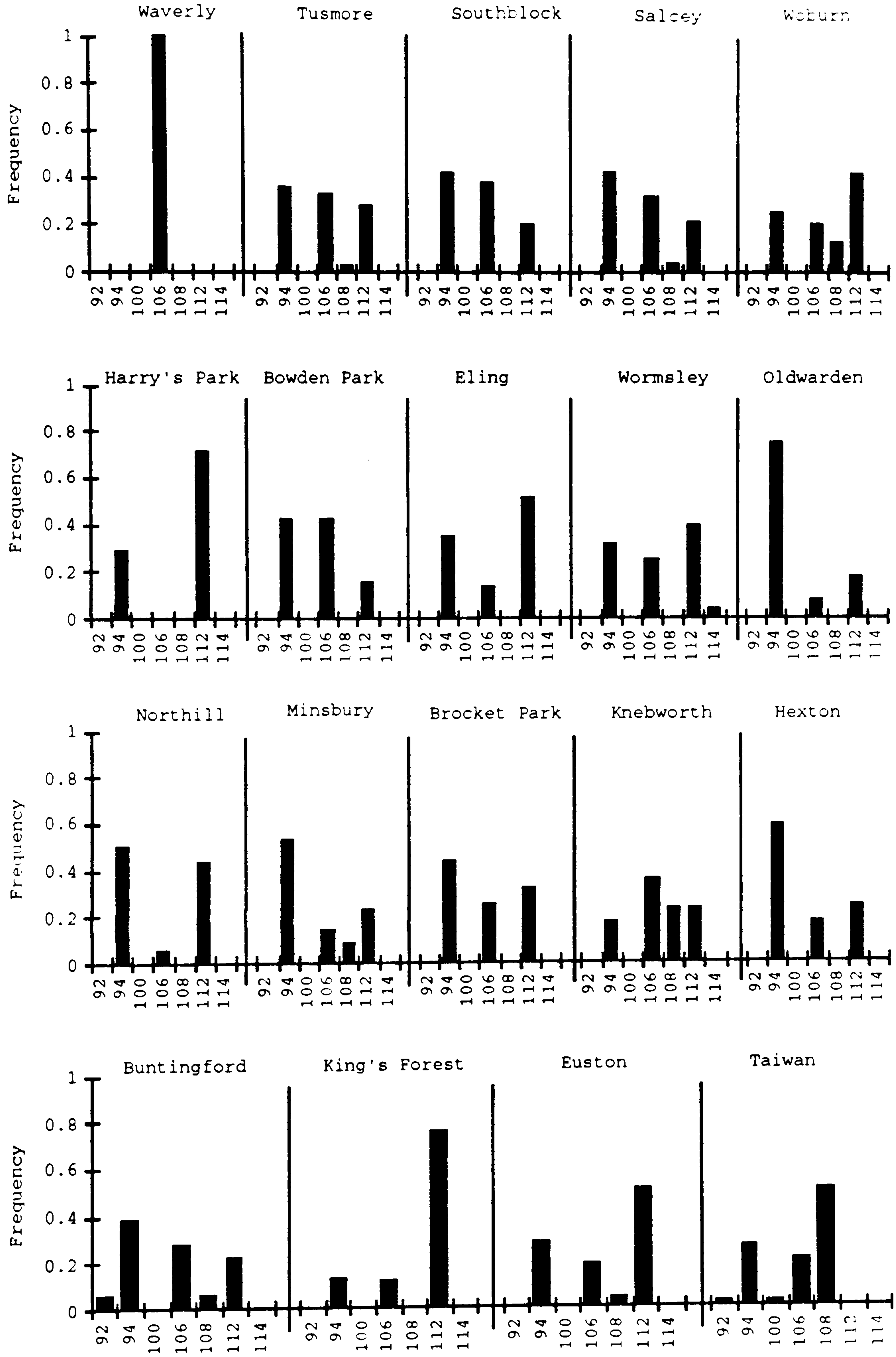


Figure 5·7a - d

Allele length (bp)

Locus DRBP1

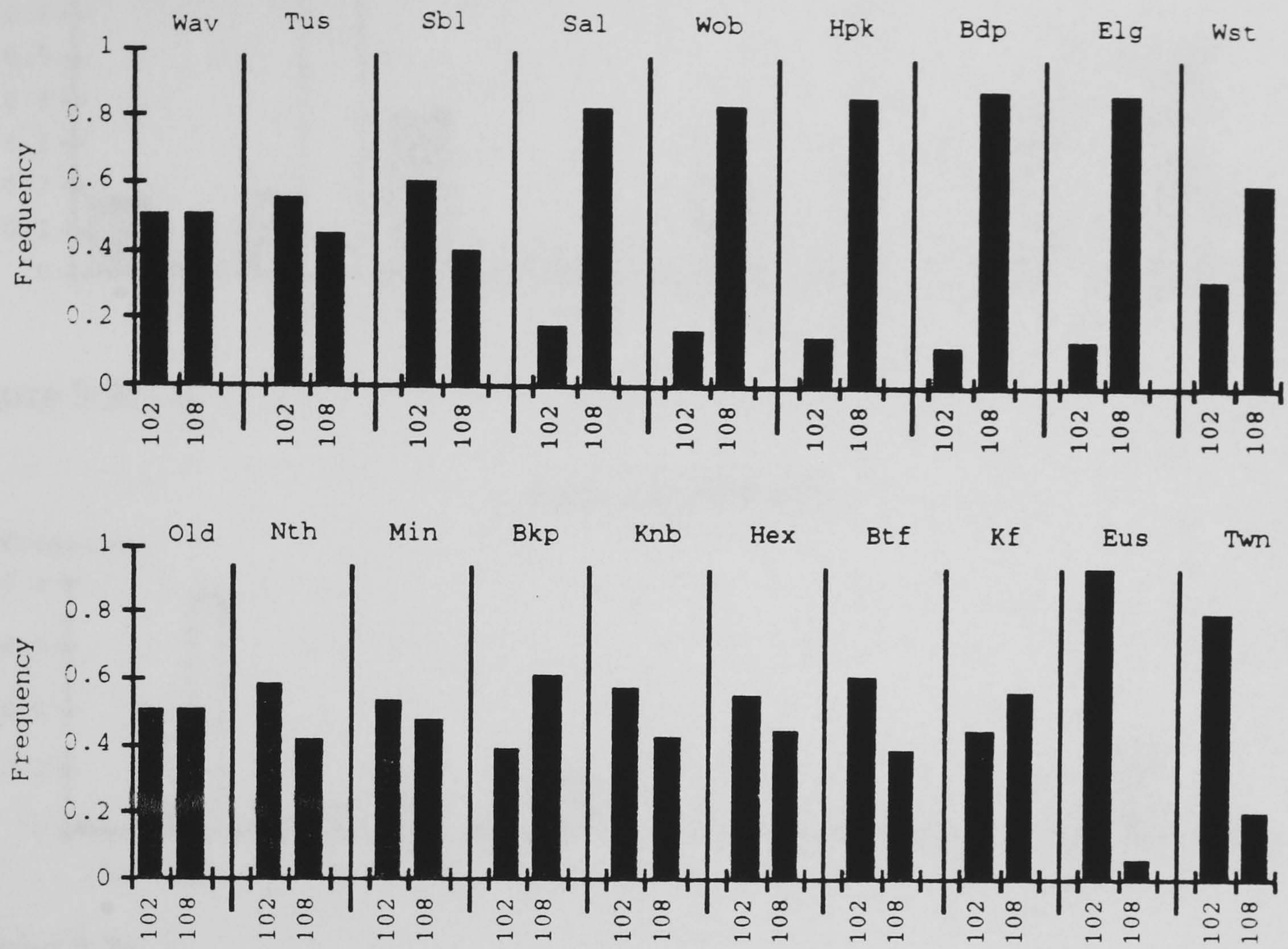


Figure 5·8a - b Allele length (bp)

Mean allele frequencies were calculated from the observed genotypes of individual animals for the 133 combinations of seven polymorphic loci and 19 sub-populations. Figures 5·9a - g describe the mean allele frequencies from the 18 study area locations (hatched). Included in each graph are the allele frequencies from the Taiwan samples (plain).

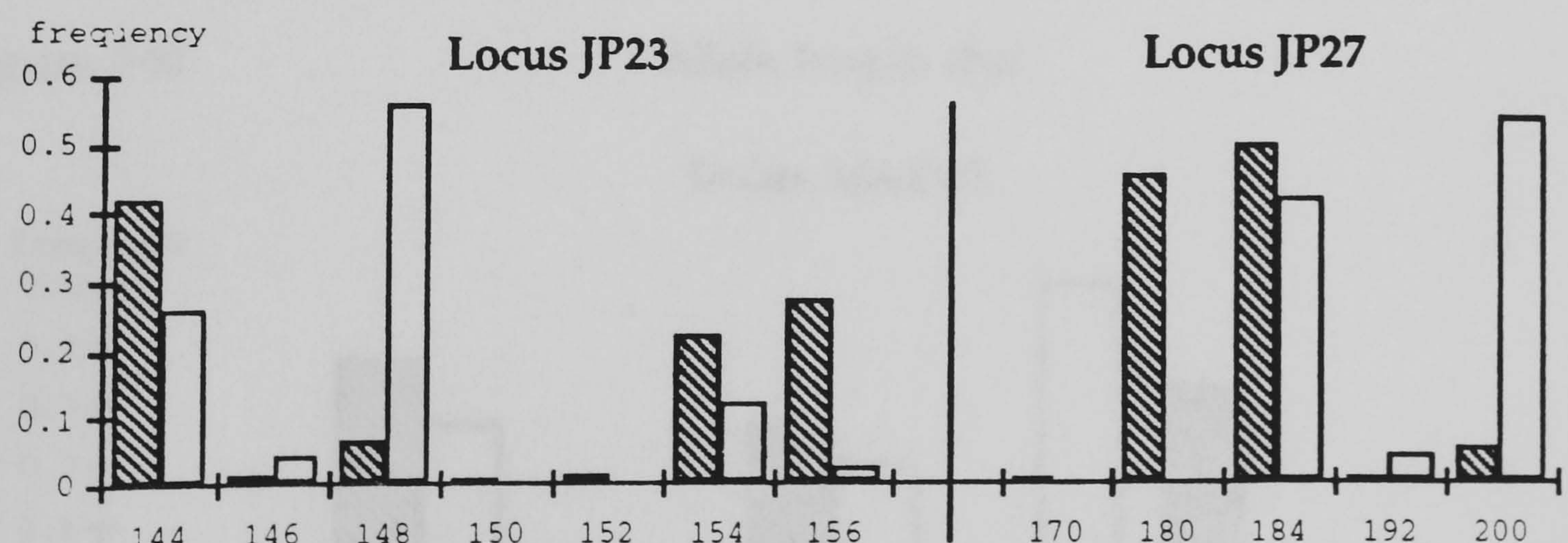


Figure 5·9a - b

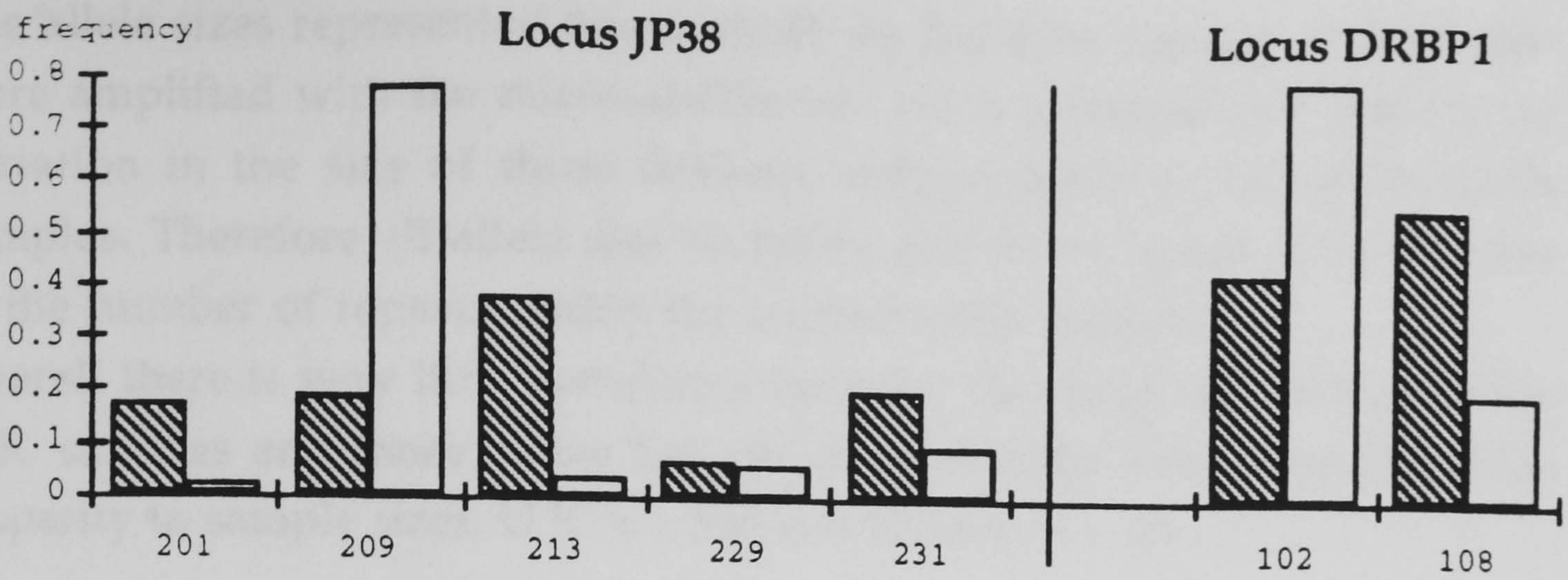


Figure 5·9c - d

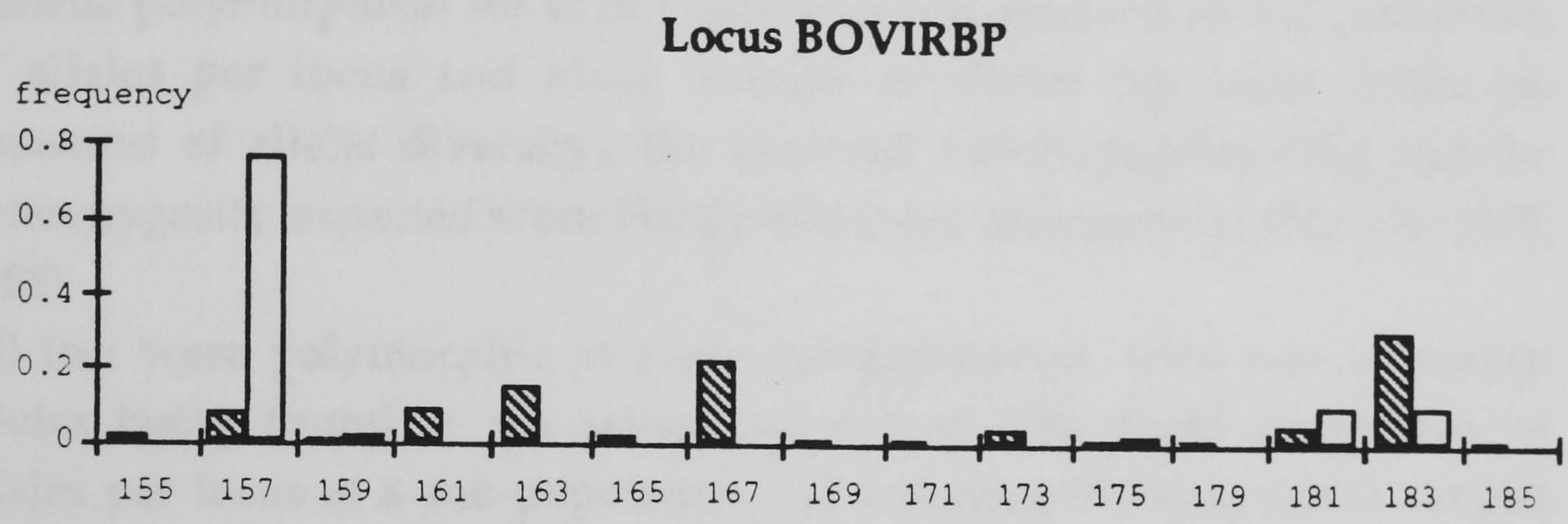


Figure 5·9e

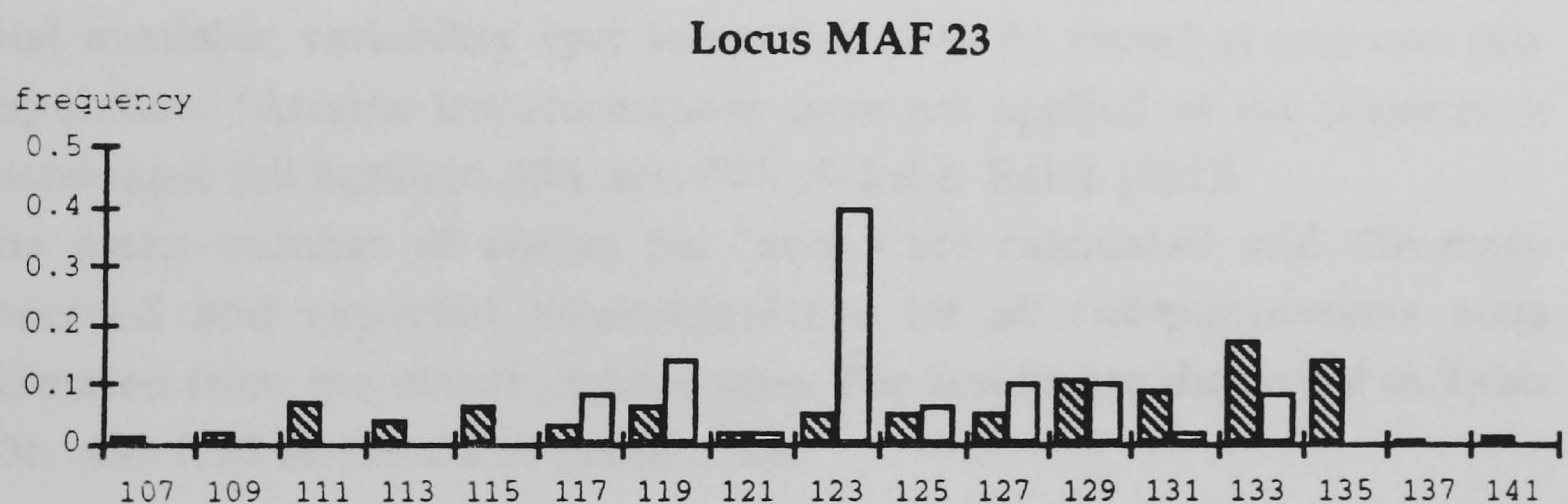


Figure 5·9f Allele length (bp)

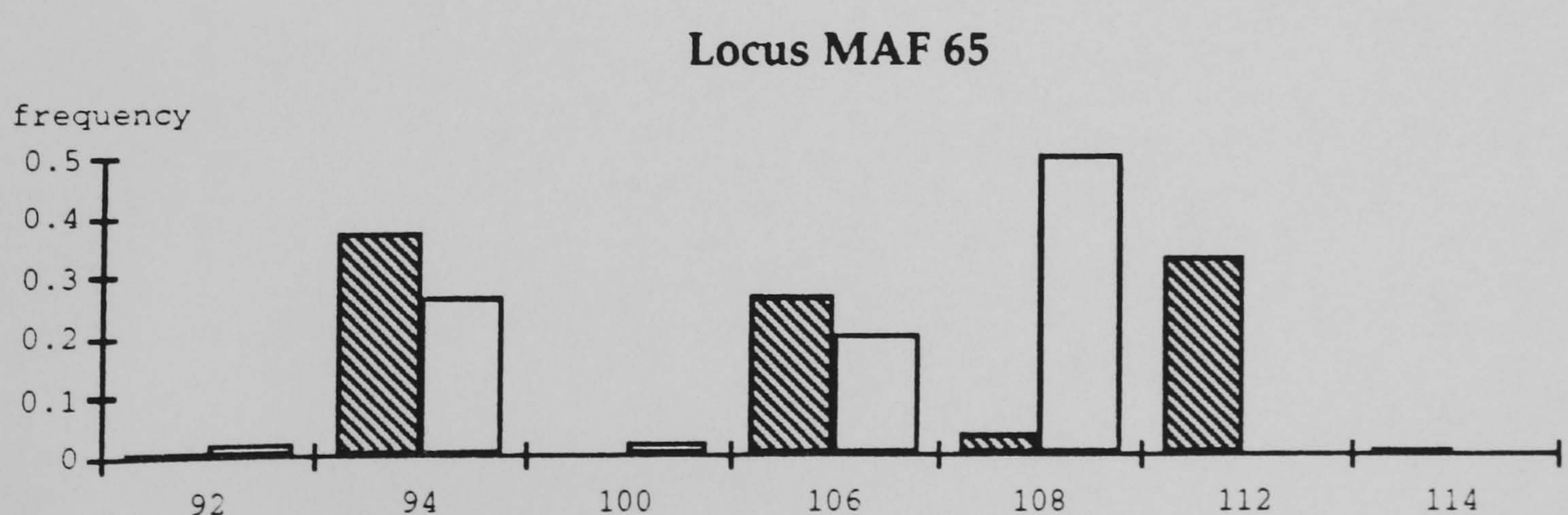


Figure 5·9g Allele length (bp)

The allele sizes represented here include the flanking regions of DNA that were amplified with the microsatellite loci - it is assumed that there is no variation in the size of those flanking regions between the individuals samples. Therefore all allele size variation is a direct result of differences in the number of repeats within the microsatellite region.

Overall there is very little correlation between the allele frequencies of the U.K. samples and those of the Taiwan samples. However, there is a great disparity in sample sizes; U.K. n = 369 and Taiwan, n = 25.

**5·3c Heterozygosity.**

Genetic polymorphism for each population was assessed as the proportion of alleles per locus and mean number of alleles per locus (both are measures of allelic diversity), the observed heterozygosity ($H_O$) and the heterozygosity expected from Hardy-Weinberg assumptions ($H_E$; Nei 1978, 1988).

All loci were polymorphic at every sub-population, with two to sixteen alleles being found at the seven loci chosen. The mean proportion of alleles per locus at a sub-population, not including Waverly (n=5), ranged from 46% (Bowden Park, n=13) to 69% (Buntingford, n=18 and Woburn, n=24), Table 5·3a. The figures in Table 5·3a represent the percentage of the total available variability (per locus) that may be found at any one sub-population. (Arcsine transformations were not applied as the majority of percentages fell between 30% and 70% (Sokal & Rohlf 1981)).

The mean number of alleles per locus were calculated and the mean observed and expected heterozygosities for all sub-populations were estimated from the observed genotypes. The results are displayed in Table 5·3b, standard errors are in parenthesis.

| Proportion (%) of Alleles Present at a Sub-population | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | JP23 | JP27 | JP38 | BOVIRBP | MAF23 | MAF65 | DRBP1 | Average |
| Wav | 42.9 | 60 | 80 | 37.5 | 29.4 | 14.3 | 100 | 44.02 |
| Tus | 71.4 | 60 | 80 | 43.8 | 64.7 | 57.1 | 100 | 62.83 |
| Sbl | 71.4 | 60 | 100 | 37.5 | 58.8 | 42.9 | 100 | 61.77 |
| Sal | 42.9 | 60 | 80 | 43.8 | 58.8 | 57.1 | 100 | 57.1 |
| Wob | 71.4 | 60 | 100 | 56.3 | 70.6 | 57.1 | 100 | 69.23 |
| Hpk | 71.4 | 40 | 100 | 37.5 | 41.2 | 28.6 | 100 | 53.12 |
| Bdp | 57.1 | 40 | 80 | 18.8 | 41.2 | 42.9 | 100 | 46.67 |
| Elg | 57.1 | 60 | 100 | 56.3 | 64.7 | 42.9 | 100 | 63.5 |
| Wst | 57.1 | 60 | 100 | 43.8 | 52.9 | 57.1 | 100 | 61.82 |
| Old | 71.4 | 40 | 100 | 31.3 | 41.2 | 42.9 | 100 | 54.47 |
| Nth | 57.1 | 40 | 100 | 50 | 76.5 | 42.9 | 100 | 61.08 |
| Min | 57.1 | 40 | 100 | 31.3 | 64.7 | 57.1 | 100 | 58.37 |
| Bkp | 57.1 | 40 | 100 | 43.8 | 52.9 | 42.9 | 100 | 56.12 |
| Knb | 57.1 | 40 | 100 | 62.5 | 82.4 | 57.1 | 100 | 66.52 |
| Hex | 42.9 | 40 | 100 | 56.3 | 58.8 | 42.9 | 100 | 56.82 |
| Btf | 57.1 | 40 | 100 | 62.5 | 88.2 | 71.4 | 100 | 69.87 |
| Kf | 100 | 60 | 100 | 50 | 58.8 | 42.9 | 100 | 68.62 |
| Eus | 57.1 | 80 | 80 | 50 | 52.9 | 57.1 | 100 | 62.85 |
| Twn | 71.4 | 40 | 100 | 31.3 | 52.9 | 71.4 | 100 | 61.17 |

Table 5·3a Percentage variability of loci, by population

| Mean No of alleles/locus | | | Heterozygosity | | | |
|---|---|---|---|---|---|---|
| | | Standard Error | Direct Count | Standard Error | HdyWbg Expected | Standard Error |
| Wav | 3.43 | 0.065 | 0.57 | 0.13 | 0.63 | 0.11 |
| Tus | 5.14 | 1.14 | 0.59 | 0.06 | 0.68 | 0.04 |
| Sbl | 4.86 | 1.01 | 0.62 | 0.06 | 0.66 | 0.05 |
| Sal | 4.71 | 1.06 | 0.51 | 0.09 | 0.66 | 0.07 |
| Wob | 5.86 | 1.3 | 0.48 | 0.07 | 0.67 | 0.07 |
| Hpk | 4.14 | 0.8 | 0.63 | 0.01 | 0.65 | 0.09 |
| Bdp | 3.57 | 0.65 | 0.58 | 0.1 | 0.59 | 0.07 |
| Elg | 5.29 | 1.29 | 0.54 | 0.09 | 0.6 | 0.07 |
| Wst | 5.0 | 0.85 | 0.59 | 0.06 | 0.69 | 0.04 |
| Old | 4.14 | 0.74 | 0.64 | 0.09 | 0.62 | 0.06 |
| Nth | 5.0 | 1.45 | 0.57 | 0.07 | 0.65 | 0.06 |
| Min | 5.0 | 1.2 | 0.65 | 0.08 | 0.69 | 0.06 |
| Bkp | 4.71 | 1.06 | 0.67 | 0.07 | 0.68 | 0.05 |
| Knb | 5.57 | 1.67 | 0.74 | 0.05 | 0.68 | 0.06 |
| Hex | 4.71 | 1.19 | 0.63 | 0.06 | 0.64 | 0.06 |
| Btf | 6.0 | 1.75 | 0.71 | 0.07 | 0.70 | 0.06 |
| Kf | 5.71 | 1.27 | 0.62 | 0.07 | 0.62 | 0.07 |
| Eus | 5.0 | 0.95 | 0.55 | 0.1 | 0.61 | 0.09 |
| Twn | 4.86 | 0.83 | 0.52 | 0.06 | 0.53 | 0.06 |
| | | | Overall Mean Heterozygosity | | | |
| | | | 0.6 | 0.015 | 0.61 | 0.03 |

Table 5·3b Summary of population heterozygosity.

Both observed heterozygosity and expected heterozygosity values were high and varied within a narrow range. The sub-population carrying the most heterozygosity was found to be Knebworth (0·74) and the sub-population carrying the least - Woburn Abbey (0·48). Woburn Abbey also had the greatest deviation from the expected value under Hardy-Wienberg expectations. Overall there was no significant difference between observed and expected heterozygosity values, as seen from Figure 5·10 the observed values are generally lower but track the expected values.

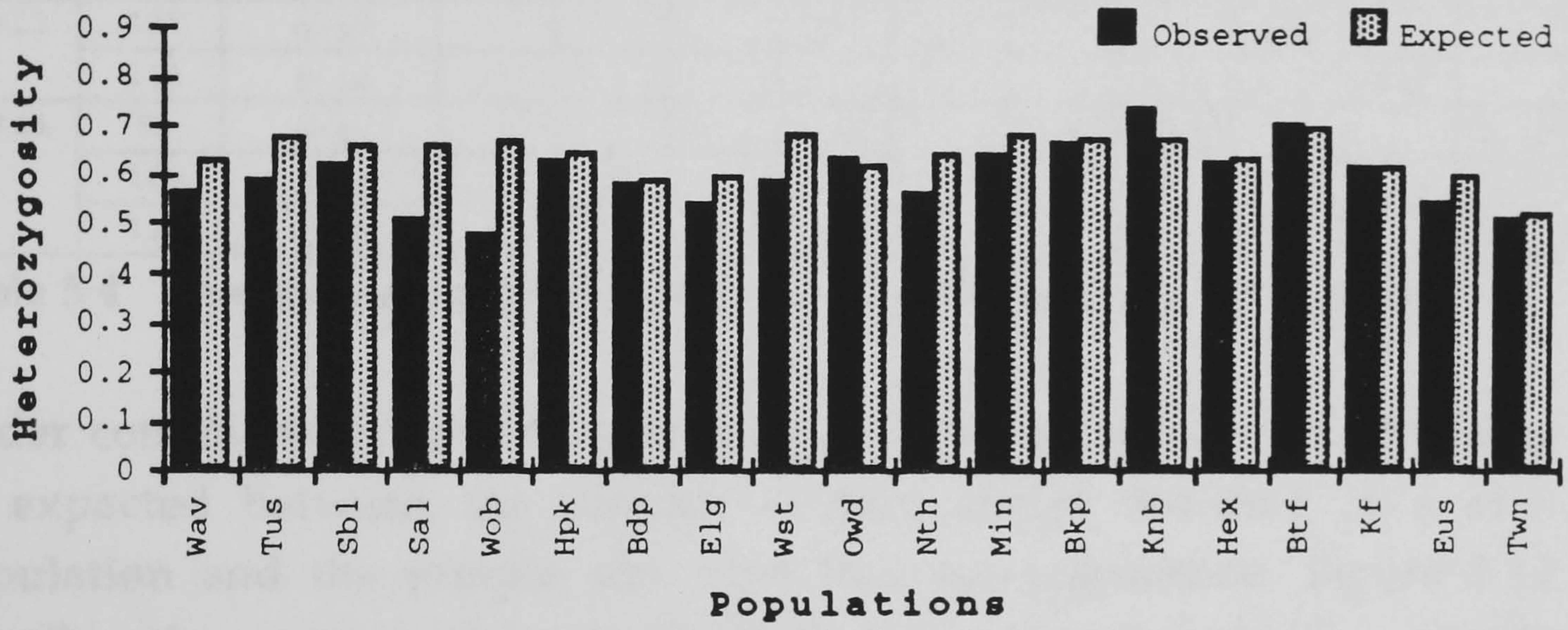


Figure 5·10 Histograms of Observed and expected heterozygosity.

### 5·3d The Occurrence of Rare Alleles

Here, rare alleles are defined as those alleles that have been found to occur at a frequency of less than 1% throughout the study area. There was no correlation between the number of alleles detected per locus and the occurrence of rare alleles in the overall population. For example, at locus BOVIRBP, 15 alleles were detected of which five were classified as rare, whereas at locus MAF23, 16 alleles were detected but only two qualified as rare. Rare alleles occurred at five of the seven loci investigated, but were not necessarily rare in the sub-population in which they occurred. Locus JP23 (allele 152) was present at 8·3% and 8·6% in the Buntingford and King's Forest sub-populations respectively. Table 5·4 summarises the occurrence and frequency of rare alleles.

| Locus | Allele | Overall % Occurrence | Percentage Occurrence in Sub-populations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Wob | Elg | Wst | Nth | Bkp | Knb | Btf | Kf | Eus | Twn |
| JP23 | 150 | 0·19 | 2·1 | | | | | | | 1·4 | | |
| | 152 | 0·94 | | | | | | | 8·3 | 8·6 | | |
| JP27 | 170 | 0·11 | | | | | | | | | 2.0 | |
| | 192 | 0·32 | 2·1 | | | | | | | | | 4.0 |
| BOVIRBP | 159 | 0·42 | | | | | 4·5 | | | 1·4 | | 2.0 |
| | 169 | 0·78 | 2·1 | | | | | | 5·6 | 7·1 | | |
| | 171 | 0·09 | | 1·7 | | | | | | | | |
| | 175 | 0·11 | | | | | | | | | | 2.0 |
| | 185 | 0·09 | | | | | | 1·8 | | | | |
| MAF23 | 137 | 0·37 | | | | 2·0 | 2·3 | | | 2·9 | | |
| | 141 | 0·54 | | | | | | | 2·8 | 1·4 | 6.0 | |
| MAF65 | 92 | 0·4 | | | | | | | 5·6 | | | 2 .0 |
| | 100 | 0·11 | | | | | | | | | | 2.0 |
| | 114 | 0·19 | | | 3·6 | | | | | | | |

Table 5·4 Frequency of rare alleles and the sub-populations in which they occur.

Under conditions of gene flow among sub-populations a correlation may be expected between the number of rare alleles detected in a sub-population and the sample size from that sub-population. Figure 5·11 describes the mean number of rare alleles detected with increasing sample size from all 19 sample sites.

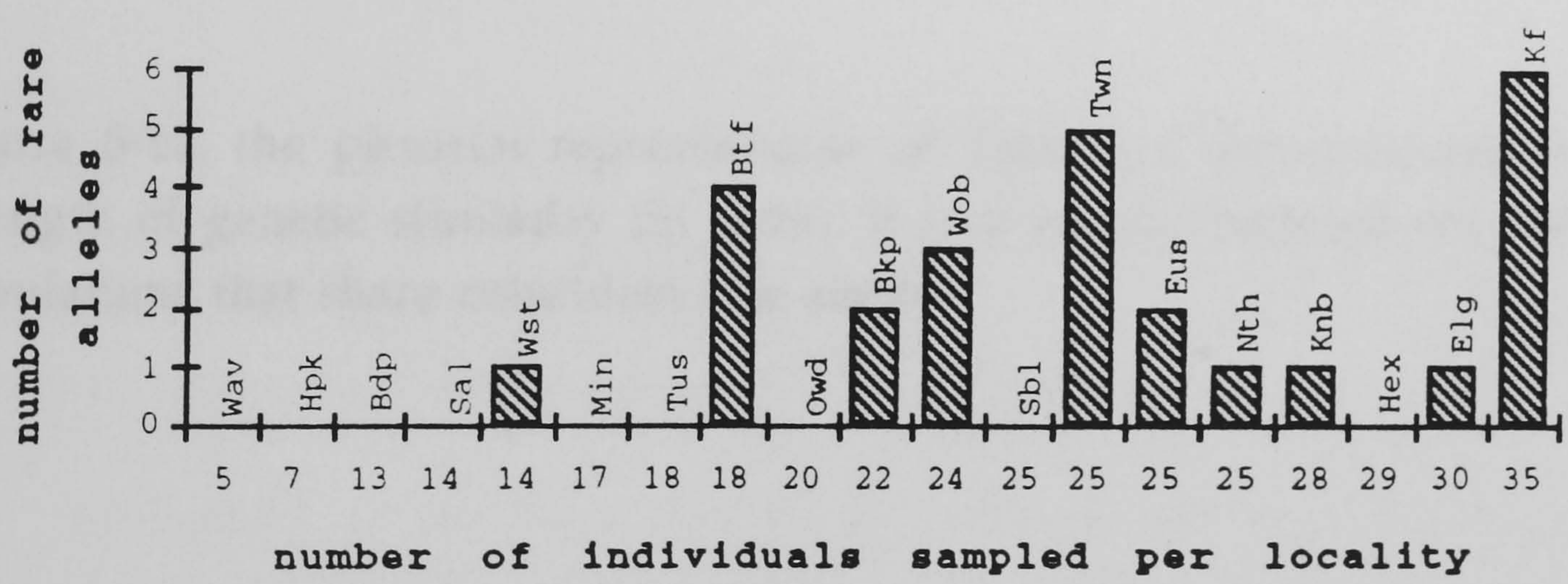


Figure 5·11 Relationship between sample size and the number of rare alleles.

To assess the association between the number of alleles detected and sample size, Kendall's coefficient of rank correlation was calculated (Campbell, 1992) with correction for ties. This produced a coefficient of 0·33, not significant at the 5% level and thus the null hypothesis that there is no association between the mean number of alleles detected and sample size is accepted.

From an inspection of Table 5·4 some groupings could be proposed based on the number of shared rare alleles. Figure 5·12 depicts these groupings.

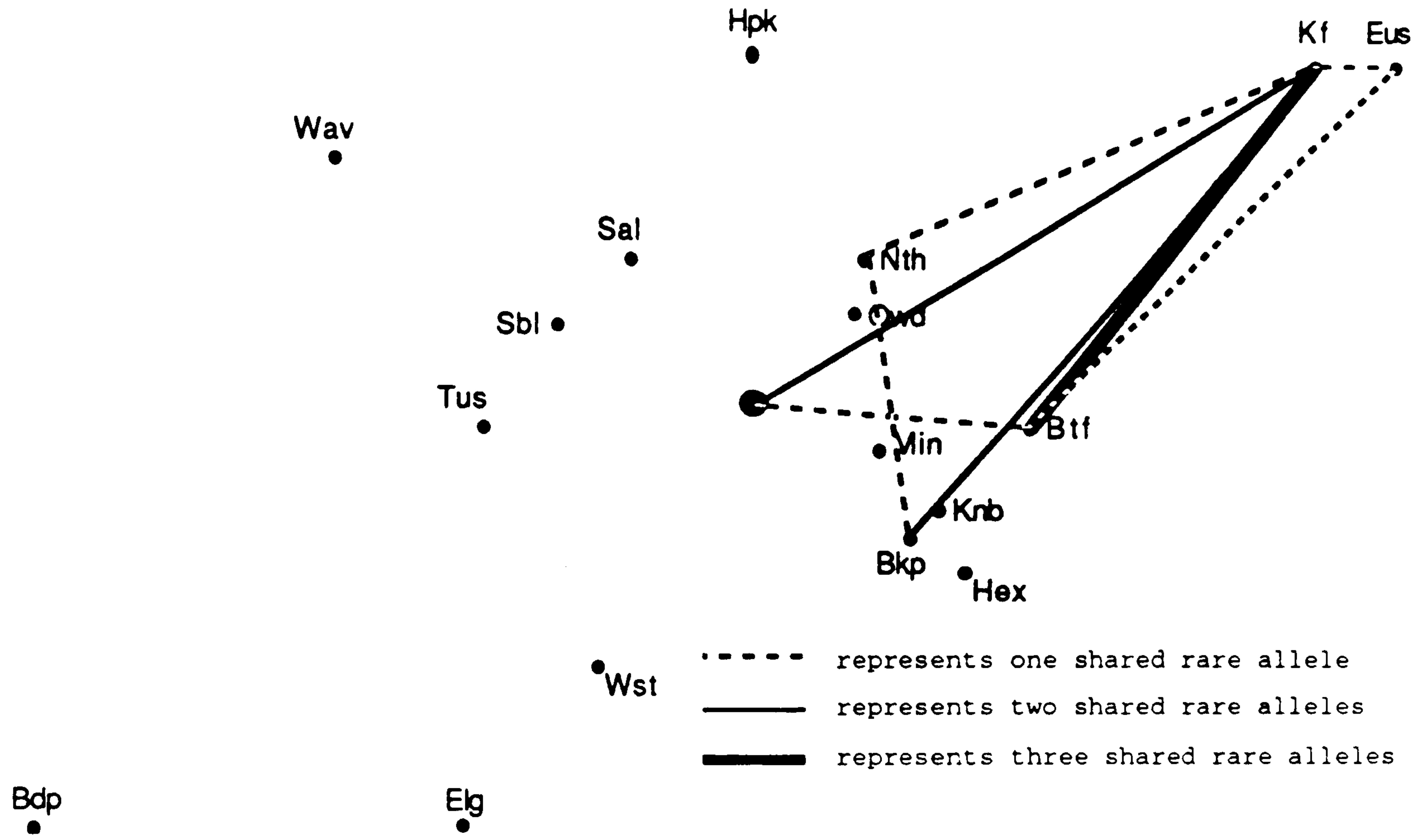


Figure 5·12 Relatedness of sub-populations according to occurrence of rare alleles.

Figure 5·12, the pictorial representation of Table 5·4 demonstrates the strength of genetic similarity (in terms of rare alleles) between the sub-populations that share coincident rare alleles.

## 5·3e Investigation of Hardy-Weinberg expectations.

Because of the low number of samples collected from several of the locations and the large number of alleles per locus, an investigation to determine if the sub-populations were in Hardy-Weinberg equilibrium could not be undertaken. The data from the sub-populations was therefore pooled and the allele frequencies analysed to determine if the total population was in Hardy-Weinberg equilibrium. There are eight assumptions made under this model of population genetic structure. The organism is diploid, reproduction is sexual, mating is random, mutation can be ignored, natural selection does not affect the gene (locus) under consideration, migration is negligible, generations are non-overlapping, population size is very large. The first six assumptions can be held to be true, however, the generations are overlapping in that there is no season for breeding, (Chapman *et al* 1984) so offspring are produced all year round and parents and offspring co-exist in the same sub-populations. Additionally the population size may not be very large, especially in locations where they have only recently become established. Observed and expected allele frequencies were calculated and the goodness of fit between the two frequencies calculated by means of the chi-squared value ($\chi^2$).

From Table 5·5 it will be seen that at some of the loci genotypes are not evident in the proportions predicted by Hardy-Weinberg equilibrium. The fixation index F (Wright 1965) was calculated as an indicator of deviation from Hardy-Weinburg equilibrium. Positive values of F indicate an increase in homozygosity whilst negative values indicate an increase in heterozygosity. Values close to zero indicate random mating.

Chi-square test for deviation from Hardy-Weinberg expectations

| | Observed frequency | Expected frequency | $\chi^2$ (1df) | P |
|---|---|---|---|---|
| **locus JP23** | | | | |
| Homozygotes for most common allele | 29 | 70.247 | 76.221 | 0 |
| Common/rare heterozygotes | 264 | 181.507 | | |
| Rare homozygotes/other heterozygotes | 76 | 117.247 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 296 Expected heterozygotes 256 Fixation index(F) -0.154 | | | | |
| **locus JP27** | | | | |
| Homozygotes for most common allele | 89 | 83.946 | 1.112 | 0.3 |
| Common/rare heterozygotes | 174 | 184.108 | | |
| Rare homozygotes/other heterozygotes | 106 | 100.946 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 189 Expected heterozygotes 201 Fixation index(F) 0.063 | | | | |
| **locus JP38** | | | | |
| Homozygotes for most common allele | 65 | 57.767 | 2.480 | 0.11 |
| Common/rare heterozygotes | 162 | 176.466 | | |
| Rare homozygotes/other heterozygotes | 142 | 134.767 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 245 Expected heterozygotes 273 Fixation index(F) 0.104 | | | | |

| | | | | |
|---|---|---|---|---|
| locus BOVIRBP | | | | |
| Homozygotes for most common allele | 55 | 36.251 | | |
| Common/rare heterozygotes | 121 | 158.499 | 20.598 | 0 |
| Rare homozygotes/other heterozygotes | 192 | 173.251 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 250 Expected heterozygotes 297 Fixation index(F) 0.16 | | | | |
| locus MAF23 | | | | |
| Homozygotes for most common allele | 29 | 11.1 | | |
| Common/rare heterozygotes | 70 | 105.799 | 42.249 | 0 |
| Rare homozygotes/other heterozygotes | 270 | 252.1 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 253 Expected heterozygotes 334 Fixation index(F) 0.244 | | | | |
| locus MAF65 | | | | |
| Homozygotes for most common allele | 90 | 51.61 | | |
| Common/rare heterozygotes | 96 | 172.780 | 72.868 | 0 |
| Rare homozygotes/other heterozygotes | 183 | 144.61 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 137 Expected heterozygotes 250 Fixation index(F) 0.453 | | | | |
| locus DRBP1 | | | | |
| Homozygotes for most common allele | 101 | 107.32 | | |
| Common/rare heterozygotes | 196 | 183.36 | 1.753 | 0.19 |
| Rare homozygotes/other heterozygotes | 72 | 78.32 | | |
| Coefficients for heterozygous deficiency or excess | | | | |
| Observed heterozygotes 196 Expected heterozygotes 184 Fixation index(F) -.064 | | | | |

Table 5·5 Chi-square test for deviation from Hardy-Weinberg expectations for the sub-populations treated as one panmictic population. Coefficients for heterozygous deficiency or excess and fixation index (F) are also shown

Only three of the seven loci were found to conform to Hardy-Weinberg equilibrium expectations. These were JP27, JP38 and DRBP1. These were also the loci with the lowest number of detected alleles; JP27 - 5 alleles, JP38 - 5 alleles and DRBP1 - 2 alleles. Of the four loci that deviated from Hardy-Weinberg equilibrium expectations JP23 (7 alleles) deviated due to heterozygous excess and MAF65 (7 alleles), BOVIRBP (15 alleles) and MAF23 (16 alleles) all deviated due to heterozygous deficiency. These observations are discussed in section 5·4.

To determine if there was any correlation between the number of alleles detected per locus and the deviation from Hardy-Weinberg equilibrium a scatter diagram was constructed of the number of alleles detected per locus against the probability (expressed as a percentage) of the loci being out of Hardy-Weinberg equilibrium.

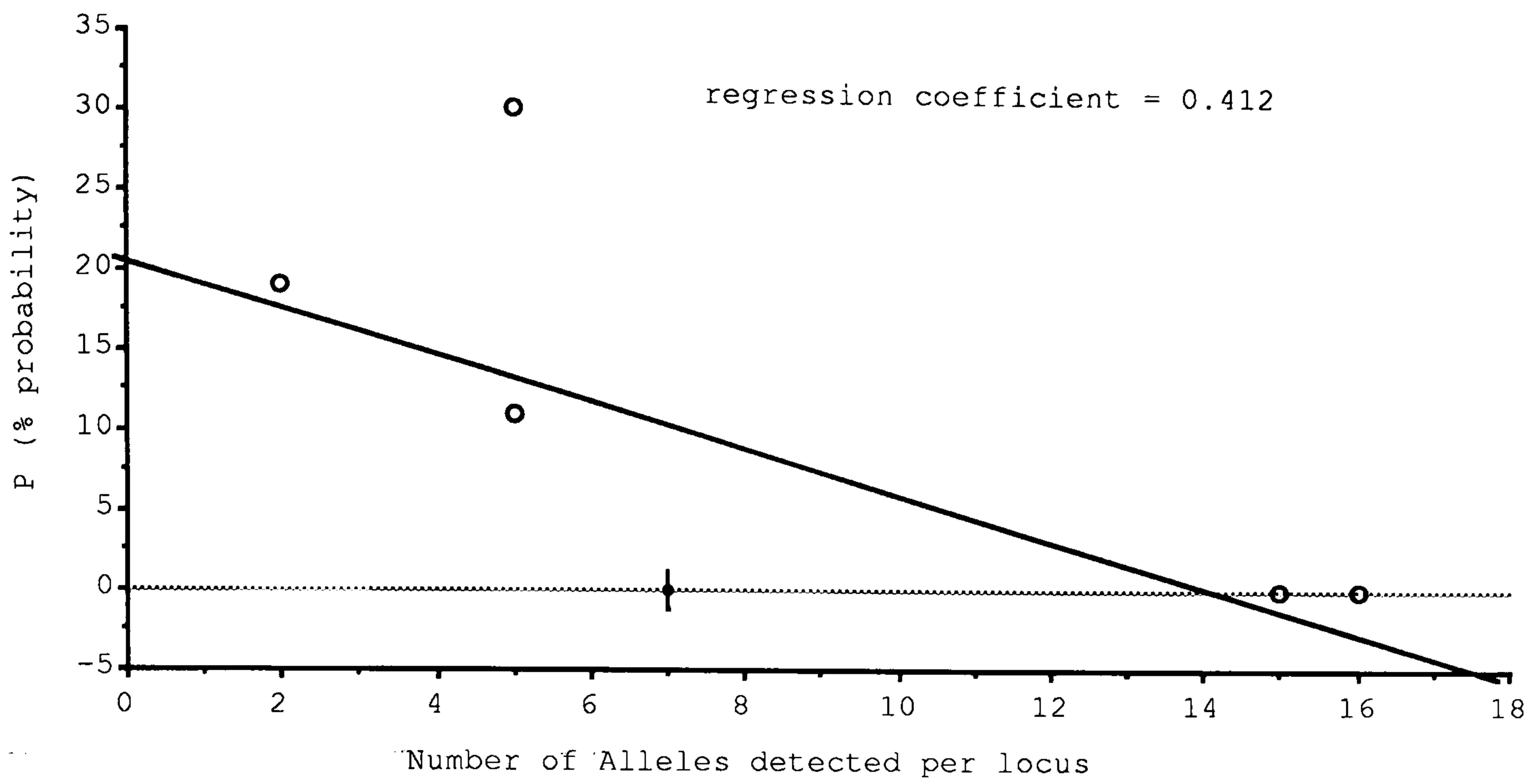


Figure 5·13 Correlation between the number of alleles detected per locus against the percentage probability of the loci being out of Hardy-Weinberg equilibrium.

From Figure 5·13 it can be seen that there is a significant correlation between the number of alleles detected per locus and the incidence of deviation from Hardy-Weinberg equilibrium (regression coefficient = 0·412 - significant at the 95% level). Loci with a greater number of alleles are more likely to deviate from Hardy-Weinberg equilibrium.

### 5·3f Comparison between Wright's $F_{ST}$ and Nei's $G_{ST}$ as measures of population subdivision.

To investigate the data produced by the microsatellite analysis, both Wright's $F_{ST}$ and Nei's $G_{ST}$ were calculated. Two data sets were chosen; (1) a comparison between three groups of sub-populations partitioned according to their radial distance from Woburn Abbey (see Figure 2·1). These three groups of populations were defined because if there has been a migratory expansion from the putative centre of origin, then there is likely to be a positive correlation between genetic differentiation and increasing geographical distance. (2) Pairwise comparison between all sub-populations.

### 5·3f1 Analysis of pairwise comparisons between all populations.

Although $F_{ST}$ and $G_{ST}$ are closely related measures of sub-population similarity, the differences between them due to interpreting the data from a bi-allelic (i.e. with pooled minor alleles) or multi-allelic perspective are quite substantial. The pairwise data are recorded in Table 5·8 (page 198), $F_{ST}$ below the diagonal and $G_{ST}$ above.

$G_{ST}$ above diagonal, $F_{ST}$ below diagonal

| | Wav | Tus | Sbl | Sal | Wob | Hpk | Bdp | Elg | Wst | Old | Nth | Min | Bkp | Knb | Hex | Btf | Kf | Eus | Twn |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav | | 0.08 | 0.09 | 0.11 | 0.12 | 0.17 | 0.14 | 0.16 | 0.13 | 0.13 | 0.12 | 0.1 | 0.1 | 0.09 | 0.12 | 0.09 | 0.15 | 0.14 | 0.2 |
| Tus | 0.06 | | 0.04 | 0.06 | 0.08 | 0.09 | 0.1 | 0.1 | 0.07 | 0.05 | 0.04 | 0.04 | 0.06 | 0.06 | 0.05 | 0.04 | 0.09 | 0.09 | 0.15 |
| Sbl | 0.07 | 0.02 | | 0.06 | 0.08 | 0.09 | 0.1 | 0.11 | 0.07 | 0.06 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.04 | 0.1 | 0.08 | 0.13 |
| Sal | 0.09 | 0.03 | 0.04 | | 0.04 | 0.05 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 | 0.04 | 0.04 | 0.07 | 0.06 | 0.05 | 0.09 | 0.13 | 0.18 |
| Wob | 0.1 | 0.06 | 0.05 | 0.05 | | 0.05 | 0.08 | 0.05 | 0.07 | 0.09 | 0.07 | 0.06 | 0.04 | 0.05 | 0.08 | 0.06 | 0.06 | 0.12 | 0.18 |
| Hpk | 0.14 | 0.07 | 0.07 | 0.04 | 0.04 | | 0.11 | 0.09 | 0.1 | 0.12 | 0.09 | 0.09 | 0.08 | 0.11 | 0.1 | 0.1 | 0.1 | 0.16 | 0.23 |
| Bdp | 0.11 | 0.07 | 0.07 | 0.04 | 0.06 | 0.07 | | 0.08 | 0.1 | 0.1 | 0.09 | 0.09 | 0.07 | 0.1 | 0.09 | 0.08 | 0.15 | 0.19 | 0.25 |
| Elg | 0.13 | 0.06 | 0.07 | 0.03 | 0.04 | 0.05 | 0.05 | | 0.07 | 0.08 | 0.07 | 0.06 | 0.04 | 0.07 | 0.07 | 0.07 | 0.07 | 0.13 | 0.23 |
| Wst | 0.09 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | 0.05 | 0.03 | | 0.06 | 0.05 | 0.04 | 0.04 | 0.06 | 0.05 | 0.05 | 0.07 | 0.1 | 0.17 |
| Old | 0.11 | 0.04 | 0.05 | 0.06 | 0.08 | 0.08 | 0.08 | 0.07 | 0.05 | | 0.04 | 0.03 | 0.06 | 0.08 | 0.04 | 0.05 | 0.12 | 0.11 | 0.2 |
| Nth | 0.11 | 0.03 | 0.03 | 0.04 | 0.05 | 0.05 | 0.06 | 0.05 | 0.04 | 0.03 | | 0.02 | 0.03 | 0.05 | 0.03 | 0.02 | 0.02 | 0.07 | 0.16 |
| Min | 0.08 | 0.03 | 0.03 | 0.03 | 0.04 | 0.05 | 0.07 | 0.05 | 0.03 | 0.02 | 0.02 | | 0.03 | 0.03 | 0.03 | 0.02 | 0.07 | 0.06 | 0.15 |
| Bkp | 0.08 | 0.04 | 0.03 | 0.03 | 0.03 | 0.05 | 0.04 | 0.03 | 0.03 | 0.05 | 0.03 | 0.02 | | 0.03 | 0.04 | 0.02 | 0.05 | 0.08 | 0.16 |
| Knb | 0.07 | 0.04 | 0.04 | 0.05 | 0.03 | 0.07 | 0.07 | 0.06 | 0.04 | 0.06 | 0.04 | 0.02 | 0.03 | | 0.05 | 0.05 | 0.06 | 0.06 | 0.15 |
| Hex | 0.11 | 0.04 | 0.05 | 0.05 | 0.06 | 0.06 | 0.06 | 0.07 | 0.04 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | | 0.03 | 0.11 | 0.09 | 0.18 |
| Btf | 0.08 | 0.03 | 0.03 | 0.04 | 0.05 | 0.06 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | | 0.08 | 0.07 | 0.15 |
| Kf | 0.12 | 0.07 | 0.08 | 0.07 | 0.04 | 0.06 | 0.12 | 0.06 | 0.05 | 0.1 | 0.07 | 0.06 | 0.05 | 0.05 | 0.08 | 0.06 | | 0.08 | 0.2 |
| Eus | 0.11 | 0.06 | 0.06 | 0.1 | 0.08 | 0.1 | 0.14 | 0.1 | 0.06 | 0.08 | 0.05 | 0.04 | 0.06 | 0.04 | 0.07 | 0.05 | 0.05 | | 0.15 |
| Twn | 0.19 | 0.14 | 0.12 | 0.16 | 0.16 | 0.18 | 0.21 | 0.21 | 0.15 | 0.19 | 0.16 | 0.14 | 0.16 | 0.14 | 0.16 | 0.14 | 0.19 | 0.15 | |

Table 5·6 Pairwise values for $G_{ST}$ and $F_{ST}$

If the population means are displayed graphically (Figure 5·14) it can be seen that the values differ, but in most populations a proportionality is preserved. The population means taken from the table of pairwise $F_{ST}$ and $G_{ST}$ are calculated as means of (popA+popB) + (popA+popC) + (popA+popD) + ..., thus the mean represents the average homozygosity of population A to all other populations. The data are arranged in ascending order.

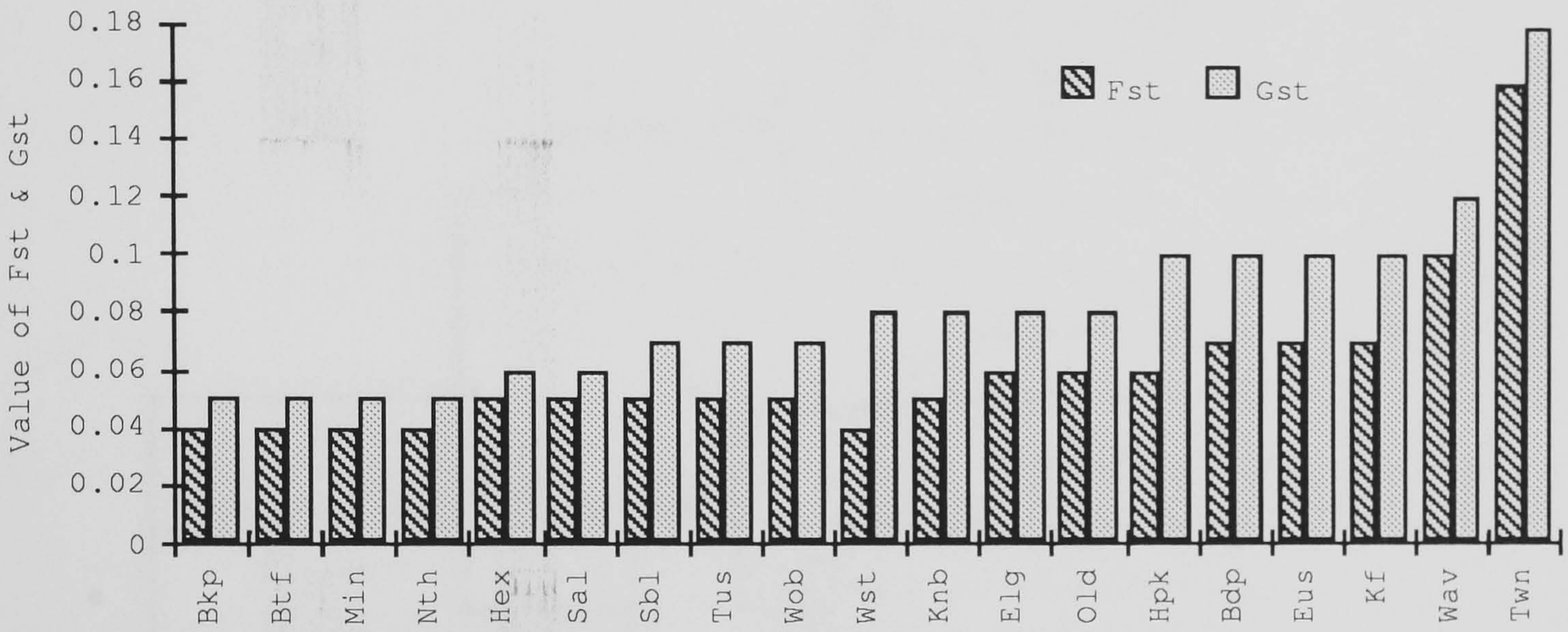


Figure 5·14 The relationship between the values of $G_{ST}$ and $F_{ST}$

However if these data are interpreted in the form of dendrograms constructed by the unweighted pair group method with arithmetic means (UPGMA - Hartl & Clark 1989), then the results are quite different, although there are some similarities. In these dendrograms (Figures 5·15 and 5·17) the lower the pairwise value the greater the level of homogeneity between sub- population 'A' and all other sub-populations.

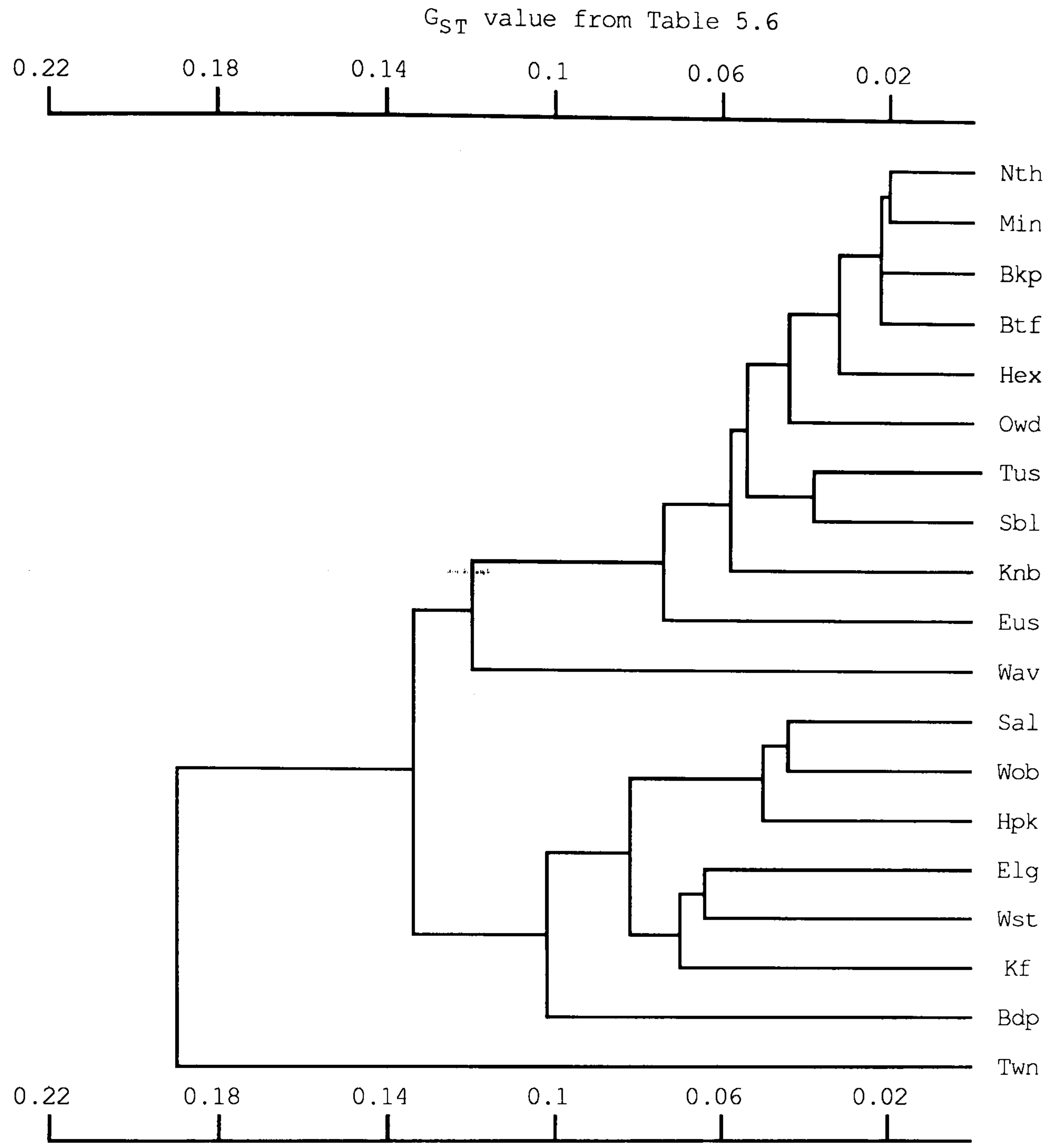


Figure 5·15 Dendrogram based on pairwise $G_{ST}$ values

From Figure 5·15, the dendrogram based on pairwise $G_{ST}$ values, it can be seen that the localities fall into two main clusters. These clusters are, in order of relatedness, [Nth-Min-Bkp-Btf-Hex-Owd-Tus-Sbl-Knb-Eus-Wav] and [Sal-Wob-Hpk-Elg-Wst-Kf-Bdp]. These groupings are interpreted in a geographic context in Figure 5·16. $G_{ST}$ and $F_{ST}$ are not measures of genetic distance, therefore the dendrograms in Figures 5·15 and 5·17 are intended to show, diagrammatically, the relationships between sub-populations in terms of the level of population sub-division.

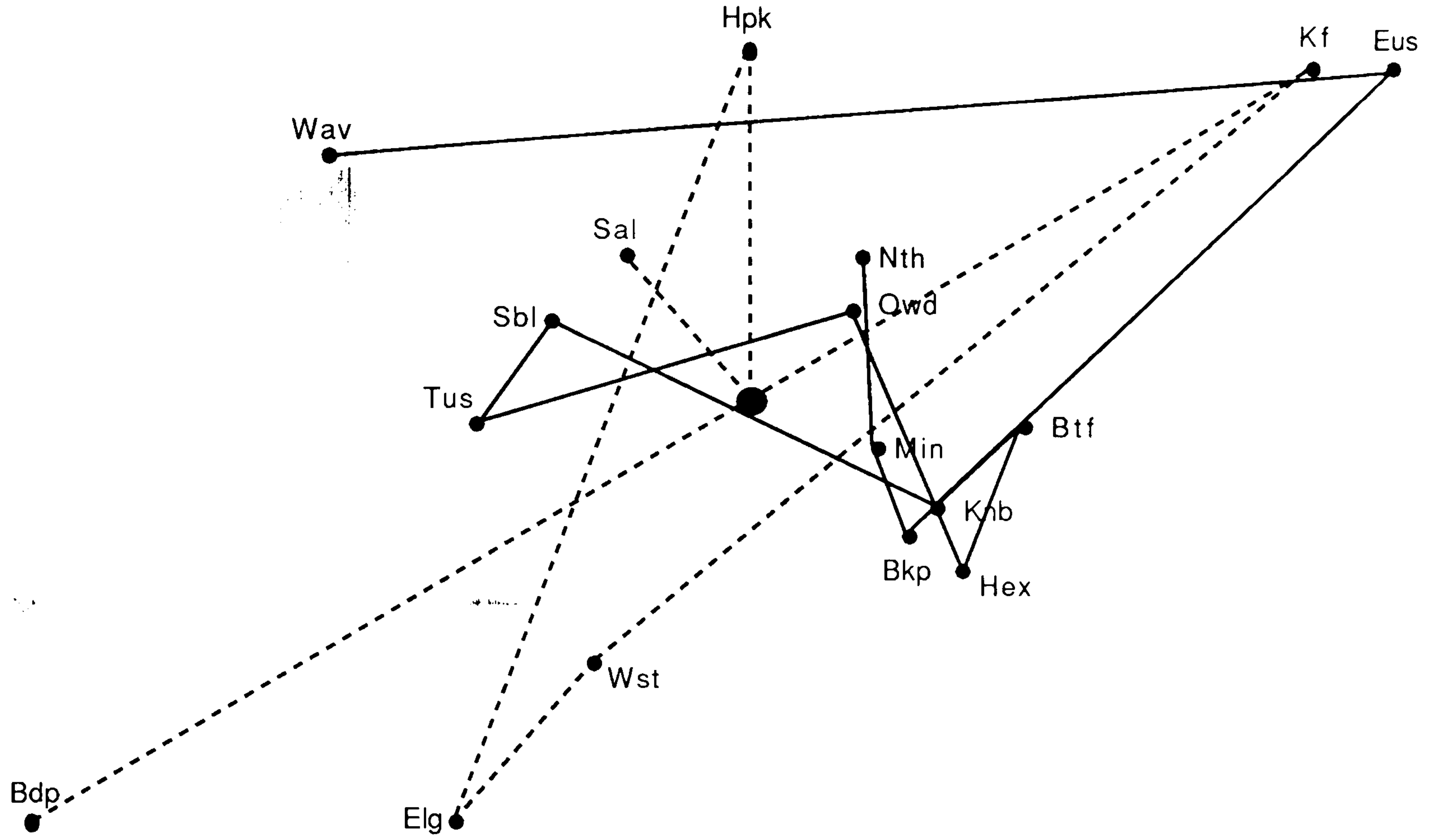


Figure 5·16 Diagrammatic representation of the two main population clusters based on $G_{ST}$ values. Different style lines are only used for clarity.

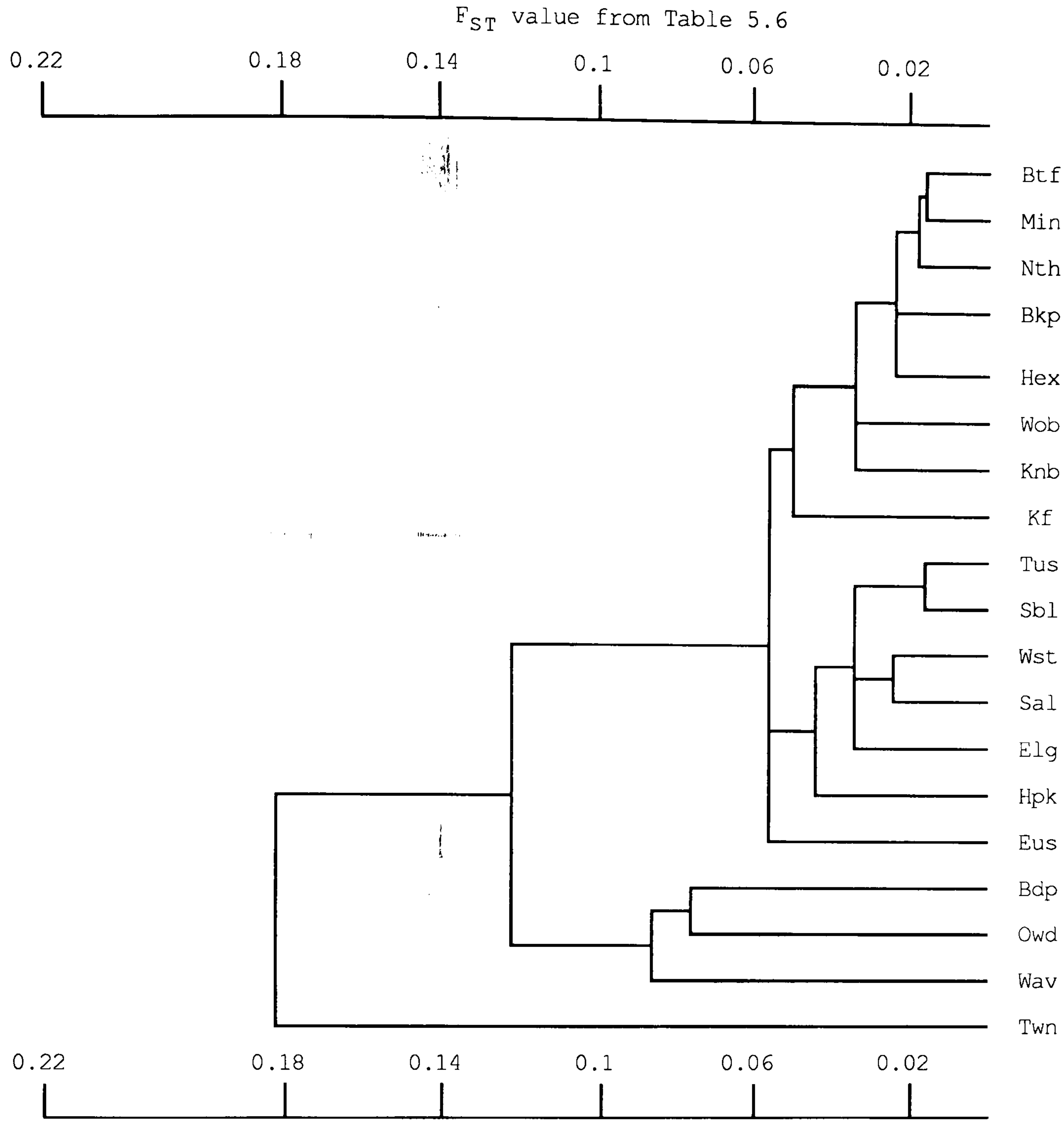


Figure 5·17 Dendrogram based on pairwise $F_{ST}$ values

From Figure 5·17, the dendrogram based on pairwise $F_{ST}$ values, it can be seen that the localities fall into three main clusters. These clusters are, in order of similarity, [Tus-Sbl-Nth-Min-Btf-Bkp-Knb-Owd-Hex], [Sal-Wst-Hpk-Wob-Elg-Bdp] and [Kf-Eus]. These groupings are interpreted in a geographic context in Figure 5·20.

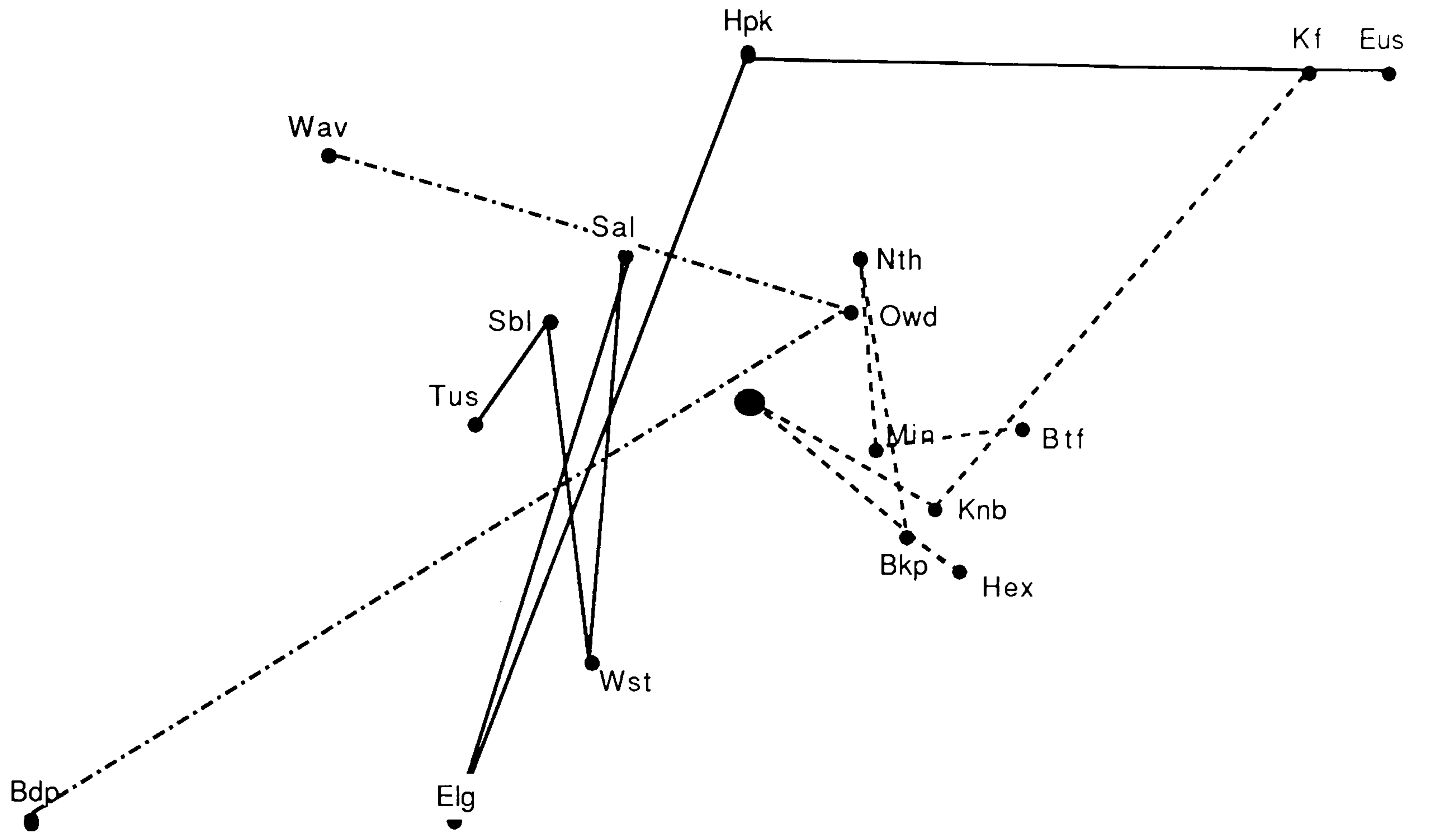


Figure 5·18 Diagrammatic representation of the three main population clusters based on $F_{ST}$ values. Different style lines are only used for clarity.

These dendrograms and associated geographic representations clearly demonstrate the disparity between $F_{ST}$ and $G_{ST}$ in terms of the levels of differentiation between sub-populations. A disparity that is not immediately obvious from Table 5·6 or Figure 5·14.

An inspection of Figures 5·15 to 5·18 will show that there are some similarities. Tusmore (Tus) and Southblock (Sbl) are shown to be closely related by both $F_{ST}$ and $G_{ST}$ analysis, as is the group Buntingford (Btf) - Minsbury Hill (Min) - Northill (Nth) - Brocket Park (Bkp) - Hexton (Hex). It is worth noting that from the latter group Nth and Bkp share two rare alleles, but it should also be noted that Btf shares three rare alleles with King's Forest (Kf) and yet overall Kf is found to be quite dissimilar.

**5·3f2 Comparison between $F_{ST}$ and $G_{ST}$ in the partitioned groups.**

To further clarify population differentiation, the localities were divided into three groups, defined in radiating circles, dependent on their distance from Woburn. The groups were labelled near, medium and far with reference to Figure 2·1 (page 23) and calculations for $F_{ST}$ and $G_{ST}$ applied to each of these groups of sub-populations. The results are listed in Table 5·7 and displayed graphically in Figure 5·19. As previously, the proportionality between the values of $F_{ST}$ and $G_{ST}$ is maintained.

| Group | mean $F_{ST}$ | mean $G_{ST}$ |
|---|---|---|
| Near | 0.052 | 0.051 |
| Near (w) | 0.064 | 0.079 |
| Med | 0.092 | 0.093 |
| Med (w) | 0.089 | 0.102 |
| Far | 0.122 | 0.135 |
| Far (w) | 0.114 | 0.127 |

Table 5·7 $F_{ST}$ and $G_{ST}$ for groups of sub-populations. The inclusion of Woburn is indicated by (w)

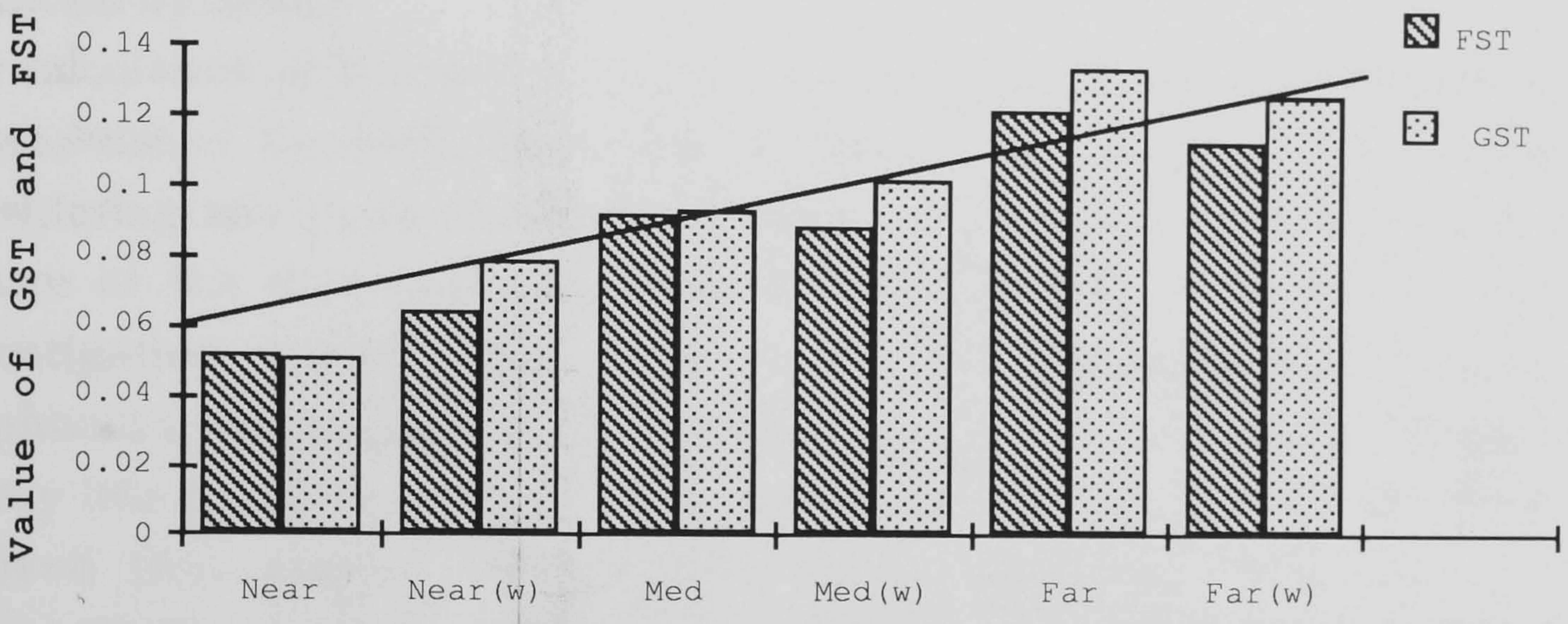


Figure 5·19 Graphic representation of $F_{ST}$ and $G_{ST}$ for groups of sub-populations, with regression line overlay.

The increase in the values of $F_{ST}$ and $G_{ST}$, that is, the increase in genetic differentiation among these groups of sub-populations, can clearly be seen. Each group was analysed twice, once in terms of the sub-populations that make up the group, and once with these populations plus the samples from Woburn Abbey. The inclusion of the Woburn samples (indicated in Table 5·7 and Figure 5·19 by (w)) caused an increase in the $F_{ST}$ and $G_{ST}$ coefficients in both the 'near' and 'medium' distance groups. A decrease in the coefficients was noted in the 'far' distance group. This demonstrates that the sub-population at Woburn Abbey (the putative centre of origin) are in fact more similar to the geographically most distant sub-populations than to the geographically nearer ones. As expected, population differentiation (in terms of these groups) increases with increasing geographical distance (regression coefficient = 0·985, P = 0·0002 t-test [significant])

### 5·3g Spatial autocorrelation between sub-populations.

The first step in obtaining the spatial autocorrelation coefficient of a variable is to map all the localities and then correlate the values for pairs of localities (points). The next step is to calculate a correlation coefficient, in this case Moran's 'I' (Moran 1950), after the methods of Sokal (1978a, 1978b). The final step is to display the data graphically as a 'correlogram'.

Genetic similarity was measured between sub-populations using Moran's autocorrelation coefficient 'I'. For this statistic a value of 1 indicates maximum similarity among neighbouring sub-populations, a value of -1 indicates maximum difference and a value of 0 indicates the null hypothesis that any two sub-populations are no more similar than predicted by chance.

The calculation of Moran's 'I' requires first that a relationship (or distance) is established between the sample locations. The choice of criteria for considering any pairs of sub-populations as neighbours depends on the nature of the data and the aspect that is being investigated. In this investigation I shall take three criteria (1) geographically nearest neighbours, (2) neighbours in terms of expanding radii from Woburn Abbey (the putative centre of origin) and (3) neighbours in terms of a cline derived from ranked microsatellite allele frequencies. Criteria (1) - geographically nearest neighbours - is in effect a two dimensional pattern in which sub-populations are assumed to be derived from the nearest preceding sub-population (in this case with regard to Woburn). There are several geographical routes by which sub-populations can become established. For example from Figure 5·20, as a result of expansion from Woburn the sub-population Sbl may equally have been derived from Sal or Tus or Woburn. If a relationship between nearest neighbours exists then the overall slope of the correlogram would be from positive (representing similarity) to negative (representing dissimilarity). The second criteria describes the expansion from Woburn in terms of a series of expanding concentric circles. In this scenario sub-populations on opposite radii of the first circle would be similar and sub-populations on opposite radii of subsequent circles would display a similarly greater degree of difference. This criteria gives a measure of directionality to the interpretation. The third criteria is based on a cline constructed from the observed allele frequencies, this analysis should show if there is a general correlation between allele frequencies at different sub-populations.

Initially a Gabriel connected graph (Gabriel & Sokal 1969) was constructed on the nearest-neighbour criteria. That is, sample sites were considered to be contiguous only with their nearest neighbour regardless of any intervening geographical features and the assumption was made that distances between any sites are symmetrical (i.e. A to B = B to A). Two assumptions about dispersal are implicit in this. (1) Gene flow occurs equally in all directions; and (2) the probability of gene flow between sites is a linear function of the distances between them. This graph is shown in Figure 5·20.

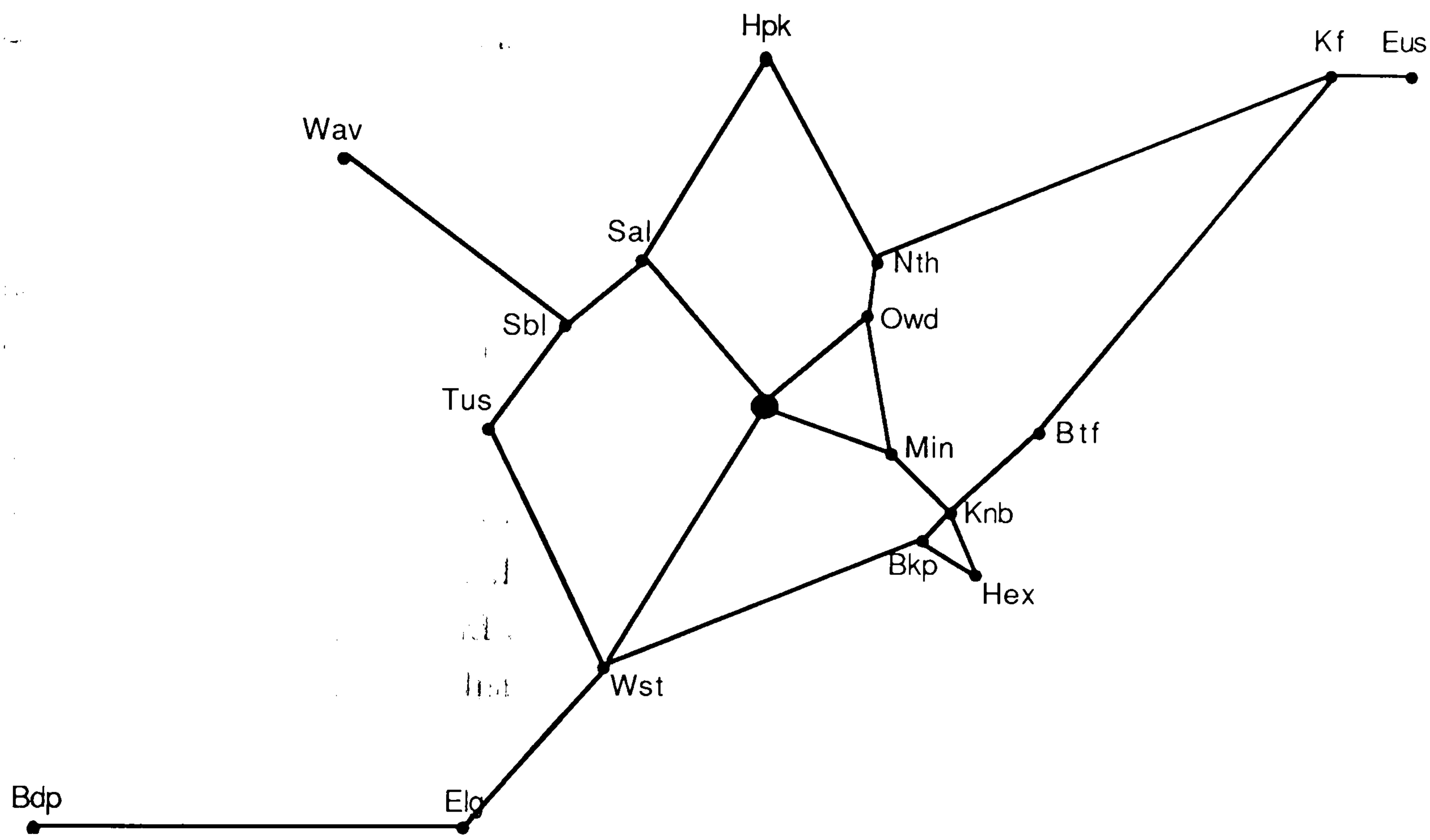


Figure 5·20 Gabriel connected graph based on nearest neighbours.

From this graph an adjacency matrix was constructed (Table 5·8) in which the number of connections traversed (or 'edges') between sub-populations was recorded. From this matrix were derived seven distance class matrices in which each matrix contained only sub-populations with the same number of edges. From this table Waverly (Wav) to Southblock (Sbl) = 1; Wav to Harry's Park (Hpk) = 2; Wav to Eling (Elg) = 4, etc. Separate matrices were then constructed for all distance classes. The matrix for distance class 1 is shown in Table 5·9, as an example.

Adjacency Matrix (Nearest Neighbour)

| | Wav | Tus | Sbl | Sal | Wob | Hpk | Bdp | Elg | Wst | Owd | Nth | Min | Bkp | Knb | Hex | Btf | Kf | Eus |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav | - | 2 | 1 | 2 | 3 | 3 | 5 | 4 | 3 | 4 | 4 | 4 | 4 | 5 | 5 | 6 | 5 | 6 |
| Tus | 2 | - | 1 | 2 | 2 | 3 | 3 | 2 | 1 | 3 | 4 | 3 | 2 | 3 | 3 | 4 | 5 | 6 |
| Sbl | 1 | 1 | - | 1 | 2 | 2 | 4 | 3 | 2 | 3 | 3 | 3 | 3 | 4 | 5 | 5 | 4 | 5 |
| Sal | 2 | 2 | 1 | - | 1 | 1 | 4 | 3 | 2 | 2 | 2 | 2 | 3 | 3 | 4 | 4 | 3 | 4 |
| Wob | 3 | 2 | 2 | 1 | - | 2 | 3 | 2 | 1 | 1 | 2 | 1 | 2 | 2 | 3 | 3 | 3 | 4 |
| Hpk | 3 | 3 | 2 | 1 | 2 | - | 5 | 4 | 3 | 2 | 1 | 3 | 4 | 4 | 5 | 3 | 2 | 3 |
| Bdp | 5 | 3 | 4 | 4 | 3 | 5 | - | 1 | 2 | 4 | 5 | 4 | 3 | 4 | 4 | 5 | 6 | 7 |
| Elg | 4 | 2 | 3 | 3 | 2 | 4 | 1 | - | 1 | 3 | 4 | 3 | 2 | 3 | 3 | 4 | 5 | 6 |
| Wst | 3 | 1 | 2 | 2 | 1 | 3 | 2 | 1 | - | 2 | 3 | 2 | 1 | 2 | 2 | 3 | 4 | 5 |
| Owd | 4 | 3 | 3 | 2 | 1 | 2 | 4 | 3 | 2 | - | 1 | 1 | 3 | 2 | 3 | 3 | 2 | 3 |
| Nth | 4 | 4 | 3 | 2 | 2 | 1 | 5 | 4 | 3 | 1 | - | 2 | 4 | 3 | 4 | 2 | 1 | 2 |
| Min | 4 | 3 | 3 | 2 | 1 | 3 | 4 | 3 | 2 | 1 | 2 | - | 2 | 1 | 2 | 2 | 3 | 4 |
| Bkp | 4 | 2 | 3 | 4 | 2 | 4 | 3 | 2 | 1 | 3 | 4 | 2 | - | 1 | 1 | 2 | 3 | 4 |
| Knb | 5 | 3 | 4 | 3 | 2 | 4 | 4 | 3 | 2 | 2 | 3 | 1 | 1 | - | 1 | 1 | 2 | 3 |
| Hex | 5 | 3 | 4 | 4 | 3 | 5 | 4 | 3 | 2 | 3 | 4 | 2 | 1 | 1 | - | 2 | 3 | 4 |
| Btf | 6 | 4 | 5 | 4 | 3 | 5 | 5 | 4 | 3 | 3 | 2 | 2 | 2 | 1 | 2 | - | 1 | 2 |
| Kf | 5 | 5 | 4 | 3 | 3 | 2 | 6 | 5 | 4 | 2 | 1 | 3 | 3 | 2 | 3 | 1 | - | 1 |
| Eus | 6 | 6 | 5 | 4 | 4 | 3 | 7 | 6 | 5 | 3 | 2 | 4 | 4 | 3 | 4 | 2 | 1 | - |

Table 5·8 Adjacency Matrix (Nearest neighbours)

Distance Class 1 (1 edge)

| | Wav | Tus | Sbl | Sal | Wob | Hpk | Bdp | Elg | Wst | Owd | Nth | Min | Bkp | Knb | Hex | Btf | Kf | Eus | *W* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav | - | | 1 | | | | | | | | | | | | | | | | 1 |
| Tus | | - | 1 | | | | | | 1 | | | | | | | | | | 2 |
| Sbl | 1 | 1 | - | 1 | | | | | | | | | | | | | | | 3 |
| Sal | | | 1 | - | 1 | 1 | | | | | | | | | | | | | 3 |
| Wob | | | | 1 | - | | | | 1 | 1 | | 1 | | | | | | | 4 |
| Hpk | | | | 1 | | - | | | | | 1 | | | | | | | | 2 |
| Bdp | | | | | | | - | 1 | | | | | | | | | | | 1 |
| Elg | | | | | | | 1 | - | 1 | | | | | | | | | | 2 |
| Wst | | 1 | | | 1 | | | 1 | - | | | | 1 | | | | | | 4 |
| Owd | | | | | 1 | | | | | - | 1 | 1 | | | | | | | 3 |
| Nth | | | | | | 1 | | | | 1 | - | | | | | | 1 | | 3 |
| Min | | | | | 1 | | | | | 1 | | - | | 1 | | | | | 3 |
| Bkp | | | | | | | | | 1 | | | | - | 1 | 1 | | | | 3 |
| Knb | | | | | | | | | | | | 1 | 1 | - | 1 | 1 | | | 4 |
| Hex | | | | | | | | | | | | | 1 | 1 | - | | | | 2 |
| Btf | | | | | | | | | | | | | | 1 | | - | 1 | | 2 |
| Kf | | | | | | | | | | | 1 | | | | | 1 | - | 1 | 3 |
| Eus | | | | | | | | | | | | | | | | | 1 | - | 1 |
| | | | | | | | | | | | | | | | | | | | 46 |

Table 5·9 Distance Class 1 Matrix (1 edge)

*W* = matrix weight

Moran's coefficient 'I' was then calculated according to equation 5·1

$$I = \frac{n\sum_{ij} w_{ij} z_i z_j}{W\sum_{i=1}^{n} z_i^2} \qquad 5\cdot1$$

Where: n = number of sub-populations
$w_{ij}$ = the weight between the pairs of sub-populations i,j
$z_i$ = the deviate of i
$z_j$ = the deviate of j
W = the sum of the matrix weights (equals the number of ones in the adjacency matrix, or twice the number of 'edges' connecting the sub-populations in the graph).

The variance, standard error, expected value and difference from expectation were calculated after Sokal and Oden (1976).
Calculation of the critical value of 'I' was after Sokal and Oden (1978a).
Except for rare alleles, which were excluded, these series of calculations were performed for every distance class of every allele at every locus.
The values for 'I' were visualised graphically using 'correlograms' (Sokal & Oden 1978a, 1978b). Composite correlograms were constructed for each locus. Figures 5·21a to 5·21g display plots for Moran's 'I' as a function of distance class (number of edges between sub-populations), and include all alleles at all seven loci. In each figure the ordinate is Moran's 'I' and the abscissa the distance classes one to seven.

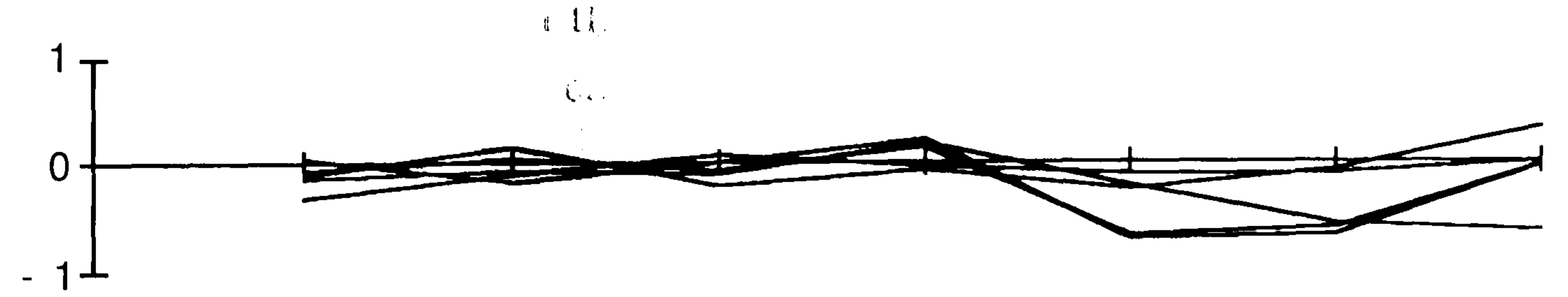


Figure 5·21a Nearest Neighbour Intervals - Locus JP23

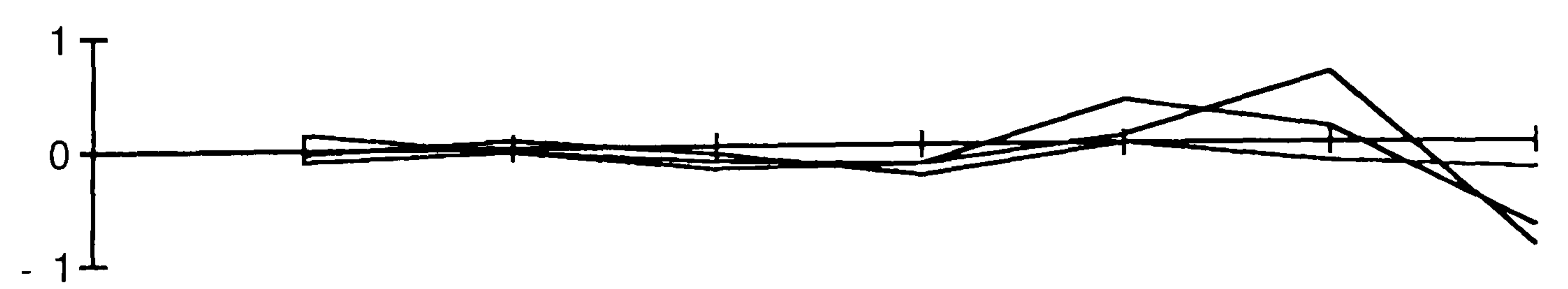


Figure 5·21b Nearest Neighbour Intervals - Locus JP27

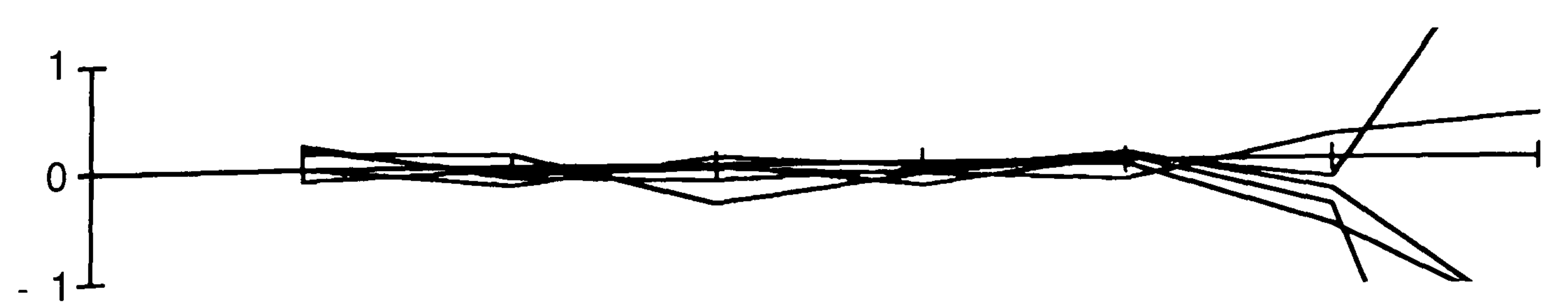


Figure 5·21c Nearest Neighbour Intervals - Locus JP38

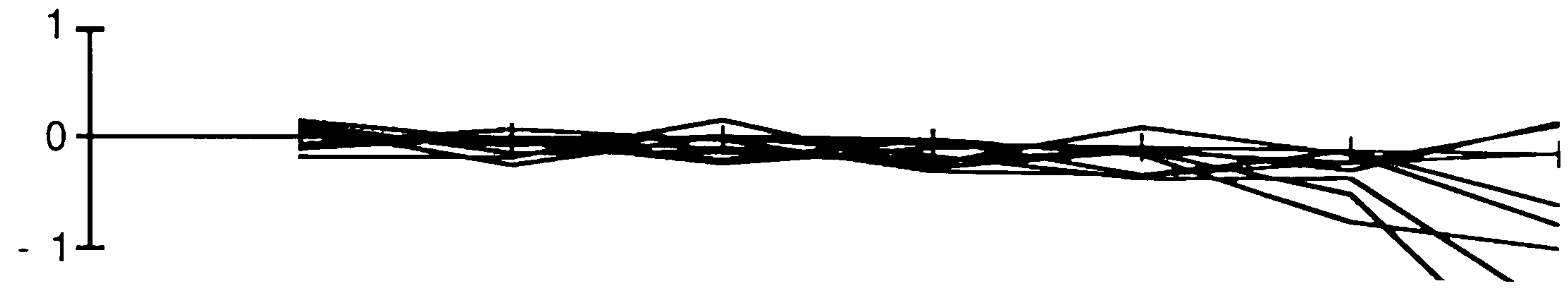


Figure 5·21d Nearest Neighbour Intervals - Locus BOVIRBP

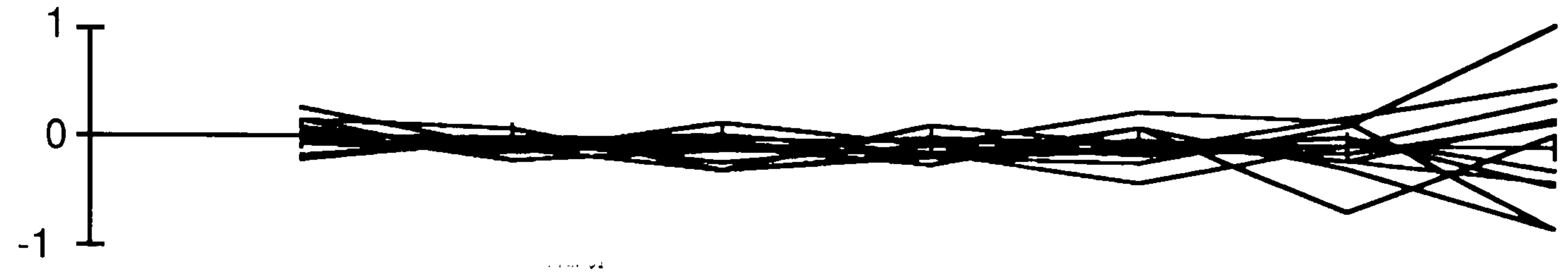


Figure 5·21e Nearest Neighbour Intervals - Locus MAF23

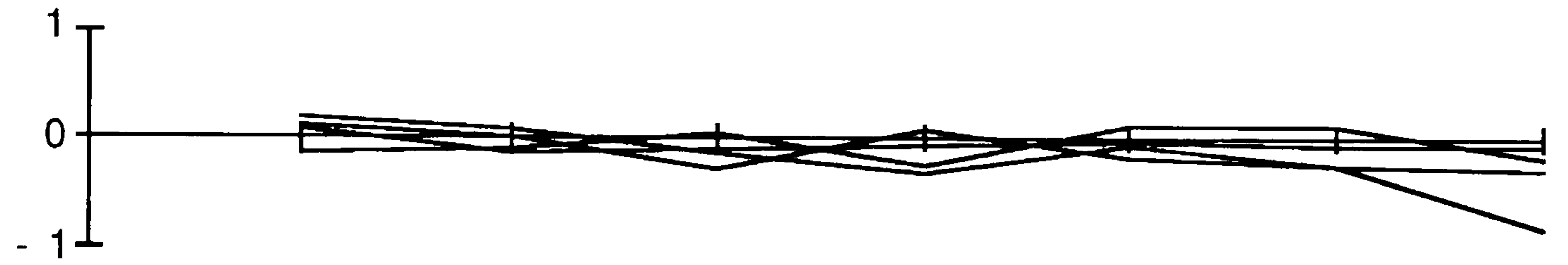


Figure 5·21f Nearest Neighbour Intervals - Locus MAF65

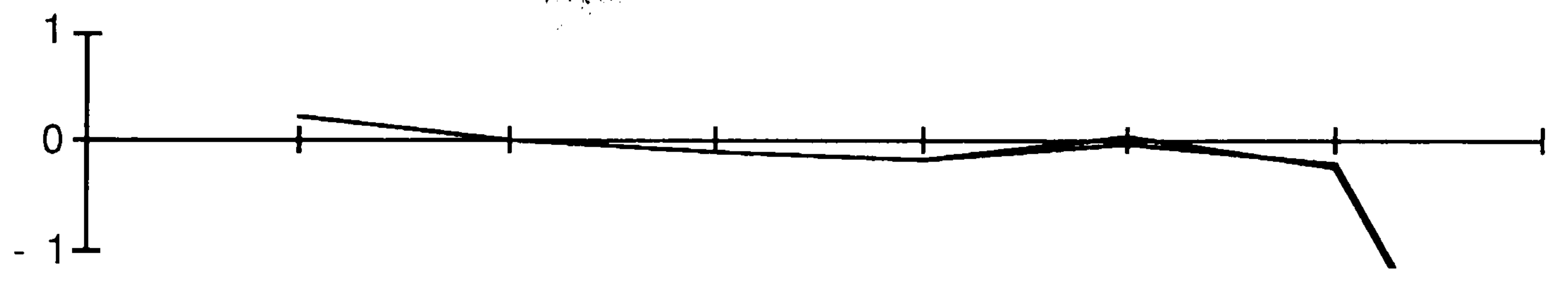


Figure 5·21g Nearest Neighbour Intervals - Locus DRBP1

These correlograms are essentially the same, the values for Moran's 'I' oscillating about the zero value, with some excursions after distance class five. In all distance classes there are no significant correlations and thus I conclude that the distribution of the alleles amongst the sub-populations is no more than would be expected by random chance. If the muntjac had radiated in a regular way from the original location (Woburn Abbey) and assuming that there was still a limited amount of gene flow, then genetic distance would be expected to increase the further away the deer migrated. Nearest neighbours would be the most similar (distance class 1) and furthest neighbours the most dissimilar (distance class 7). Spatial autocorrelation would be able to identify this and a general result should appear as in Figure 5·22.

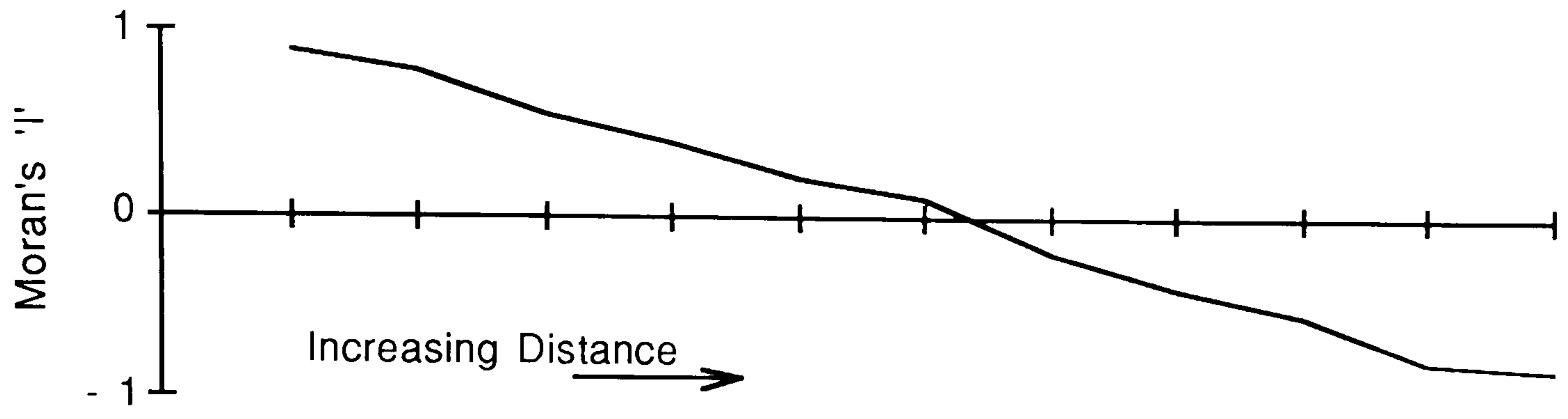


Figure 5·22 Generalised graph of 'I' against some distance.

However, the geographically nearest-neighbour criteria may be more appropriate for a linear migration rather than a radiational one - which may have occurred here. If this is the case then individuals equi-distant from the centre of origin, yet far apart (i.e. on opposite radii) may be very similar. Genetic differences should therefore be apparent concentrically. To test this hypothesis distance classes were erected based on the straight-line distance (in kilometres) between sub-populations. A distance matrix was constructed and the data partitioned into five distance classes: up to 30km, 31km to 60km, 61km to 90km, 91km to 120km and over 121km. The statistics were performed as described above and the value of Moran's 'I' visualised graphically, as shown in Figures 5·23a to 5·23g.

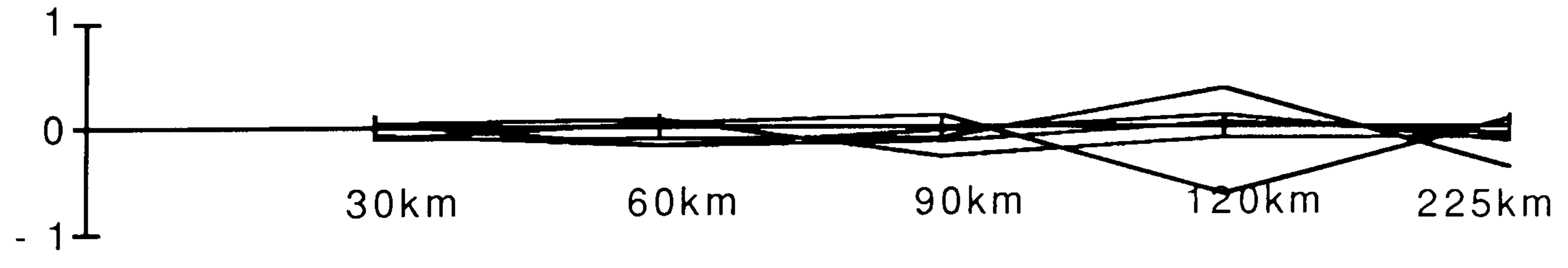


Figure 5·23a Distance Classes - Locus JP23

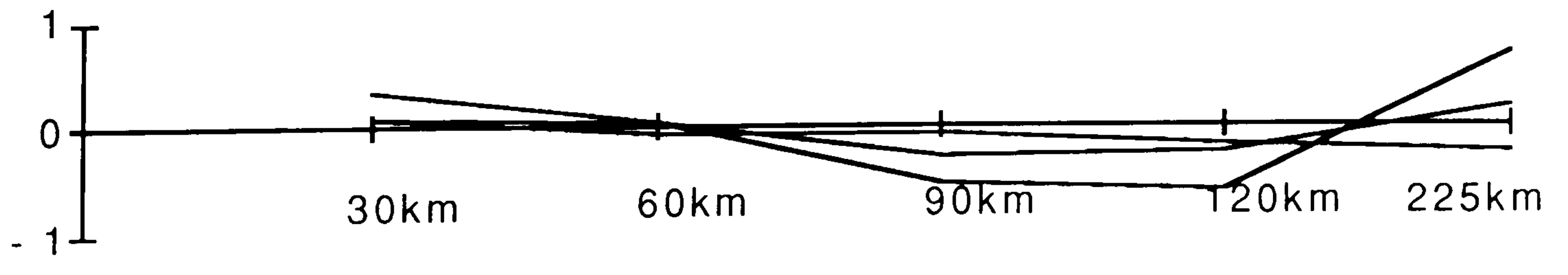


Figure 5·23b Distance Classes - Locus JP27

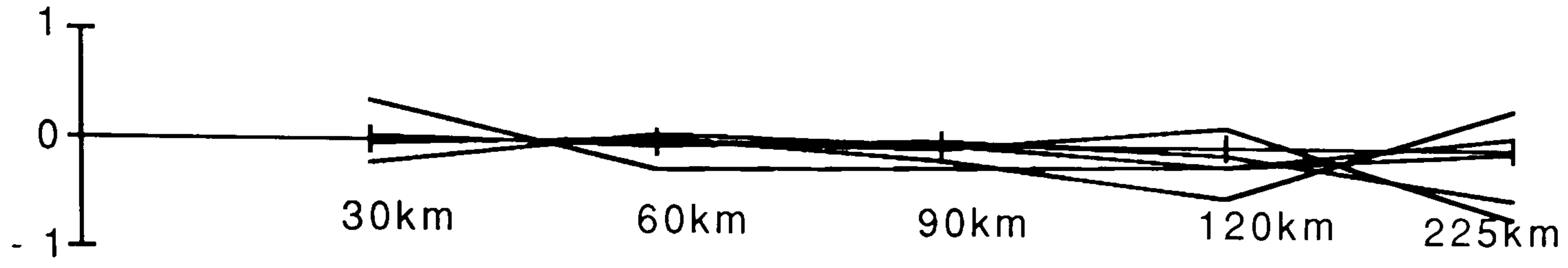


Figure 5·23c Distance Classes - Locus JP38

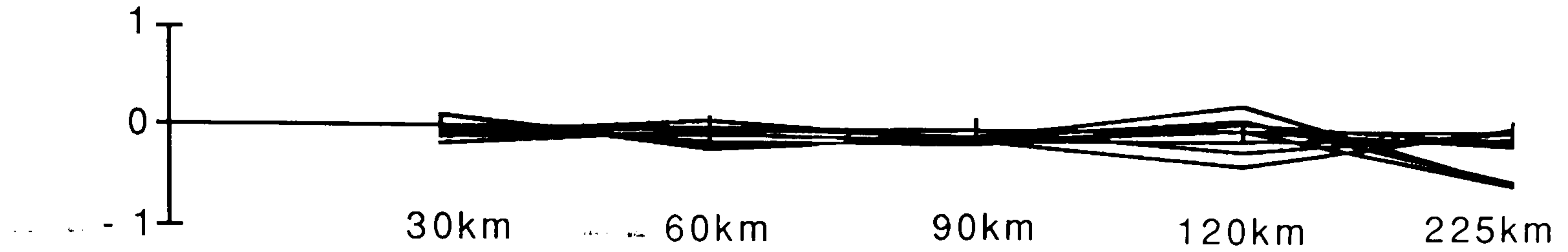


Figure 5·23d Distance Classes - Locus BOVIRBP

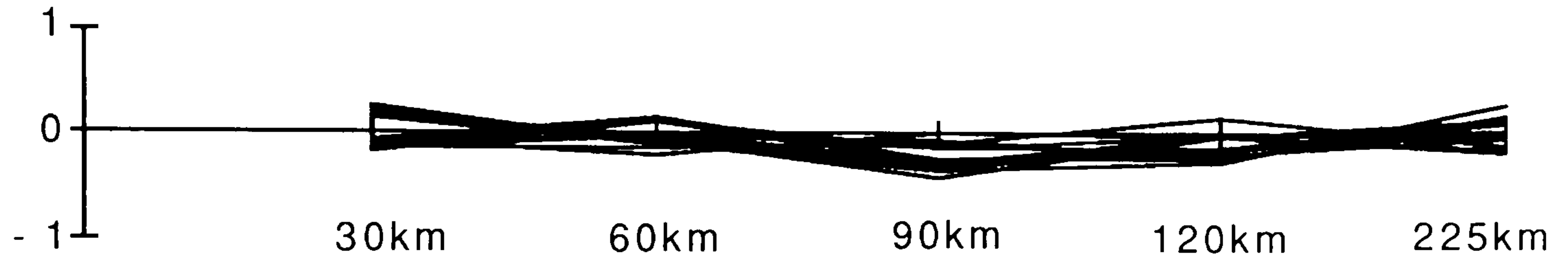


Figure 5·23e Distance Classes - Locus MAF23

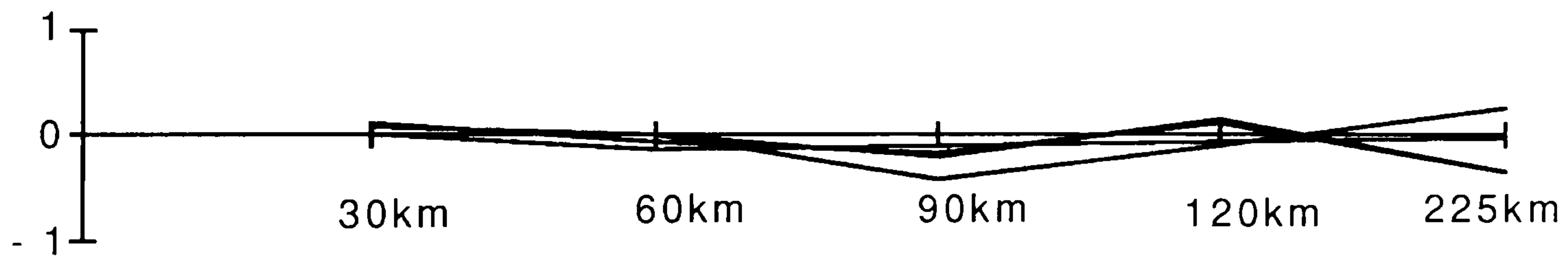


Figure 5·23f Distance Classes - Locus MAF65

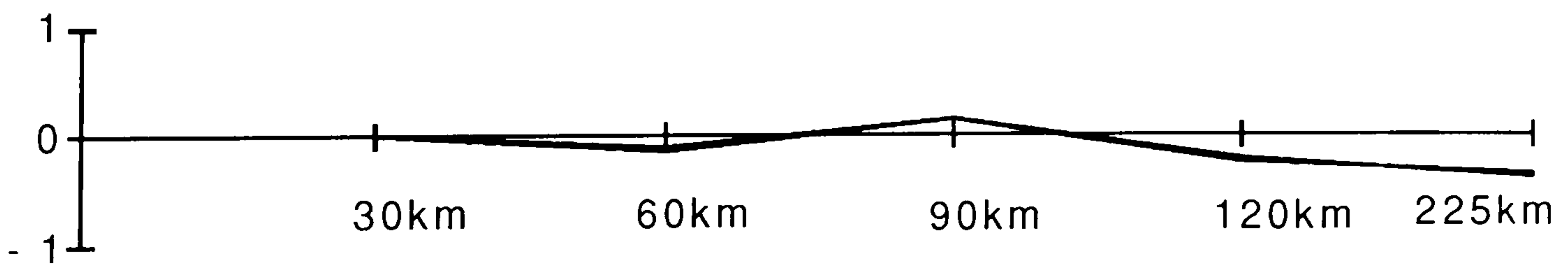


Figure 5·23g Distance Classes - Locus DRBP1

From an inspection of the preceding graphs it can be seen that the results from both of these two approaches are essentially similar, and as neither had found evidence of a migration pattern, another approach was considered. The relationships between sub-populations were interpreted from the frequency of occurrence of two particular alleles - allele 144 from locus JP23 and allele 183 from locus BOVIRBP. These two alleles were selected on the basis of being present in every sub-population, and both presenting a definite cline. In both cases the sub-populations were connected by a Gabriel-connected graph in a cline following the frequency of the two alleles. For JP23 (144), starting at Waverly - being the lowest frequency (0·20) and progressing to Hexton, being the sub-population with the greatest frequency (0·586). For BOVIRBP(183), starting at Hexton - being the lowest frequency (0·052) and progressing to Southblock, being the sub-population with the greatest frequency (0·479). If other alleles at the same loci, and other loci, are distributed in a similar fashion then their correlograms should follow these clines. This method produced 12 distance classes in the case of JP23 and 14 distance classes in the case of BOVIRBP. Only the correlograms for JP23 and BOVIRBP are shown as the correlograms for the other loci follow the same general pattern. The Gabriel-connected graph and correlogram for locus JP23 (144) are shown in Figures 5·24 and 5·25. Those for locus BOVIRBP(183) are shown in Figures 5·26 and 5·27.

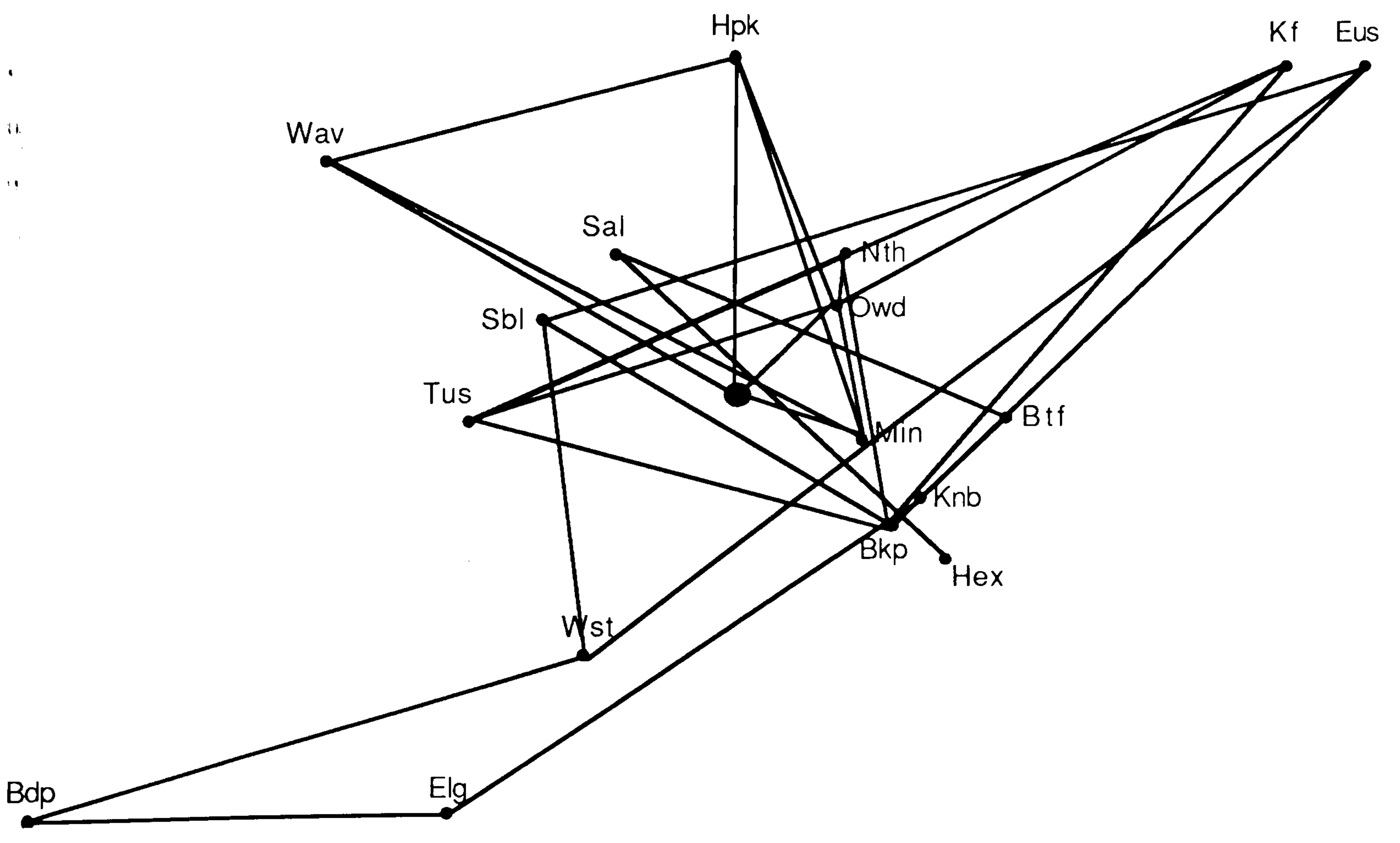


Figure 5·24 Gabrial connected graph based on locus JP23 - frequency of allele 144 (28 edges)

The statistical analysis was performed as previously described.

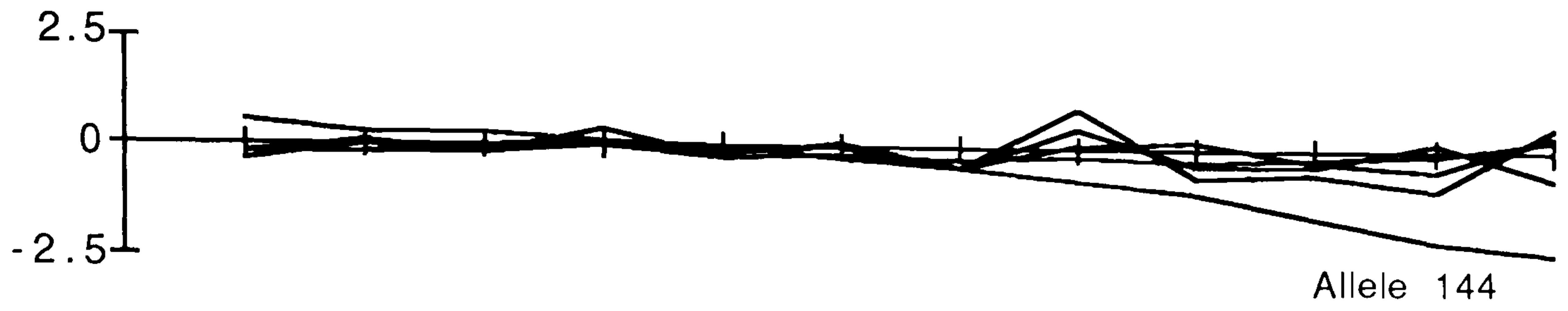


Figure 5·25 Distance Classes based on JP23(144)

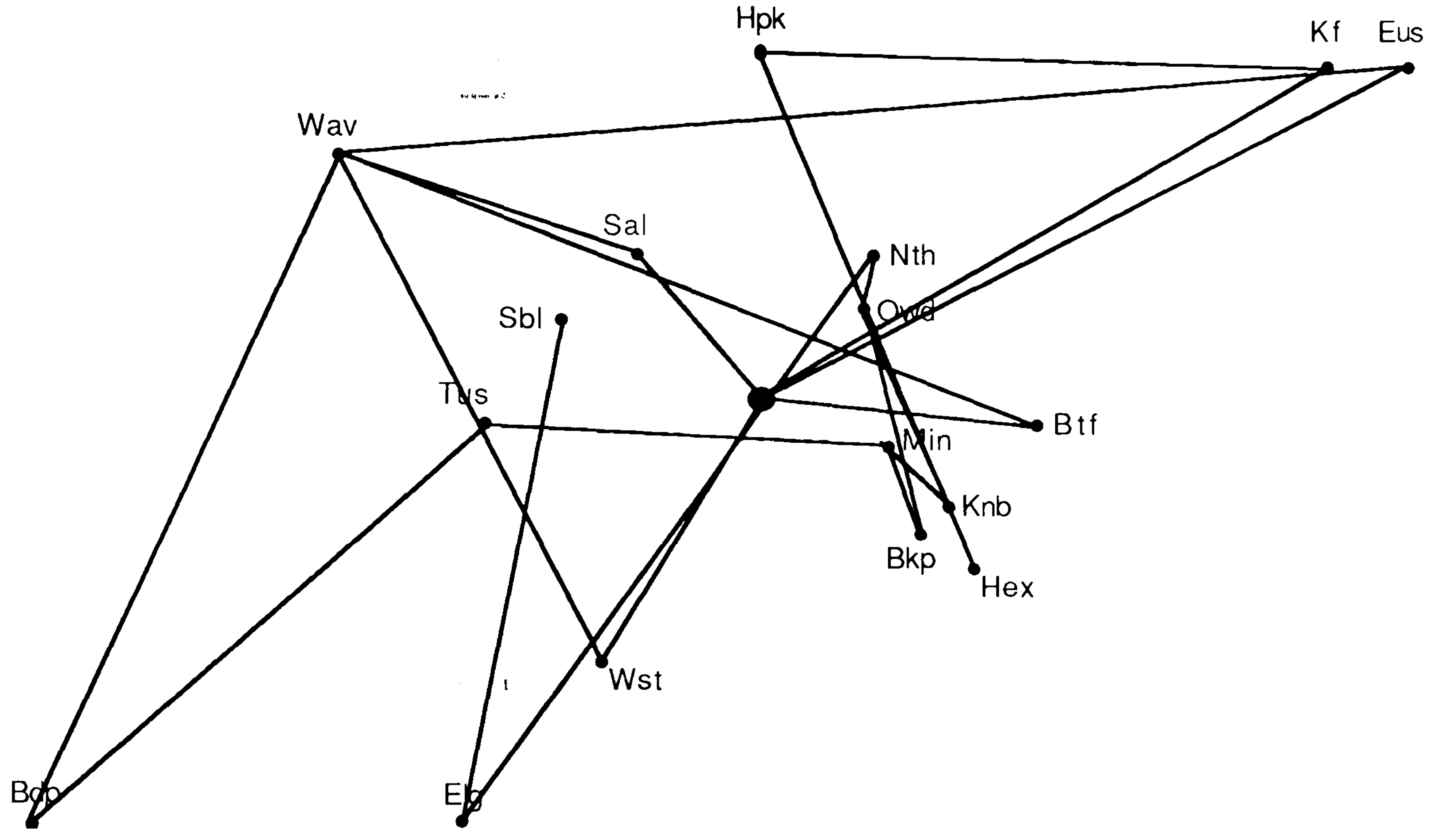


Figure 5·26 Gabriel connected graph based on locus BOVIRBP - frequency of allele 183 (21 edges)

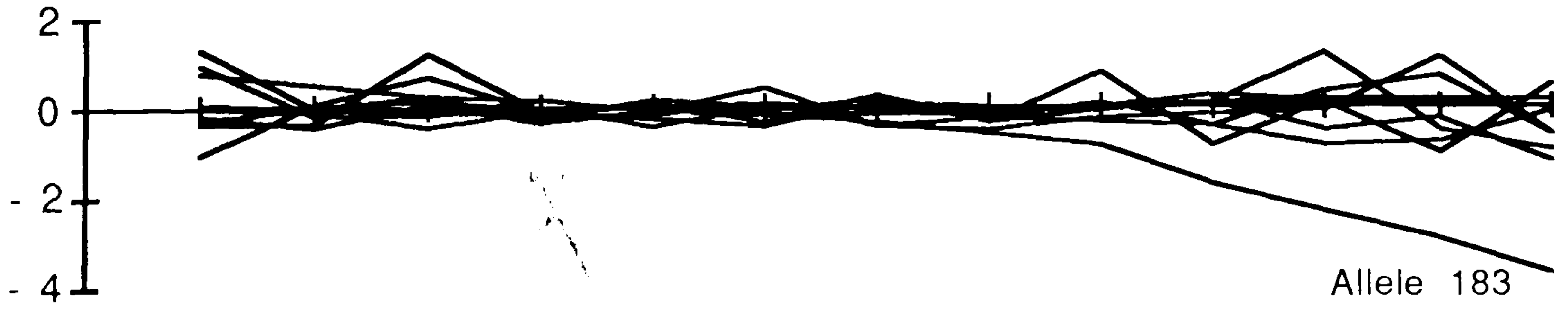


Figure 5·27 Distance Classes based on BOVIRBP(183)

In these sets of correlograms, as well as in the two previous sets, the data oscillate about the zero value thereby demonstrating that there is no correlation between sub-populations that may be considered as neighbours in terms of the arranged clines. With reference to the investigation based on the created clines of JP23(144) and BOVIRBP(183), other alleles at these loci fail to follow the cline, demonstrating that there is no correlation between alleles at the same locus.

### 5·3h Sub-Population Differentiation

An estimate of the migration rate between sub-populations (N*m*) was obtained as previously described in Chapter four (section 4·3h, page 113). Table 5·11 lists the results with values of $G_{ST}$ above the diagonal and values of N*m* below the diagonal.

Scatter diagrams (Figure 5·28a-s) were constructed as previously described - to investigate the correlation between estimated migration rate (N*m*) and geographical distance from each sub-population to all other sub-populations, and for all sub-populations together. Table 5·10 summarises the results of the scatter plots.

| Sub-population | $r^2$ | Sub-population | $r^2$ |
|---|---|---|---|
| Waverly | 0.149 | Tusmore | 0.154 |
| Southblock | 0.097 | Salcey Forest | 0.028 |
| Woburn Abbey | 0.058 | Harry's Park | 0.168 |
| Bowden Park | 0.160 | Eling Common | 0.042 |
| Wormsley Estate | 0.074 | Oldwarden | 0.371 |
| Northill | 0.190 | Minsbury Hill | 0.598 * |
| Brocket Park | 0.531 * | Knebworth | 0.398 |
| Hexton | 0.533 * | Buntingford | 0.486 * |
| King's Forest | 0.082 | Euston | 0.341 |

Table 5·10 Summary of correlation coefficients from scatter plots in Figure 5·30a-r

Four of the sub-populations (marked by an asterisk in table 5·11) show a significant (at the 95% level) correlation between values for N*m* and geographic distance. The remaining sub-populations have $r^2$ values that are not significantly different from zero. This result suggests that differentiation by distance may be significant from the 'point-of-view' of four of the sub-populations.

$G_{ST}$ above the diagonal, N*m* below the diagonal

| | Wav | Tus | Sbl | Sal | Wob | Hpk | Bdp | Elg | Wst | Old | Nth | Min | Bkp | Knb | Hex | Btf | Kf | Eus |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav | | 0.079 | 0.09 | 0.112 | 0.12 | 0.169 | 0.141 | 0.155 | 0.127 | 0.13 | 0.121 | 0.096 | 0.102 | 0.087 | 0.124 | 0.091 | 0.146 | 0.14 |
| Tus | 2.915 | | 0.037 | 0.059 | 0.08 | 0.091 | 0.1 | 0.099 | 0.072 | 0.054 | 0.043 | 0.041 | 0.056 | 0.06 | 0.053 | 0.042 | 0.09 | 0.09 |
| Sbl | 2.528 | 6.507 | | 0.059 | 0.076 | 0.092 | 0.095 | 0.105 | 0.072 | 0.059 | 0.037 | 0.036 | 0.045 | 0.05 | 0.049 | 0.037 | 0.098 | 0.08 |
| Sal | 1.982 | 3.987 | 3.987 | | 0.044 | 0.051 | 0.055 | 0.051 | 0.056 | 0.064 | 0.055 | 0.044 | 0.038 | 0.065 | 0.055 | 0.05 | 0.093 | 0.13 |
| Wob | 1.833 | 2.875 | 3.039 | 5.432 | | 0.047 | 0.075 | 0.052 | 0.069 | 0.087 | 0.072 | 0.055 | 0.037 | 0.05 | 0.075 | 0.064 | 0.055 | 0.12 |
| Hpk | 1.229 | 2.497 | 2.467 | 4.652 | 5.069 | | 0.105 | 0.087 | 0.095 | 0.117 | 0.091 | 0.09 | 0.083 | 0.112 | 0.098 | 0.103 | 0.102 | 0.16 |
| Bdp | 1.523 | 2.25 | 2.382 | 4.295 | 3.083 | 2.131 | | 0.077 | 0.096 | 0.099 | 0.089 | 0.091 | 0.067 | 0.1 | 0.086 | 0.084 | 0.149 | 0.19 |
| Elg | 1.363 | 2.275 | 2.131 | 4.652 | 4.558 | 2.624 | 2.997 | | 0.065 | 0.082 | 0.068 | 0.063 | 0.038 | 0.072 | 0.071 | 0.066 | 0.07 | 0.13 |
| Wst | 1.719 | 3.222 | 3.222 | 4.214 | 3.373 | 2.382 | 2.354 | 3.596 | | 0.059 | 0.045 | 0.039 | 0.037 | 0.309 | 0.053 | 0.046 | 0.07 | 0.1 |
| Old | 1.673 | 4.38 | 3.987 | 3.656 | 2.624 | 1.887 | 2.275 | 2.799 | 3.987 | | 0.044 | 0.034 | 0.061 | 0.075 | 0.039 | 0.045 | 0.121 | 0.11 |
| Nth | 1.816 | 5.564 | 6.507 | 4.295 | 3.222 | 2.497 | 2.559 | 3.426 | 5.306 | 5.432 | | 0.021 | 0.034 | 0.046 | 0.027 | 0.023 | 0.023 | 0.07 |
| Min | 2.354 | 5.848 | 6.694 | 5.432 | 4.295 | 2.528 | 2.497 | 3.718 | 6.16 | 7.103 | 11.65 | | 0.029 | 0.032 | 0.028 | 0.022 | 0.072 | 0.06 |
| Bkp | 2.201 | 4.214 | 5.306 | 6.329 | 6.507 | 2.762 | 3.481 | 6.329 | 6.507 | 3.848 | 7.103 | 8.371 | | 0.027 | 0.037 | 0.023 | 0.046 | 0.08 |
| Knb | 2.624 | 3.917 | 4.75 | 3.596 | 4.75 | 1.982 | 2.25 | 3.222 | 0.559 | 3.083 | 5.185 | 7.563 | 9.009 | | 0.05 | 0.05 | 0.059 | 0.06 |
| Hex | 1.766 | 4.467 | 4.852 | 4.295 | 3.083 | 2.301 | 2.657 | 3.271 | 4.467 | 6.16 | 9.009 | 8.679 | 6.507 | 4.75 | | 0.034 | 0.107 | 0.09 |
| Btf | 2.497 | 5.702 | 6.507 | 4.75 | 3.656 | 2.177 | 2.726 | 3.538 | 5.185 | 5.306 | 10.62 | 11.11 | 10.62 | 4.75 | 7.103 | | 0.078 | 0.07 |
| Kf | 1.462 | 2.528 | 2.301 | 2.438 | 4.295 | 2.201 | 1.428 | 3.321 | 3.321 | 1.816 | 10.62 | 3.222 | 5.185 | 3.987 | 2.086 | 2.955 | | 0.08 |
| Eus | 1.523 | 2.657 | 2.726 | 1.658 | 1.905 | 1.284 | 1.087 | 1.688 | 2.354 | 2.108 | 3.321 | 3.718 | 2.915 | 3.917 | 2.438 | 3.271 | 2.915 | |

Table 5·11 Pairwise Values for $G_{ST}$ and N*m*

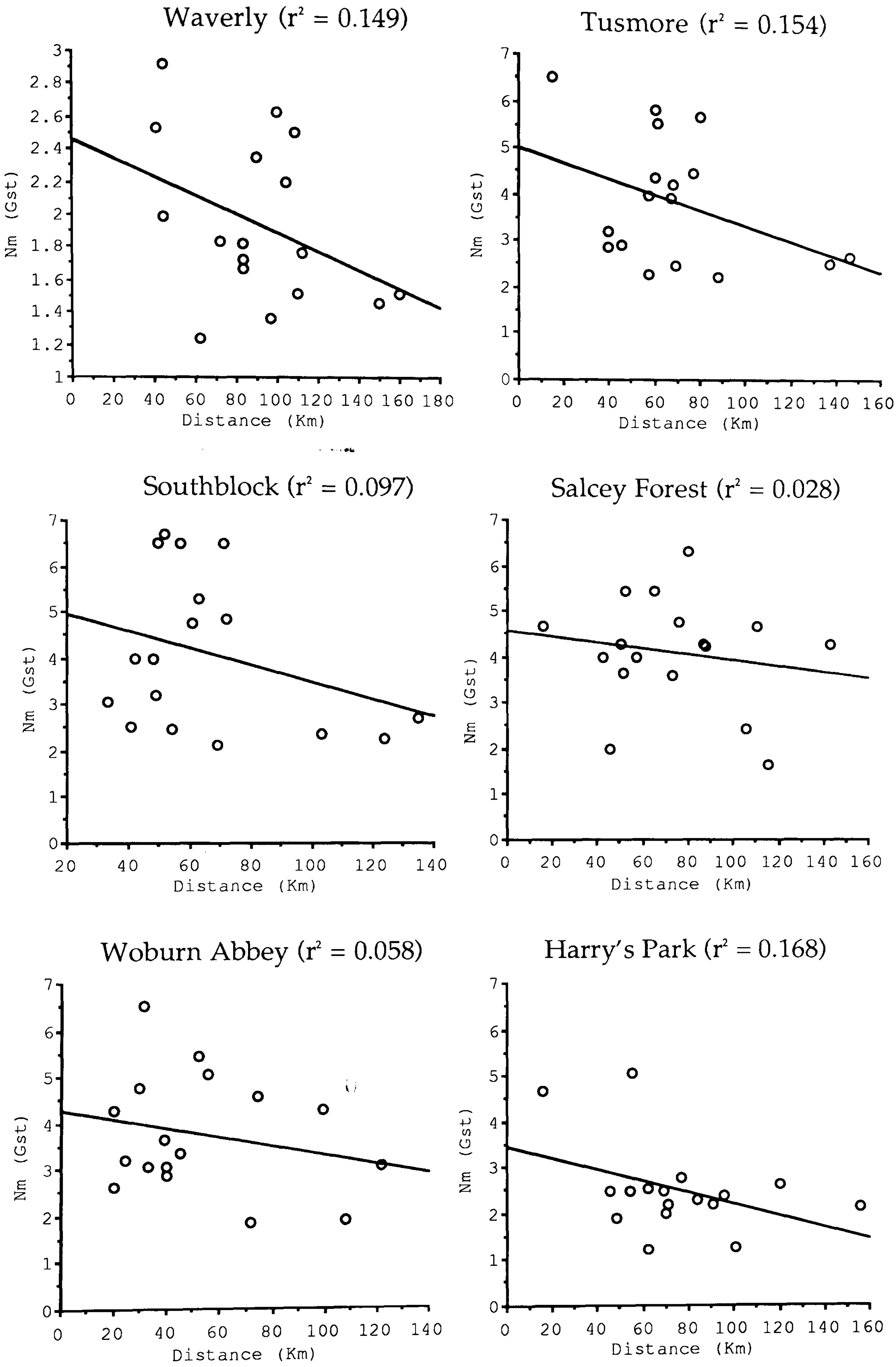


Figure 5·28a - f

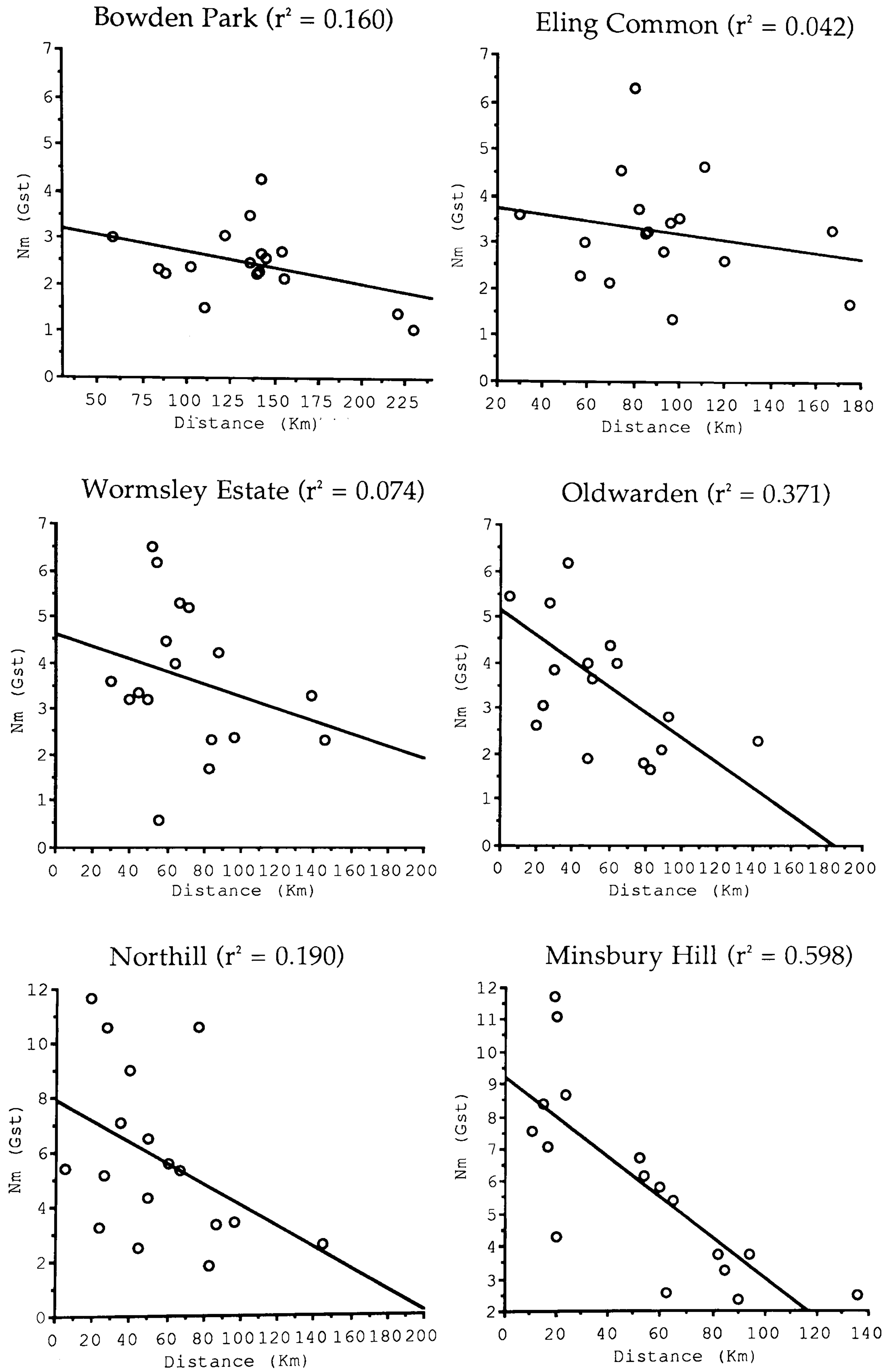


Figure 5·28g - 1

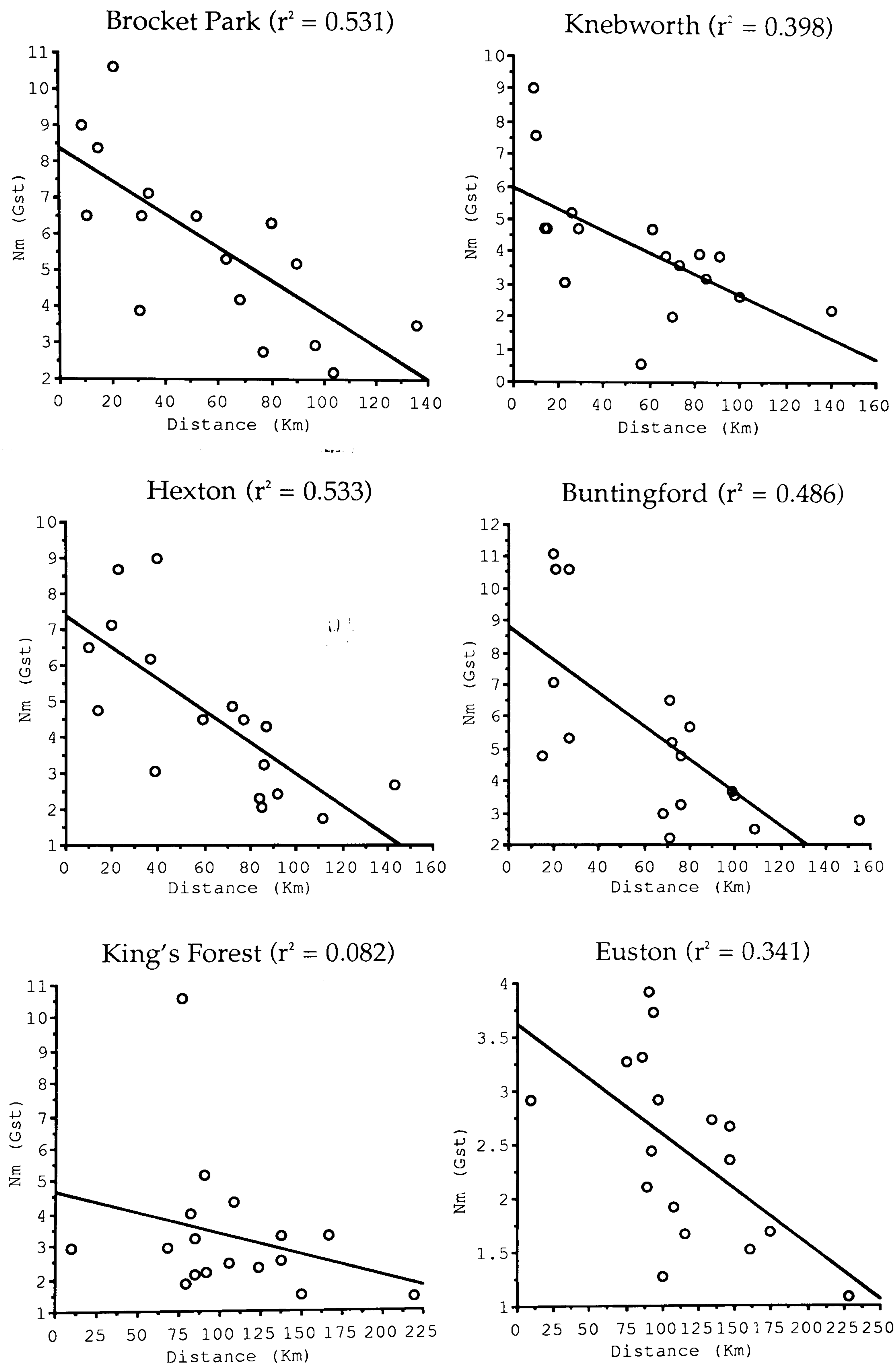


Figure 5·28m - r

All Sub-populations ($r^2$ = 0.006)

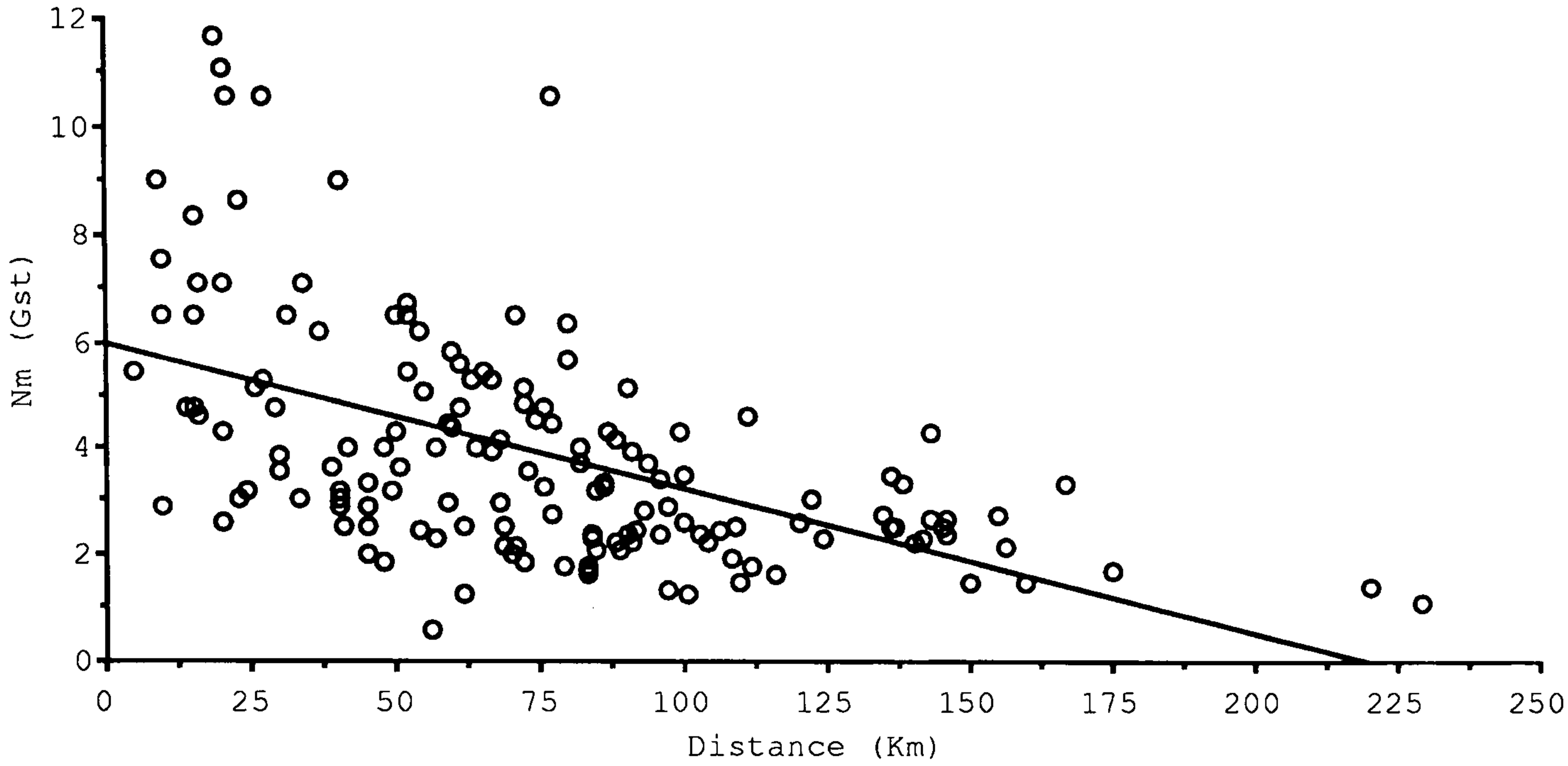


Figure 5·28s

## 5·4 DISCUSSION

In this chapter I have investigated the spread and distribution of muntjac deer throughout the study area by analysis of microsatellite loci. Twenty one primer pairs were screened but only seven proved to be informative in terms of the requirements for this study. The remaining primer pairs were found to be monomorphic or could not be amplified. The primer pairs used in this work were originally discovered and used in analysis of Cervine (JP23 - JP27 - JP38), Ovine (MAF23 - MAF65) and Bovine (BOVIRBP - DRBP1) investigations. Previous investigations at the level of population and sub-population have generally used more loci and primers taken from the species under investigation - or a closely related species. MacHugh *et al* (1994) in their investigation of microsatellite variation in European cattle breeds used bovine primer pairs to amplify twelve microsatellite loci and produced a range of two to eleven alleles.

Roy *et al* (1994) in a study of the pattern of differentiation and hybridisation in North American wolf-like canids used ten microsatellite primer pairs identified from a domestic dog genomic library (Ostrander *et al* , 1993). These ten loci produce a range of four to twenty alleles per locus.

Taylor *et al* (1994) in an analysis of the genetic variation of simple sequence loci in the northern hairy-nosed wombat used sixteen wombat derived loci with a range of four to ten alleles.

Gottelli *et al* (1994) in an appraisal of the molecular genetics of the Ethiopian wolf used nine microsatellite loci identified from the domestic dog genomic library. These loci produced two to six alleles per locus.

Although a lesser number of loci have been used in this investigation they have proved sufficiently informative to be able to elucidate relationships between sub-populations. In this study the primer pair JP23, originally derived from red deer (*Cervus elaphus*), produced seven alleles - whereas in red deer it produced five visible alleles plus a non-amplifying allele (J. Pemberton pers. comm.). From the same source JP27 produced five alleles (six in red deer) and JP38 also produce five alleles (eight in red deer). The primers for locus MAF65 (Buchanan *et al* 1992) produced six alleles in muntjac deer, yet only three alleles in the sheep from which they were derived. Although Buchanan *et al* (1992) reported that the MAF65 locus could not be amplified in red deer, this locus has been found to produce two visible alleles plus a non-amplifying allele in red deer (J. Pemberton pers. comm.). The primers MAF23 (Swarbrick *et al* 1990) produced 17 alleles in muntjac deer and only four in the sheep from which they were derived. The locus BOVIRBP (J. Pemberton pers. comm., primers originally from D. MacHugh) was derived from a bovine genome (from the GenBank data base), this locus produced six visible alleles plus a non-amplifying allele in red deer (J. Pemberton pers. comm.) and sixteen visible alleles in muntjac.

A report by Moore *et al* (1991) found that levels of polymorphism at simple sequence loci may decrease with increasing phylogenetic distance from the species used to construct the genomic library. However in some cases the primers were tested on highly inbred domestic stock and this may have biased the results. In contrast to this, the investigation of muntjac has found an increase in polymorphism. Along the same lines Gottelli *et al* (1994) and Roy *et al* (1994) both used primers derived from a domestic dog genomic library (Ostrander *et al* 1993) and found no reduction in polymorphism between the domestic dog and the canid species they each studied. Schlötterer *et al* (1991) discovered that microsatellite primers designed for the long-finned pilot whale (*Globicephala melas*) could be successfully applied to a diverse range of whales representing most of the major cetacean radiation.

All the primer pairs used have previously been reported to detect dinucleotide repeat microsatellite sequences (see Table 5·1). It may be expected that alleles would be found at every size step within the confines of the upper and lower allele size limits. This however is not the case here and it must remain a matter of conjecture whether the alleles not represented in this study exist but have not been detected, or in fact do not exist. In microsatellite DNA the mutational process can act in either direction to increase or decrease allele length. In human microsatellite studies Huang *et al* (1992) reported a new allele at locus DXS453 2bp larger than parental alleles and another new allele at locus DXS454 4bp shorter. In a separate study of 50 chromosomes evaluated with several markers, two new alleles were observed representing a gain of 2bp in one case and a loss of 4bp in the other (Kwaitkowski *et al* 1992).

With the exception of locus DRBP1, which was dimorphic, all the microsatellite loci were found to be highly polymorphic - having five to sixteen alleles. There was substantial variation in the total number of alleles detected per sub-population, ranging from 24 alleles at Waverly (n=5) to 43 alleles at Buntingford (n=18). Heterozygosity values, both observed and expected were high, ranging from 0·48 at Woburn (n=24) to 0·74 at Knebworth (n=28).

Heterozygosity, defined as the proportion of heterozygous individuals at a locus, has been widely used as a measure of genetic variation because it is proportional to the amount of genetic variance at a locus and lends itself readily to theoretical considerations of the effect of limited population size on genetic variation (Allendorf 1986). In general, heterozygosity is found to be a reliable measure of the loss of genetic variation caused by bottlenecks in population size and the level of heterozygosity provides a good measure of the capability of the population to respond to selection following a population crash or bottleneck event. However, as pointed out by Allendorf (1986), heterozygosity can be misleading in some instances in that it may be insensitive to the actual number of different genotypes at a locus. If the observed heterozygosity data is displayed graphically against the total number of alleles detected per sub-population (Figure 5.29) it can be seen that many comparisons can be made in which an increase in the number of alleles (and hence genetic diversity) is associated with a comparative decrease in heterozygosity.

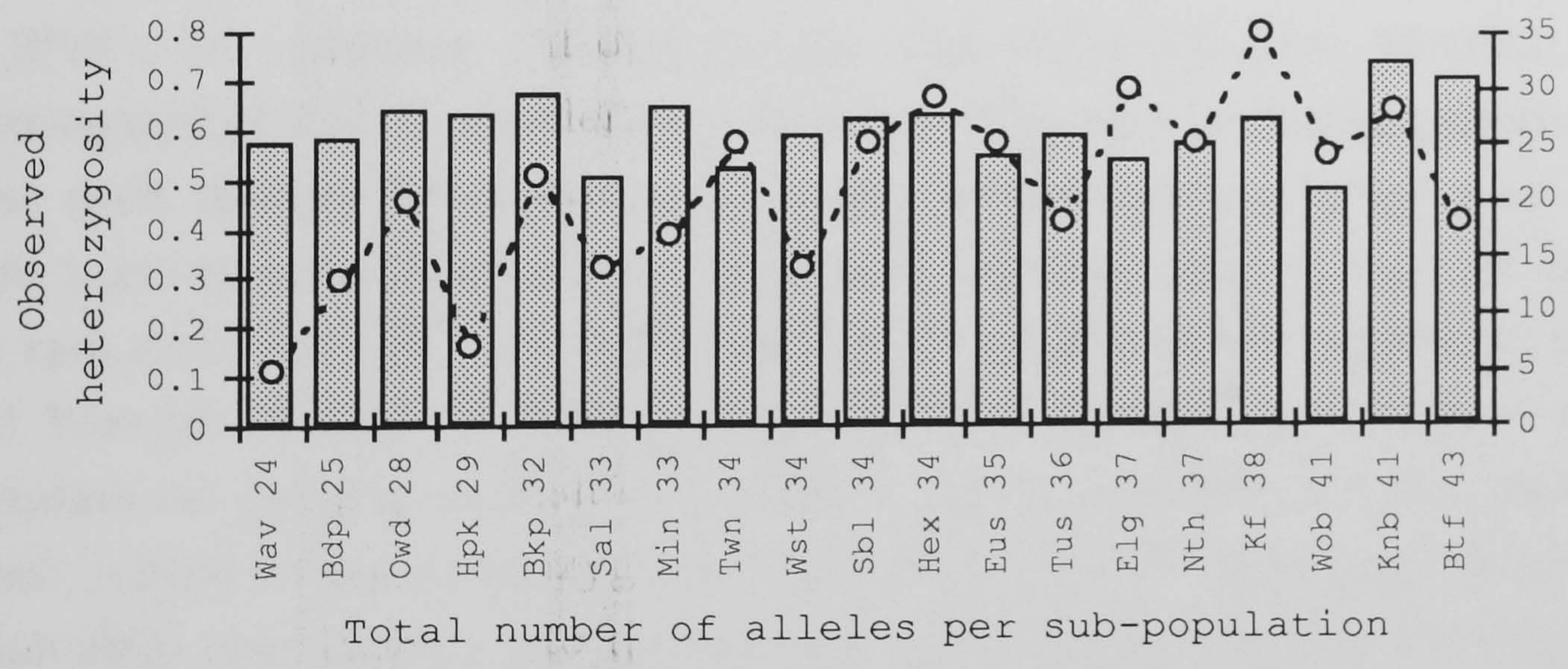


Figure 5·29 Comparison between total number of alleles per sub-population and observed heterozygosity. The overlay line graph represents the sample size.

For example, Minsbury Hill (Min) exhibits a greater level of heterozygosity (0·65) than most of the sub-populations to the right (on the graph) and yet has a lesser number of alleles.

Rare alleles can be used to determine sub-population relationships if sufficient quantity have been uncovered. In this study 14 rare alleles were found distributed between ten sub-populations (including Taiwan). From this some tentative relationships can be proposed.

There are four routes by which rare alleles can appear in a population, either as a result of mutation, migration, genetic drift or a diminution of a more common alleles by some selective process. In the populations considered in this study mutation is discounted as a possibility. Although the reported rate of mutation in microsatellite DNA is high, between $10^{-3}$ and $10^{-6}$ (Dietrich *et al* 1992, Edwards *et al* 1992, Kwaitkowski *et al* 1992, Weissenbach *et al* 1992), the time scale since the introduction of muntjac deer is very limited and the same mutations would have had to occur coincidentally in several geographically separate localities (unless there is some mechanism by which mutation arise more rapidly - as in the fragile X syndrome in humans (Fu *et al* 1991)). Diminution by selection is also discounted as microsatellite DNA, being non-coding, is likely to be selectively neutral. The majority of rare alleles detected in the various sub-populations are therefore considered to be a result of migration and/or genetic drift. Taking migration as the major course of the occurrence of rare alleles at the various sub-populations, then these rare alleles can be used to define genetic relationships between sub-populations. For example, from Table 5·4, a group formed by Woburn (Wob), Buntingford (Btf) and King' Forest (Kf) all share the same allele - BOVIRBP(169) as well as JP23 (150) Woburn /King's Forest, and JP23(152) plus MAF23(137) Buntingford/King's Forest. This group could be seen as being genetically close even though they are geographically distant. King's Forest may also form a group with Brocket Park (Bkp) and Northill (Nth) as they all share the rare allele MAF23(137). Interestingly no rare alleles were detected from Old Warden (Owd), Minsbury Hill (Min), or Hexton (Hex) three sub-populations geographically very close to the two groups. At Knebworth (Knb), which is equally geographically close to these two groups, only one (unshared) rare allele was detected. The grouping formed by Buntingford and King's Forest is particularly strong. Both these sub-populations share JP23(152) at a frequency of 8·3% and 8·6% respectively plus the rare alleles BOVIRBP(169) and MAF23(141). From an inspection of Table 5·4 other groupings could be defined.

The Woburn/Buntingford/King's Forest group, in which there are a total of four coincident rare alleles, is probably not the result of natural migration as these sub-populations are far apart. Woburn to King's Forest 75km, Buntingford to King's Forest 50km and Woburn to Buntingford 30km. As mentioned in chapter one, the migration rate is estimated to be about 1km/year and muntjac were absent from King's Forest until 1963. It was not until 1980 that they were recorded throughout the forest (Chapman *et al* 1994). There is no directional information in the distribution of rare alleles except that which can be inferred from the known times of colonisation. With this information the population at King's Forest can be seen to be at least partly descended from those populations with which it has coincident rare alleles, i.e. Buntingford, Woburn, Northill and Brocket Park. The muntjac population at King's Forest may therefore be the product of several releases rather than migration. The coincidence of the rare allele MAF23(141) between King's Forest and Euston may be a product of natural migration as these sub-populations are only about 10km apart. Any other attempts to determine directionality must remain speculative.

The two most interesting sub-populations are those of King's Forest and Buntingford. Six rare alleles were uncovered at King's Forest, this represents 15% of the total of 40 alleles found at this location, or 0·17 rare alleles per individual. At Buntingford four rare alleles were uncovered, which represents 9·5% of the total of 42 alleles, or 0·22 rare alleles per individual. This indicates that the population at Buntingford could be more diverse than that at King's Forest.

Three out of the five rare alleles from the Taiwan population are shared with four of the U.K. populations. These groups of muntjac are separated by at least 100 years if some of the U.K. founding population had ancestors in common with the Taiwan population. This observation supports the contention that (at least) some of these rare alleles have been extant in the population for some time and are not the product of recent mutational events.

Rare alleles may be the product of genetic drift, founder effects or the result of predation (selection) in the form of culling heavily in a particular area. In these cases alleles may become rapidly lost in small populations and what we may be seeing here are the remains of once common alleles, but in general the data here indicate that the number of rare alleles in a sub-population probably reflect the diversity of the origins of the local populations.

One explanation for the excess of homozygotes detected at MAF23, MAF65 and BOVIRBP is that one or more non-amplifying alleles segregate at these loci, such that heterozygotes between visible and non-amplifying alleles were recorded as homozygotes. Interestingly, these three loci are suspected of having non-amplifying alleles in red deer (J. Pemberton and D. Cheeseman, pers. comm.). Confirmation of this point requires samples from known parent-offspring sets and/or further molecular research.

Assuming that non-amplifying alleles are not the explanation for the excess of homozygotes observed here, the Hardy-Weinberg test has the ability to define how far the populations are from equilibrium and thus make the investigator aware that some process(es) may be occurring that are not immediately obvious. However, natural populations are almost never going to comply with all the assumptions necessary to maintain Hardy-Weinberg proportions. If a natural population does exhibit exact Hardy-Weinberg proportions then that may imply that all the assumptions have been met, or it may imply that a deviation in one direction is being balanced by a deviation in another.

In the investigation of Muntjac deer there are very few data about population size but it is certain to vary considerably from one sampling locality to another. King's Forest, for example, is a large area of undisturbed woodland ca 2360 ha of mainly coniferous plantations, whereas Bowden Park is a private estate. These different regions are likely to be able to support vastly different numbers of muntjac. So the assumption of very large population size (at this sub-population level) is violated. About migration again very little is known, but from the evidence accumulated in this study a number of deliberate releases and/or transfers have occurred. A third assumption which can not be maintained is that of the species under investigation having non-overlapping generations. Muntjac are well known to have no breeding season, the female coming into post-partum œstrus within eighteen hours to three days (Chapman 1991b) and having a gestation period of about 210 days. Male muntjac have been found to be fertile all year round (Chapman & Harris 1991).

Because of the statistical problems presented by having many genotypes represented in relatively small sample sizes the data were pooled and treated as one panmictic population Of the seven loci tested in this work only three were found not to deviate from Hardy-Weinberg expectations; these were DRBP1 (2 alleles), JP27 (5 alleles) and JP38 (5 alleles). Of the remaining four loci locus JP23 (7 alleles) was found to deviate significantly from Hardy-Weinberg expectations in the direction of heterozygous excess and the three other loci; MAF 23 (16 alleles), MAF 65 (7 alleles) and

BOVIRBP (15 alleles) deviated in the direction of homozygous excess (Table 5·5).

It has become well known that the existence of Hardy-Weinberg equilibrium can be demonstrated at many polymorphic loci in outbreeding organisms (Nei 1987). However, Hardy-Weinberg equilibrium may be disturbed by a number of factors, such as inbreeding, assortive mating and natural selection. The result of inbreeding will be to increase the frequency of homozygotes and, in the absence of other factors, the fixation index 'F' will be equal to Wright's (1969) inbreeding coefficient. If this were to be the case then F would always be positive and become equal to 1 when all genotypes eventually became homozygous. Assortive mating is another case of non-random selection. Here, mating occurs selectively among individuals having similar phenotypic characters. The effect is similar to inbreeding in that there is an increase in the frequency of homozygotes, but the effect is limited to only those loci concerned with the character with which assortive mating occurs. As microsatellites are neutral (unless deleterious) with respect to phenotypic characters, assortive mating can not have any measurable effect here - unless of course the regions of microsatellite DNA used in this investigation are linked to loci which would be affected by assortive mating.

Another consideration is the Wahlund effect. In a sub-divided population the genotype frequency may deviate from Hardy-Weinberg equilibrium in the entire population even if gene frequencies maintain Hardy-Weinberg equilibrium in each sub-population. It can be shown that if a population is divided into breeding units the frequency of homozygotes tends to be higher than the Hardy-Weinberg proportion (Nei 1987). This property was first noted by Wahlund (1928).

Assuming that assortive mating is irrelevant, the incidences of homozygosity may be taken as an indication of the level of inbreeding. A situation that would be expected given the small number of founders and the short time since their introduction. Overall the deviations from Hardy-Weinberg expectations found in the total population are due to both heterozygous excess and homozygous excess, indicating the homogenous nature of the sub-populations, taken in totality. The muntjac can therefore be seen as being in a dynamically expanding phase, increasing both their numbers and range. From the original 25 animals imported to the U.K. about 100 years ago the numbers have grown to an estimated 10,000 (S. Harris pers. comm.). This rate of expansion coupled with what appears to be an ad-hoc basis of releases would naturally tend to counter any approach to conformation to Hardy-Weinberg expectations.

The heterozygous excess uncovered at locus JP23 could be the result of selection in favour of heterozygotes, or due to the mixing of sub-populations previously quite divergent (see section 1.5). The latter explanation seems to be unlikely though as other loci would be expected to show the same effect. The former explanation is difficult to assess without more data on the locus and its genetic associations.

Overall, the values derived for Wright's $F_{ST}$ and Nei's $G_{ST}$ are reasonably similar and both indicate that a small degree of genetic differentiation exists among the sub-populations. Nei's $G_{ST}$ is probably the more accurate of these two measures in this work as the formula takes into account the existence of multiple alleles. Wright's $F_{ST}$ was derived for bi-allelic loci, and to accommodate multiple alleles, the minor alleles have to be pooled.
To determine if there was any partitioning among sub-populations three groups of sub-populations were erected based on their distance from Woburn, as previously described in Chapter 2 (Figure 2·1, page 23). From Table 5·7 and Figure 5·19, these groups show an increase in genetic differentiation that is strongly correlated with increasing distance from Woburn. The groups were analysed twice, once just the group and once the group plus Woburn. In the groups 'near' and 'medium' the inclusion of the Woburn samples caused a significant increase in the values of both $F_{ST}$ and $G_{ST}$, thus showing that the Woburn individuals are genetically distinct from these groups. A result that intuitively would not be expected as these groups are geographically near to Woburn. However the inclusion of Woburn in the 'far' group had the opposite effect. That is, the values of $F_{ST}$ and $G_{ST}$ decreased, demonstrating the greater genetic relatedness of these groups to Woburn. This is another counter-intuitive result. In terms of genetic similarity the 'far' group seem to have 'leap-frogged' the two other groups. The only explanation for this result is that the muntjac have been transported from place to place and released.
At a fine level of investigation pairwise values of $F_{ST}$ and $G_{ST}$, were calculated for all possible pairs of sub-populations. Figure 5·14 demonstrates that a proportionality is maintained between $F_{ST}$ and $G_{ST}$, for most sub-populations.

But if the pairwise data are interpreted in the form of UPGMA dendrograms the results are different in a number of respects. Taking $F_{ST}$ first, the sub-populations are seen to fall into three main groups (Figure 5·17) which are depicted in a geographical context in Figure 5·18. There is some evidence of geographical organisation to the two larger clusters. The cluster beginning with Btf and Min form a group to the west of Woburn whilst the group beginning with Tus and Sal form a group to the east (with the exception of the inclusion of Euston). From this data it would seem that radiation from Woburn has proceeded in two separate directions. In a north/south direction to the west and in a north/south direction to the east.

From an interpretation of the data using Nei's $G_{ST}$ the sub-populations can be seen to fall into two clusters (Figure 5·15), which are depicted in a geographical context in Figure 5·16. In contrast to the $F_{ST}$ interpretation there is no geographical organisation. For reasons previously given, the analysis using $G_{ST}$ probably gives a more accurate interpretation of the similarities between sub-populations. There is no correlation between distance from Woburn and genetic similarity, in fact from Figure 5·16 the lines connecting sites of genetic similarity are seen to criss-cross in a random fashion.

However there are some agreements in these two interpretations in that the grouping Northill / Oldwarden / Minsbury Hill / Brocket Park / Knebworth / Hexton / Buntingford (plus or minus some additions) can be seen in both, and Tusmore-Southblock are also consistently together. But there are also many differences. For example Euston is seen as being genetically similar to Harry's Park and quite distant from Woburn and Knebworth ($F_{ST}$), or, from the $G_{ST}$ interpretation Euston is genetically similar to Waverly and Knebworth and distant from Harry's Park. A consistently larger genetic dissimilarity than would be expected is that between King's Forest and Euston. These two localities are only about 12km apart and yet are found to be genetically more closely related to geographically distant localities than they are to each other.

The value of applying Wright's $F_{ST}$ and Nei's $G_{ST}$ to this data has been found to be somewhat contentious in comparison. Nei's $G_{ST}$ is taken as being most relevant, but for both $F_{ST}$ and $G_{ST}$ the values are small. The sub-populations that form two of the three major clusters as defined by $F_{ST}$ are separated on the basis of $F_{ST}$ values that fall below 0·06. The sub-populations that form the two major clusters as defined by $G_{ST}$ are separated on values of $G_{ST}$ that are equal to or less than 0·135. This range of values may be too low to give a clear picture of sub-population relationships.

In the autocorrelation calculations in which simple nearest neighbours were considered there was no significant deviation from zero until distance classes four to six. These deviations were not uniquely positive or negative but an increase in the oscillations about zero. The same situation was seen in the 'expanding radii' from Woburn. When a cline of ranked microsatellite allele frequency is arranged a positive to negative correlation may be expected. From Figures 5·25 and 5·27 the only correlation present is that of the alleles that were used to arranged the cline. The other alleles behave as before. A degree of autocorrelation evident between various localities would be evidence that Wright's (1969) isolation-by-distance model could be applied to the population of muntjac covering the sample area.

Together these different approaches to investigating spatial autocorrelation between sub-populations provides powerful evidence of the random nature of the dispersal of muntjac. The lack of correlation discovered here is strong evidence that a natural migration and radiation from Woburn (the putative centre of origin) has indeed not taken place.

The interpretation of the microsatellite allele frequencies using an estimate of the migration rate (N*m*) shows that for four sub-populations (Minsbury Hill, Buntingford, Hexton and Brocket Park) differentiation by distance may be significant. This result is surprising given that these four sub-populations are grouped within a circle of 20km diameter, which also includes Knebworth, the $r^2$ value of which approaches the levels of the other four sub-populations.

In conclusion: the analysis of microsatellite DNA using only seven loci has provided a wealth of data from which genetic similarity between sub-populations can be ascertained. The primers used were from deer, sheep and cattle yet, with the exception of locus DRBP1 which was bi-allelic, all loci were highly polymorphic, showing between five and sixteen alleles. The sub-populations themselves were found to carry a high level of genetic diversity as shown by the high values of heterozygosity. A substantial number of rare alleles were discovered. On the basis of alleles shared between sub-populations some associations between various sub-populations were suggested but these were not borne out by any of the subsequent analyses. The application of Wright's (1951) $F_{ST}$ and Nei's (1965) $G_{ST}$ were contentious in that each produced substantially different interpretations of the data. On the grounds that Nei's $G_{ST}$ takes multiple alleles into consideration, this was the interpretation thought most likely to be more accurate.

Chapter 6

# GENERAL DISCUSSION

The purpose of this study was to investigate the level of molecular genetic variation in sub-populations of muntjac deer in order to establish the origins and mode of dispersal in present day sub-populations. Three molecular genetic approaches were used, those of DNA fingerprinting, restriction fragment analysis of a region of mitochondrial DNA, and microsatellite DNA analysis. Genetic information of this kind can be used to identify relationships both within and between sub-populations as well as elucidate the spatial/migratorial relationships.

The combined use of three molecular genetic techniques in this work has provided strong evidence that muntjac in the U.K. have not dispersed in a regular way to occupy their present-day range. Previous to this work, the sightings of muntjac over large distances from the supposed centre of origin (Woburn Abbey) were taken to indicate rapid dispersal and colonisation and hence the consensus opinion arose that muntjac are rapid colonisers (e.g. Pickvance & Chard 1960, Anderson & Cham 1987). The evidence gathered in this work suggests that dispersal has been human mediated, with intentional translocations.

In a naturally dispersing population an observed differentiation between sub-populations would be expected to reflect the dispersal distance of individuals over time. If differentiation among sub-populations is a function of the dispersal of individuals, then the observed variation in allele frequencies between sub-populations should increase with increasing geographic distance and genetic differences within sub-populations would be expected to diminish. Not all possible sub-populations within the study area have been sampled therefore some fine gradations in genetic similarity between sub-populations may have been missed.

The assessment of similarity within and between sub-populations by DNA fingerprinting, even though a limited number of sample sites were investigated, has proved informative and the potential for a productive wider study been demonstrated. The sub-population at Knebworth was found to have the most diverse grouping of individuals (in terms of departure from

inbreeding) and the sub-population at Brocket Park the most closely related. King's Forest and Brocket Park were found to be more closely related to each other than to any of the other sub-populations, yet these two localities are geographically the furthest apart. In the analysis of mtDNA the four sub-populations analysed by DNA fingerprinting were found to be consistently genetically far apart. Had a complete data set been available then the actual levels of similarity of these four sub-populations would have been more meaningful. All four sub-populations carried rare microsatellite alleles (see Table 5·4), King's Forest sharing two with Brocket Park and three with Buntingford, thus reinforcing the relationships previously noted.
The restriction fragment analysis of the D-loop of mtDNA uncovered eight maternal lines from a theoretically possible 18 (theoretically possible within the confines of the restriction profiles produced here). The 'missing' genotypes may never have existed although it is quite possible that some of the 'missing' maternal lines may have died out, especially as so many animals perished in hard winters. In 1947, 70 muntjac corpses were found in Hazelborough Forest, from the Southblock sample site in Northamptonshire (Pickvance & Chard 1960) and in the winter of 1962/63 a total of 83 corpses were reported from various localities, which is supposed to be a minimum number (Chapman *et al* 1994). Of great value to the mtDNA analysis in this work would be the occurrence of rare genotypes. Unfortunately only three rare genotypes were found and only one of these was shared; between Knebworth and King's Forest (Table 4·3). The occurrence of this rare genotype indicates a strong connection between these two localities and, as they are about 75km apart and both the intervening and surrounding sub-populations do not display this genotype, the suggestion of human mediated dispersal appears to be valid.
All the analysis methods used here have associated sub-populations on the basis of the frequency of occurrence of the various alleles and genotypes. Whilst this would undoubtedly be highly informative in a naturally dispersing population, in this case it is less useful as there has been a certain degree of human intervention. From the low level of diversity (eight genotypes) many sub-populations are likely to be defined as being genetically close whereas they are only related matriliniarly and may be in fact many generations apart. The value of mtDNA analysis has been found to lie in its ability to indicate a gross structure to dispersal and to define historical

relationships between sub-populations - as depicted by Figure 4·14 (page 113). From this figure it can be seen that the inter-relatedness of the sub-populations is more complex than was initially expected from the assumption given in the introduction to this thesis.

For example, in all mtDNA analyses the four sub-populations Northill-Old Warden-Minsbury Hill-Bowden Park are constantly seen as being genetically very close, if not contiguous. If fact they are genetically very close but only by virtue of sharing the same genotype - which is also the most common genotype. An inspection of Table 4·3 will show that the founder(s) of these sub-populations could in fact have originated from any of the other sub-populations (with the exception of King's Forest, in which this genotype does not occur). From this grouping Northill-Old Warden-Minsbury Hill could well be the product of dispersal as they are all geographically within 10km of each other, but it is highly unlikely that the sub-population at Bowden Park is a product of natural dispersal given that this sample site is about 120km to the South West of the group.

Similar observations could be made for other sub-populations and groups of sub-populations. This general inability of the mtDNA analysis to proceed any further than identify at best tentative relationships between sub-population is a serious weakness, but one that is probably a consequence of human mediated dispersal.

In muntjac, microsatellites were found to be moderately to highly polymorphic, having two to sixteen alleles per locus and heterozygosity values that averaged 0·6. Such high heterozygosity may be expected from microsatellite loci, given their estimated mutation rate of $10^{-4}$ (Dallas 1992, Dietrich *et al* 1992) and large population size. At equilibrium, heterozygosity (H) = $1-[1/(1+8N_e\mu)^{0.5}]$, where $N_e$ = the effective population size and $\mu$ = the single step mutation rate (Ohta & Kimura 1973). From this equation an effective population size in the region of 6,500 individuals (within the confines of this study area) would be sufficient to maintain heterozygosity values of 0·6.

The microsatellite data collected here show that differentiation by distance is not significant across the geographic range of the sample area, indicating that the muntjac have not dispersed in a 'natural' way. The values calculated for $G_{ST}$ are low, with an average of 0·073 ± 0·033 (not including Taiwan) and a median of 0·07, indicating that there is effectively no genetic partitioning between sub-populations.

A recent paper by Chapman *et al* (1994) has documented the spread of Reeve's muntjac in the U.K. The authors of this paper have been able to make use of archive material and distribution records not previously available, as well as establishing personal communications with retired park workers from Woburn.

Chapman *et al* report that Reeve's muntjac were first imported to the U.K. at the Zoological Society of London in 1838, and by 1845 there was also a collection of muntjac at Knowsley Hall (Merseyside), the origins of which are unknown. Reeve's muntjac were thereafter imported to London Zoo in 1867 (from Taiwan), 1873 (Ningpo, south-east China) and 1874 (Taiwan, Ningpo and Hong Kong). In the early days of the presence of muntjac in the U.K. *'all that is certain is that during the latter half of the nineteenth century there was a trade in Reeve's muntjac to and from Britain...'* and *'... it seems probable that at least some of the animals supplied to Woburn were descended of stock from London Zoo.'* (Chapman *et al* 1994).

The paucity of early records confounds an understanding of the spread of Reeve's muntjac, clearly they are not all descended from the Woburn stock, as was originally postulated. What proportion the stock at Woburn originated from animals imported from their native range of China and Taiwan, or animals obtained from other collections in Britain and Europe is unknown. In the late 1930s and 1940s a number of well established colonies appeared - often at some distance from Woburn, and with intervening region apparently not colonised. This relatively sudden appearance of a number of breeding populations coupled with scattered records of occurrences at some distance from Woburn led to the consensus opinion that Reeve's muntjac were rapid colonisers.

However, during the late 1930s and the late 1940s many animals were purposefully translocated from Woburn to several surrounding regions. Other nearby collections at Tring (20km south of Woburn) and Whipsnade Park (38km north) are also probably sources of feral muntjac. A chronological history of escapes and deliberate releases may be found in Table six of Chapman *et al* (1994). From this table Figure 6·1 has been constructed to show the relationship between the sub-populations sampled in this study and the known release sites.

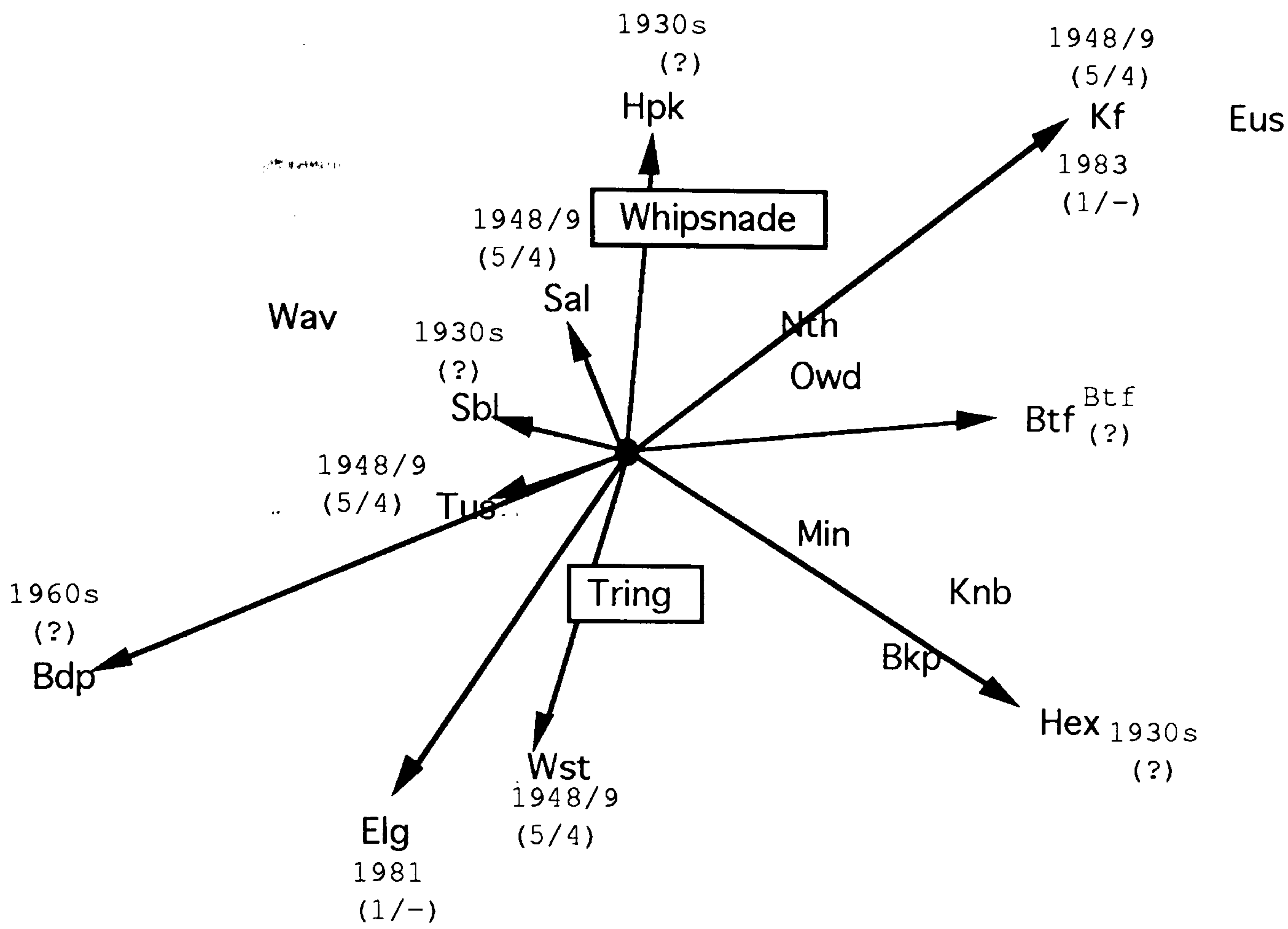


Figure 6·1 Known release sites with the approximate dates of release and the ratio of males to females released (♂/♀). Tring and Whipsnade also maintain populations of muntjac deer.

Many of the relationships between Woburn Abbey and the release sites are reflected in the mtDNA data analysis of Chapter 4. All the release sites noted in Figure 6·1 share at least one mtDNA genotype with Woburn and three of the sites share two genotypes. However, with the exception of Bowden Park (from which only one genotype was uncovered) all the sites have one, two or three mtDNA genotypes not found at Woburn. This may be a sampling deficiency - as previously noted.

Throughout the data analysis one group of sub-populations has consistently been grouped together. This group is formed from Bowden Park, Northill, Oldwarden and Minsbury Hill. All share one or two of the mtDNA genotypes found at Woburn (Table 4·3) yet only Bowden Park is known to have had deer translocated from Woburn. The origins of the other three sites are unknown but, as they are all within 23km of Woburn it is likely that they are the product of dispersal following escapes or local (to Woburn) releases.

Another consistent group is of Harry's Park and Wormsley Estate. These two sites are grouped both in the mtDNA analysis and in the microsatellite analysis, and both sample sites received muntjac from Woburn (Hpk in the 1930s and Wst in 1948/9). This may again be an artefact caused by inappropriate sampling as only seven individuals were taken from Hpk and 14 from Wst.

Out of the (theoretically) possible 18 mtDNA genotypes only eight were uncovered in the 369 individuals investigated during the course of this investigation. Some maternal lines may have died out, or be so rare as to evade detection. Fifteen females are known to have been introduced to Woburn (Chapman *et al* 1994) and assuming that they were all unrelated and contributed to the feral population, these introductions would be sufficient to account for all the mtDNA genotypes observed. However, several considerations suggest that this may not be the case. A high proportion of the muntjac introduced into the park at Woburn died within their first year. Between 1894 and 1903 a total of 13 males and 14 females were introduced to Woburn, of which 10 males and 11 females died (Chapman *et al* 1994 [Table 3]). Only two females were recorded as being released from the enclosures in the park (Chapman *et al* 1994). It seems likely that either there were more introductions to Woburn than the records suggest, or that the U.K. Reeves' muntjac population is descended from other sources.

Previously, assumptions about the rapid rate of dispersal of Muntjac were based only on sporadic sightings. As the locations of these sightings were sometimes far apart the consensus opinion developed that the Muntjac had in fact migrated from one sighting location to another and thus were rapid colonisers (Chapman *et al* 1994). There is very little evidence for these assumptions as the radio-tracking work by Harris *et al* (1990) and Chapman *et al* (1993) shows. Harding (1986) estimate their mean rate of spread to be 2·5 km per year but Chapman *et al* (1994) have estimated the rate to be in the order of 1 km per year, which is comparable to Harding's (1986) figure of 0·8 km per year for Roe Deer (*Capreolus capreolus*) and Chapman and Chapman's (1980) figure of 0·8 ± 0·15 km per year for Fallow Deer (*Dama dama*). Although it must be born in mind that this radio-tracking data was only taken in Kings'

Forest - the animals may in fact disperse more rapidly in more open country or where they are subjected to human disturbance.

In this study, the three data sets are combined to provide a comprehensive picture of the genetic structure of muntjac deer populations and the process of colonisation in southern England. The possibilities and advantages of using DNA fingerprinting at the level of populations are demonstrated and discussed. Two advantages of the mtDNA approach are illustrated. Firstly, the phylogenetic information from the data gave some insight into the origins of the various sub-populations. And secondly, the value of mtDNA analysis in association with microsatellite DNA analysis was used to give an indication of the timing of founding events.

The mtDNA data also suggested that a number of founding events had occurred within the sub-populations, especially those sub-populations that are remote from others. However, the genetic differences found among some of the sub-populations may be due to the extent of colonisation and survival rates of the founding populations in each.

A high level of genetic similarity taken in isolation from other data does not differentiate between a dispersal event that led to the founding of a new colony or a migration from one colony to a second pre-existing colony. Long term field observations and radiotelemetry will be needed or regular genetic sampling of the sub-populations. This could be accomplished non-invasively by sampling hair and amplifying the DNA using the polymerase chain reaction, a technique which also opens up the potential for analysing ancient museum samples or poorly stored samples that could not otherwise be analysed.

However, considering the dangers of assuming genetic relationships from a single data set it is suggested that other molecular genetic approaches are considered for concurrent use.

In summary, the analysis of restriction fragments from the D-loop of mtDNA has provided a 'snap-shot' of the present distribution of matrilinear lines with the implicit inclusion of a historical perspective. It can be show that certain sub-populations are genetically close, and thus one is probably derived from the other. No temporal directionality can be deduced from these data.

The microsatellite analysis can again be shown to interpret the genetic relationships at the time the samples were taken, again with no temporal content. However, when the data from the two techniques are combined, in a general sense, a second dimension to the interpretation becomes available, that of time. Sub-populations that can be shown to belong to the same maternal line, and therefore initially closely related sub-populations can be differentiated in the general way of; sub-population 'A' has been isolated from sub-population 'B' long enough for to allow the effects of genetic drift to become apparent.

## LITERATURE CITED IN THE TEXT


ALLARD, M.W. & Honeycutt, R.L. (1992). Nucleotide sequence variation in the mitochondrial 12S rRNA gene and the phylogeny of African mole-rats (Rodentia: Bathyergidae). *Molecular Biology and Evolution*, **9**(1): 27-40.

ALLENDORF. F.W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology* **5**: 181-190.

ALLENDORF, F.W., Christiansen, F.B., Bodson, T., Eanes, W.F. & Frydenberg, O. (1979). Electrophoretic variation in large mammals. I The polar bear *Thalarctos maritimus*. *Hereditas* **91**: 19-22.

AMOS, D. (1992). Muntjac Deer in Leicestershire. *Leicestershire Museum Publication No 119*, Leicester.

AMOS, W. (1990). Analysis of Polygamous Systems Using DNA Fingerprinting. *Symposium of the Zoological Society of London*. **64**: 151-165.

AMOS, W., Barrett, J. & Dover, G. (1991). Breeding system and social structure in the Faroese pilot whale as revealed by DNA fingerprinting. *Report of the International Whaling Commission, Special Issue* **12**: 79-85.

AMOS, W., Schlötterer, C & Tautz, D. (1993a). Social structure of pilot whales revealed by analytical DNA profiling. *Science* **260**: 670-672.

AMOS. W,. Twiss, S., Pomeroy, P.P. & Anderson, S. S. (1993b). Male mating success in the grey seal *Halichoerus grypus*: a study using DNA fingerprinting. *Proceeding of the Royal Society of London*. **252**: 199-207.

ANDERSON, D. & Cham, S.A. (1987). Muntjac deer (*Muntiacus reevesi*) - the early years. *Bedfordshire Naturalist*, **42**: 14-18.

ANDERSON, R.P. & Roth, J. (1977). Tandem genetic duplications in phage and bacteria. *Annual Review of Microbiology* **31**: 4730505.

ANDERSON, R.P. & Roth, J. (1981). Spontaneous tandem genetic duplications in Salmonella typhimurium arise by unequal recombination between rRNA (rrn) cistrons. *Proceedings of the National Academy of Science, USA* **78**: 3113-3117.

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

ANDERSON, S., deBruijn, M.H.L., Coulson, A.R., Eperon, I.C. & Sanger, F. (1982). Complete sequence of bovine mitochondrial DNA. *Journal of Molecular Biology*, **156**: 683-717.

AQUADRO, C.F. & Greenberg, B.D. (1982). Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics*, **103**: 287-312.

AQUADRO, C.F., Kaplan, N. & Risko, K. J. (1984). An analysis of the dynamics of mammalian mitochondrial DNA sequence evolution. *Molecular Biology and Evolution*, **1**(5): 423-434.

ARCHIE, J.W. (1985). Statistical analysis of heterozygosity data: independent sample comparisons. *Evolution* **39**: 623-637.

ARNASON, U. & Johnsson, E. (1992). The complete mitochondrial DNA sequence of the Harbour seal, *Phoca vitulina*. *Journal of Molecular Evolution*, **34**: 493-505.

ARNASON, U., Gullberg, A. & Widegren, B. (1991). The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. *Journal of Molecular Evolution*, **33**(6): 556-568.

ATTARDI, G. (1985). Animal mitochondrial DNA: An extreme example of genetic economy. *International Review of Cytology*, **93**: 93-145.

AVISE, J.C. (1986). Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society London*, **312**(B): 325-342.

AVISE, J.C. (1989). Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution*, **43**(6): 1192-1208.

AVISE, J.C. (1994). *Molecular Markers, Natural History and Evolution*. Chapman and Hall, London.

AVISE, J.C. & Lansman, R.A. (1983). Polymorphism of mitochondrial DNA in populations of higher animals. In: M. Nei & R.E. Kohen (eds.) *Evolution of Genes and Proteins.* Sinauer, Sunderland, MA.

AVISE, J.C., Lansman, R.A. & O'Shade, R. (1979). The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations I. Population structure and evolution in the genus *Peromyscus. Genetics,* **92**: 279-295.

AVISE, J.C., Reeb, C.A. & Saunders, N.C. (1987). Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). *Evolution,* **41**(5): 991-1002.

BAIRD, R.W. & Stacey, P.J. (1988). Variation in saddle patch pigmentation in populations of killer whales (*Orcinus orca*) from British Columbia, Alaska and Washington State. *Canadian Journal of Zoology* **66**: 2582-2585.

BAKER, C.S., MacCarthy, M., Smith, P.J., Perry, A.P. & Chambers, G.K. (1992). DNA fingerprints of orange roughy, *Haplostethus atlanticus*: a population comparison. *Marine Biology* **113**: 561-567.

BAKER, C.S., Palumbi, S.R., Lambertsen, R.H., Weinrich, M.T., Calambokidas,J. & O'Brien, S.J. (1990). Influence of seasonal migration on geographic distribution of mitochondrial haplotypes in humpback whales. *Nature* **343**: 238-240.

BAKER, M.C. & Fox, S.F. (1978). Dominance, survival and enzyme polymorphism in dark-eyed juncos, *Junco hyemalis. Evolution* **32**: 697-711.

BAKER, M.C. & Marler, P (1980). Behavioural adaptations that constrain the gene pool in vertebrates. In: Markl, H. (ed.) *Evolution of Social Behaviour: Hypotheses and Empirical Tests.* Konferenzen 1980, Weinheim, Verlag Chemie GmbH.

BALLINGER, S.W., Blankenship, L.H., Bickham, J.W. & Carr, S.M. (1992). Allozyme and mitochondrial DNA analysis of a hybrid zone between white-tailed deer and mule deer (*Odocoileus*) in West Texas. *Biochemical Genetics,* **30**(1/2): 1-11.

BALLOU, J. (1991). Small population overview. In: *Pink Pigeon. Conservation Viability Assessment.* Species Survival Commission/IUCN: 2-11.

BALLOU, J. & Ralls, K. (1982). Inbreeding and juvenile mortality in small populations of ungulates: a detailed analysis. *Biology and Conservation* **24**: 239-272.

BARONE, M.A., Roelke, M.E., Howard, J., Brown, J.l., Anderson, A.E. & Wildt, D.E. (1993). Reproductive fitness of the male Florida panther: Comparative studies of *Felis concolor* from Florida, Texas, Colorado, Chile and North American Zoos. *Journal of Mammology* **75**(1): 150-162

BARRETTE, C. (1987). The comparative behaviour and ecology of chevrotains, musk deer and morphologically conservative deer. In: Wemmer, C.M. (ed.) *Biology and Management of the Cervidae.* Smithsonian Institution Press, Washington D.C. pp 200-213.

BARTLETT, S.E. & Davidson, W.S. (1991). Identification of *Thunnus* tuna species by the polymerase chain reaction and direct sequence analysis of their mitochondrial cytochrome b genes. *Canadian Journal of Fisheries and Aquatic Sciences,* **48**: 309-317.

BECKMANN, J.S. & Weber J.L (1992). Survey of human and rat microsatellites. *Genomics* **12**: 627-631.

BECKWITT, R. & Petruska, J. (1985). Variation in mitochondrial DNA genome size among fishes of the family Scorpaenidae. *Copeia,* **4**: 1056-1058.

BELL, G.I., Selby, M.J. & Rutter, W.J. (1982). The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* **295**: 31-35

BENGTSSON, B.O. (1978). Avoid inbreeding: at what cost? *Journal of Theoretical Biology* **73**: 439-444.

BENTZEN, P., Leggett, W.C. & Brown, G.C. (1988). Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (*Alosa sapidissima*). *Genetics,* **118**: 509-518.

BERG, W.J. & Ferris, S.D. (1984). Restriction endonuclease analysis of salmonid mitochondrial DNA. *Canadian Journal of Fish and Aquatic Science* **41**(7): 1041-1047.

BERMINGHAM, E. & Avise, J.C. (1986)Molecular zoogeography of freshwater fishes in the Southeastern United States. *Genetics,* **113**: 939-965.

BERNATCHEZ, L., Guyomard, R. & Bonhomme, F. (1992). DNA sequence variation in the mitochondrial control region among geographically and morphologically remote European brown trout (*Salmo trutta* L.) populations. *Molecular Ecology* **1**: 161-173

BERZIN A.A. & Vladimirov, V.L. (1983). A new species of killer whale from Antarctic waters. *Zoologicheskií Zhurnal* **62**(2): 287-295. (In Russian with abstract in English).
BIBB, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. & Clayton, D.A. (1981). Sequence and gene organisation of the mouse mitochondrial DNA. *Cell*, **26**: 167-180.
BIGG, M.A., Olesiuk, P.F., Ellis, G.M., Ford, J.K.B. & Balcomb, K.C. (1990). Social organization and genealogy of resident killer whales (*Orcinus orca*) in coastal waters in British Columbia and Washington State. *International Whaling Commission Special Report Series* **12**: 383-406.
BIJU-DUVAL, C., Ennafaa, H., Dennebouy, N., Monnerot, M., Mignotte, F., Soriguer, R.C., El Gaaied, A., El Hili, A. & Mounolou, J.C. (1991). Mitochondrial DNA evolution in lagomorphs: origin of systematic heteroplasmy and organisation of diversity in European rabbits. *Journal of Molecular Evolution*, **33**: 92-102.
BIRD, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research* **8**: 1499-1504.
BIRD, A.P. & Taggart, M.H. (1980). Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Research* **8**: 1485-1497.
BITTLES, A.H. (1979). Incest re-assessed. *Nature* **208**: 107.
BONNELL, M.L. & Selander, R.K. (1974). Elephant seals: genetic variation and near extinction. *Science* **184**: 908-909.
BONNE-TAMIR, B., Johnson, M.J., Natali, A., Wallace, D.C. & Cavalli-Sforza, L.L. (1986). Human mitochondrial DNA types in two Israeli populations - a comparative study at the DNA level. *American Journal of Human Genetics*, **38**: 341-351.
BOWCOCK., A., Osbourne-Lawrence, S., Barnes, R., Chakravarti, A., Washington, S. & Dunn, C. (1993). Microsatellite polymorphism linkage map of human chromosome 13. *Genomics* **15**: 376-386.
BRAATEN, D.C., Thomas, J.R., Little, R.D., Dickson, K.R., Goldberg, I., Schlessinger, Ciccodicola, A. & D'Urso, M (1988). Locations and contexts of sequences that hybridise to poly(dG-dT)•(dC-dA) in mammalian ribosomal DNAs and two X-linked genes. *Nucleic Acids Research* **16**: 865-881.
BRAY, D.W. (1981). Taking Muntjac Seriously. *Deer*. 184-185.
BRITTEN, R.J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science*, **231**: 1393-1398.
BROCK, M.K. & White, B.N. (1991). Multifragment alleles in DNA fingerprints of the parrot, *Amazona ventralis*. *Journal of Heredity* **82**: 209-212
BROCK, M.K. & White, B.N. (1992). Application of DNA fingerprinting to the recovery program of the endangered Puerto Rican Parrot. *Proceedings of the National Academy of Science USA*. **89**: 11121-11125.
BROOK, J.D., McCurrach, M., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stantion, V.P., Thirion, J.P. Hudson T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J. & Houseman, D.E. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**: 799-808.
BROWN, A.H.D. (1979). Enzyme Polymorphisms in Plant Populations. *Theoretical Population Biology*. **15**: 1-42.
BROWN, T.A. (1990). *Gene Cloning*. Chapman and Hall, London.
BROWN, W.M. (1981). Mechanisms of evolution in animal mitochondrial DNA. *Annals of the New York Academy of Sciences*, **361**: 119-134.
BROWN, W.M. (1983). Evolution of animal mitochondrial DNA. In: M. Nei & R.E. Kohen (eds.) *Evolution of Genes and Proteins*. Sinauer, Sunderland, MA.
BROWN, W.M. (1985). The mitochondrial genome of animals. In: MacIntyre, J.R. (ed.). *Molecular Evolutionary Genetics*. New York: Plenum, 95-131.
BROWN, W.M. George, M. & Wilson, A.C. (1979). Rapid evolution of mitochondrial DNA. *Proceedings of the National Academy of Science, USA* **76**: 1967-1971.
BROWN, W.M., Gadelata, G., Pepe, G., Saccone, C & Sbisá, E (1986). Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *Journal of Molecular Biology*, **192**: 503-511.

BROWN, W.M., Prager, E.M., Wang, A. & Wilson, A.C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *Journal of Molecular Evolution*, **18**: 225-239.

BROWN, W.M. & Vinograd, J. (1974). Restriction endonuclease cleavage maps of animal mitochondrial DNAs. *Proceedings of the National Academy of Sciences USA*, **71**(11): 4617-4621.

BROWN, W.M., & Wright, J.W. (1979). Mitochondrial DNA analysis and the origin and relative age of parthenogenetic lizards. (genus *Cnemidophorum*). *Science* **203**: 1247-1249.

BRUFORD, M.W. & Altmann, J. (1993). DNA fingerprinting and problems of paternity determination in an inbred captive population of Guinea baboons (*Papio hamadryas papio*). *Primates* **34**(3): 403-411.

BRUFORD, M.W., Hanotte, O., Brookfield, J.F.Y. & Burke, T. (1992). Single-locus and multilocus DNA fingerprinting. In: A.R. Hoelzel (ed.) *Molecular Genetic Analysis of Populations: A Practical Approach.* IRL Press, Oxford. pp 225-269.

BRUFORD, M.W. & Wayne R.K. (1993). Microsatellites and their application to population genetics. *Current Opinions in Genetics and Development* **3**: 939-943.

BUCHANAN, F.C., Swarbrick, P.A. & Crawford, A.M. (1991). Ovine dinucleotide polymorphism at the MAF48 locus. *Animal Genetics* **22**: 379-380.

BUCHANAN, F.C., Swarbrick, P.A. & Crawford, A.M. (1992). Ovine dinucleotide polymorphism at the MAF65 locus. *Animal Genetics* **23**: 85

BUCHANAN, F.C. & Crawford, A.M. ( 1992a). Ovine dinucleotide repeat polymorphism at the MAF209 locus. *Animal Genetics* **23**: 183

BUCHANAN, F.C. & Crawford, A.M. ( 1992b). Ovine dinucleotide repeat polymorphism at the MAF70 locus. *Animal Genetics* **23**: 185

BUCHANAN, F.C. & Crawford, A.M. ( 1992c). Ovine dinucleotide repeat polymorphism at the MAF33 locus. *Animal Genetics* **23**: 186

BUCHANAN, F.C. & Crawford, A.M. ( 1992d). Ovine dinucleotide repeat polymorphism at the MAF214 locus. *Animal Genetics* **23**: 394.

BUITKAMP, J., Ammer, H. & Geldermann, H. (1991a). DNA fingerprinting in domestic animals. *Electrophoresis* **12**: 169-174.

BUITKAMP, J., Zischler, H., Epplen, J.T. & Geldermann, H. (1991b). DNA fingerprinting in cattle using oligonucleotide probes. *Animal Genetics* **22**: 137-146.

BURGER, G. (1985). Unassigned reading frames in the mitochondrial DNA of *Neurospora crassa*. In : E. Quagliariello et. al. (eds.) *Achievements and Perspectives of Mitochondrial Research* Volume II: Biogenesis: Elsevier Science Publishers B.V. (Biomedical Division) pp 305-316.

BURKE, T. (1989). DNA fingerprinting and other methods for the study of mating success. *Trends in Ecology and Evolution.* **4**:139-144.

BURKE, T. & Bruford, M.W. (1987). DNA fingerprinting in birds. *Nature* **327**: 139-144.

BURKE, T., Davies, N.B., Bruford, M.W. & Hatchwell, B.J. (1989). Parental care and mating behaviour of polyandrous dunnocks *Prunella modularis* related to paternity by DNA fingerprinting. *Nature* **338**: 249-251.

BUROKER, N.E., Brown, J.R., Gilbert, T.A., O'Hara, P.J., Beckenbach, A.T., Thomas, W.K. & Smith, M.J. (1990). Length heteroplasmy of sturgeon mitochondrial DNA: An illegitimate elongation model. *Genetics*, **124**: 157-163.

CALLEN, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C. & Sutherland, G.R. (1993). Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*. **52**: 922-927.

CAMERON, D.G. & Vyse, E.R. (1978). Heterozygosity in Yellowstone Park elk *Cervus canadensis*. *Biochemical Genetics* **16**(7/8): 651-657.

CAMPBELL, R.C. (1992). *Statistics For Biologists*. Cambridge University Press. Cambridge.

CANN, R.L., Brown, W.M. & Wilson, A.C. (1984). Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* **106**: 479-499.

CANN, R.L., Stoneking, M. & Wilson, A.C. (1987). Mitochondrial DNA and human evolution. *Nature*, **325**: 31-36.

CAPON, D.J., Chen, E.Y., Levinson, A.D., Seeburg, P.H., & Goeddel, D.V. (1983). Complete nucleotide sequence of the T24 human bladder carcinoma oncogene and its normal homologue. *Nature* **302**: 33-37.

CARNE, P.H. (1968). The spread of the muntjac. *Country Life* **143**(3696): 24-25.

CARR, S.M. & Marshall, H.D. (1991). Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome b gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**: 48-52.

CARSON, H.L. (1968). The population flush and its genetic consequences. In: Lewontin, R.C. (ed.), *Population Biology and Evolution*. Syracuse university Press, New York.

CARSON, H.L. (1975). The genetics of speciation at the diploid level. *American Naturalist* **109**: 83-92.

CAVALLI-SFORZA, L.L. & Bodmer, W.F. (1971). *The Genetics of Human Populations.* W.H. Freeman & Company, San Francisco.

CAVALLI-SFORZA, L.L. & Edwards, A.W.F. (1967). Phylogenetic analysis: Models and estimation procedures. *American Journal of Human Genetics*. **19**: 233-257.

CHAKRABORTY, R. (1981). Gene-diversity analysis in nested subdivided populations. *Genetics* **96**: 721-726.

CHAKRABORTY, R., de-Andrade, M., Daiger, S.P. & Budowle, B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics* **56**: 45-57.

CHANG, D.D. & Clayton, D.A. (1985). Priming of human mitochondrial DNA replication occurs at the light strand promoter. Proceedings *of the National Academy of Sciences USA*, **82**: 351-355.

CHANG, D.D., Wong, T.W., Hixson, J.E. & Clayton, D.A. (1985). Regulatory sequences for mammalian mitochondrial transcription and replication. In: E. Quagliariello *et. al.* (eds.) *Achievements and Perspectives of Mitochondrial Research*, **II**: 135-144.

CHAPLIN, R.E. (1977a). *Wild Deer in Britain*. Jarrod, Norwich.

CHAPLIN, R.E. (1977b). *Deer*. Butler & Tanner, London.

CHAPMAN, D.I. & Chapman, N.G. (1982). The Taxonomic Status of Feral Muntjac Deer (Muntiacus sp. ) in England. *Journal of Natural History*. **16**: 381-387.

CHAPMAN, D.I., Chapman, N.G. & Dansie, O. (1984). The periods of conception and parturition in feral Reeves' muntjac (*Muntiacus reevesi*) in southern England, based on juvenile animals. *Journal of Zoology*, London. **189**: 259-273.

CHAPMAN, D.I., Chapman, N.G., Mathews, J.G. & Wurster-Hill, D.M. (1983). Chromosomal Studies of Feral Muntjac Deer (Muntiacus sp.) in England. *Journal of Zoology, London* **201**: 557-559

CHAPMAN, N.G. (1991a). The Chinese muntjac, *Muntiacus reevesi*. In Corbet, G.B. & Harris, S. (eds.) Handbook of British Mammals. 3rd edn. Blackwell Scientific Publishers, Oxford. pp 526-532

CHAPMAN, N.G. (1991b). Breeding performance of female Chinese Muntjac deer in England. In: Brown, D. (ed.) *The Biology of Deer*. Springer-Verlag, New York. pp 359.

CHAPMAN, N.G., Claydon, K., Claydon, M., & Harris, S. (1985). Distribution and habitat selection by Muntjac and other species of deer in a coniferous forest. *Acta Theriologica* **30**(18): 287-303.

CHAPMAN, N.G., Claydon, K., Claydon, M., Forde, P.G. & Harris, S. (1993). Sympatric populations of muntjac (*Muntiacus reevesi*) and roe deer (*Capreolus capreolus*): a comparative analysis of their ranging behaviour, social organisation and activity. *Journal of Zoology, London* **229**: 623-640.

CHAPMAN, N.G. & Harris, S. (1991). Evidence that the seasonal antler cycle of adult Reeves' muntjac (*Muntiacus reevesi*) is not associated with reproductive quiescence. *Journal of Reproduction and Fertility* **92**: 361-369.

CHAPMAN, N.G., Harris, S. & Stanford, A. (1994). Reeves' muntjac *Muntiacus reevesi* in Britain: their history, spread, habitat selection, and the role of human intervention in accelerating their dispersal. *Mammal Review* **24**(2): 113-160.

CHAPMAN, R. W., Stephens, J.C., Lansman, R. A. & Avise, J.C. (1982). Models of mitochondrial DNA transmission genetics and evolution in higher eukaryotes. *Genetic Research*, **40**: 41-57.

CHAPMAN, R.W. & Powers, D.A. (1984). A method for the rapid isolation of mitochondrial DNA from fishes. Technical Report No. UM-SG-TS-84-05. Maryland Sea Grant Program, College Park, Md.

CHESSON, P.L. & Murdoch, W.W. (1986). Aggregation of risk: relationships among host-parasitoid models. *American Naturalist* **127**: 696-715.

CHOMYN, A., Mariottini, P., Clecter, M.W.J., Ragan, C.J., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R.F. & Artardi, G. (1985). Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH complex. *Nature* **314**: 592-597.

CHOUDHARY, M., Strassmann, J.E., Solis, C.R. & Queller, D.C. (1993). Microsatellite variation in a social insect. *Biochemical Genetics* **31**: 87-96.

CLARKE, M. (1974). The survey of mammals, reptiles and amphibia in Hertfordshire. *Transactions of the Hertfordshire Natural History Society* **27**: 265-270.

CLARY, D. O., Wolstenholme, D. R. (1985). The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organisation, and genetic code. *Journal of Molecular Evolution*, **22**: 252-271.

CLAYTON, D.A. (1982). Replication of animal mitochondrial DNA. *Cell*, **28**: 693-705.

CLAYTON, D.A. (1984). Transcription of the mammalian mitochondrial genome. *Annual Review of Biochemistry*, **53**: 573-594.

CLIFF, A.D. & Ord, J.K. (1973). *Spatial Autocorrelation*. Pion, London.

CLIFF, A.D. & Ord, J.K. (1981). *Spatial Processes*. Pion, London.

COFFROTH, M.A., Lasker, H.R., Diamond, M.E., Bruenn, J.A. & Bermingham, E. (1992). DNA fingerprints of a gorgonian coral: a method for detecting clonal structure in a vegetative species. *Marine Biology* **114**: 317-325.

COMINS, H.N. & Noble, D.R. (1985). Dispersal, variability, and transient niches: species coexistence in a uniformly variable environment. *American Naturalist* **126**: 706-723.

COOPER, D.N. (1983). Eukaryotic DNA methylation. *Human Genetics* **64**: 315-533.

CORBET, G.B. (1971). Provisional distribution maps of British mammals. Mammal Review **1**: 95-142.

COTHRAN, E.G., Chesser, R.K., Smith, M.H. & Johns, P.E. (1983). Influences on genetic variability on foetal growth in white-tailed deer. *Evolution* **37**: 282-291.

COULONDRE, C., Miller, J.H., Farbaugh, P.J. & Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274**: 775-780.

CRANIUM, R. (1990). Studies in ignorance and absurdity for supervisors: 1. Methods of application to post-graduates. *The Journal of the Egotist* **1**: 1-1.

CRAWFORD, A.M., Buchanan, F.C & Swarbrick, P.A. (1990). Ovine dinucleotide repeat polymorphism at the MAF18 locus. *Animal Genetics* **21**: 433-434.

CREEL, S.R. & Waser, P.M. (1991). Failures of reproductive suppression in dwarf mongooses (*Helogale parvula*): accident or adaptation. *Behavioural Ecology* **2**: 7-16.

CREWS, S., Ojala, D., Posakony, J., Nishiguchi, J. & Attardi, G. (1979). Nucleotide sequence of a region of human mitochondrial DNA containing precisely identified origin of replication. *Nature*, **227**: 192-198.

CRONIN, M.A. (1991). Mitochondrial-DNA phylogeny of deer (Cervidae). *Journal of Mammalogy*. **72**(3): 553-565.

DALLAS, J.F. (1988). Detection of DNA 'fingerprints' of cultivated rice by hybridisation with a human minisatellite DNA probe. *Proceedings of he National Academy of Science, USA*. **85**: 6831-6835.

DALLAS, J.F. (1992). Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome* 3: 452-456

DANSIE, O. (1970). Muntjac. *British Deer Society* **2**: 1-19

DANSIE, O. (1983). Muntjac. In: *Muntjac and Chinese Water Deer*. British Deer Society, Warminster, Wiltshire. pp 3-24.

DANSIE, O. (1989). Muntjac Survey. *Mammal Society Newsletter* **79**: 2-3.

DAUSSE, E., Komajda, M., Fetler, L., Dubourg, O., Dufour, C., Carrier, L., Wisnewsky, C., Bercovici, J., Hengstenberg, C., al-Mahdawi, S., Isnadr, R., Hagege, A., Bonhour J-B., Desnoy, M., Beckman, J., Weissenbach, J., Schwartz, K. & Guicheney, P. (1993). Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and identification of a hotspot for mutations in the beta-myosin heavy chain gene. *Journal of Clinical Investigations* **92**: 2807-2813.

DAWID, I.B. & Blackler, A.W. (1972). Maternal and cytoplasmic inheritance of mitochondrial DNA in *Xenopus*. *Developmental Biology*, **29**: 1520161.

DAWKINS, R. (1989). *The Selfish Gene*. Oxford University Press, Oxford.

DESJARDINS, P. & Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. *Journal of Molecular Biology* **212**: 599-634.

DESJARDINS, P. & Morais, R. (1991). Nucleotide sequence and evolution of coding and noncoding regions of a quail mitochondrial genome. *Journal of Molecular Evolution*, **32**: 153-161

DEVI, A.R.R. & Rao, N.A. (1981). Consanguinity, fecundity and post-natal mortality in Karnataka, South India. *Annals of Human Biology* **8**: 469-472.

DEWEY, S.E. & Heywood, J.S. (1988). Spatial Genetic Structure in a Population of *Psychotria nervosa*. 1. Distribution of Genotypes. *Evolution* **42**(4): 834-838.

DIETRICH, W., Katz, H., Lincoln S.E., Shin, H.S., Freidman, J., Dracopoli, N. & Lander, E.S. (1992). A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**: 423-447.

DiRIENZO, A. & Wilson, A.C. (1991). Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, **88**: 1579-1601.

DODA, J.N., Wright, C. T. & Clayton, D.A. (1981). Elongation of displacement loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proceedings of the National Academy of Sciences USA*, **78**: 6116-6120.

DODD, J.G. & Straus, N.A. (1982). Repeated sequences in L-cell mRNA complementary to long deoxypolypyrimidines. *Acta Biochemica et Biophysica* **698**: 140-148.

DOVER, G.A. (1982). Molecular drive: a cohesive mode of species evolution. *Nature* **299** 111-117.

DUNNICK, W., Rabbits, T.S. & Milstein, C. (1980). An immunoglobin deletion mutant with implications of the heavy chain switch and RNA splicing. *Nature* **286**: 669-675.

ECONOMOU, E.P., Bergen, A.W., Warren, A.C. & Antonarakis, S.E. (1990). The polydeoxyadenylate tract of *Alu* repetitive elements is polymorphic in the human genome. *Proceedings of the National Academy of Science, USA* **87**: 2951-2954.

EDWARDS, A., Hammond, H., Jin, L., Caskey, C.T. & Chakraborty, R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**: 241-253

ELY, J & Ferrell, R.E. (1990). DNA "fingerprints" and paternity ascertainment in chimpanzees (*Pan troglodytes*). *Zoo Biology* **9**: 91-98

ELY, J., Alford. P & Ferrell, R.E. (1991). DNA "fingerprinting" and the genetic management of a captive chimpanzee population (*Pan troglodytes*). *American Journal of Primatology* **24**: 39-45.

ENDLER, J.A. (1977). *Geographic Variation , Speciation and Clines*. Princetown University Press, Princetown, N.J.

EPPLEN, J.T., Ammer, H., Epplen, C., Kammerbauer, C., Mitreiter, R., Roewer, L., Schwaiger, W., Steimle, V., Zischler, H., Albert, E., Andreas, A., Beyermann, B., Meyer, W., Beautkamp, J., Nanda, I., Schmid, M., Nürnberg, P., Pena, S.D., Pöche, H., Sprecher, W., Schartl, M., Weising, K & Yassouridis, A. (1991). Oligonucleotide fingerprinting using simple repeat motifs: a convenient, ubiquitously applicable method to detect hypervariability for multiple purposes. In: Burke, T., Dolf, G., Jeffreys, A.J. & Wolff, R. (eds.) *DNA Fingerprinting. Approaches and Applications*. Birkhaüser, Basel. pp 50-69.

EPPLEN, J.T., Cellini, A., Romero, S, & Ohno S. (1983). An attempt to approach the molecular mechanisms of primary sex determination: w- and y-chromosomal conserved simple repetitive DNA sequences and their differential expression in mRNA. *Journal of Experimental Zoology* **228**: 305-312.

ESTOUP, A., Presa, P., Krieg, F., Vaiman, D. & Guyomard, R. (1993) $(CT)_n$ and $(GT)_n$ microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity* **71**: 488-496.

EVANS, J.D. (1993). Parentage analysis in ant colonies using simple sequence loci. *Molecular Ecology* **2**(6): 393-398

EVANS, W.E., Yablokov, A.V. & Bowles, A.E. (1982). Geographic variation in the colour pattern of killer whales. *Report for the International Whaling Commission.* **32**: 687-694.

FAHRIG, L & Paloheimo, J. (1988). Effect of spatial arrangement of habitat patches on local population size. *Ecology* **69**(2): 468-475.

FALCONER, D.S. (1981). *Introduction to Quantitative Genetics.* 2nd Edition. Longman, New York.

FAURON, C.M. & Wolstenholme, D.R. (1976). Structural heterogeneity of mitochondrial DNA molecules within the genus *Drosophila. Proceedings of the National Academy of Sciences USA,* **73**: 3623-3627.

FEINBERG, A.P. & Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Annals of Biochemistry,* **132**: 6-13.

FELDMAN, D., Gagnon, J, Hofmann, R. & Simpson, J. (1987). *Statview Version 1.01.* Abacus Concepts Berkley CA.

FELSENSTEIN, J.(1982). How can we infer geography and history from gene frequencies. *Journal of Theoretical Biology* **96**: 9-20.

FEY, M.F. & Tobler, A. (1991). Assessment of DNA 'fingerprinting' as a method for validating the identity of cancer cell lines maintained in long term culture. *Nucleic Acids Research,* **19** (12): 3464.

FOOSE, T.J. (1991). Interactive management of small wild and captive populations. In: *Pink Pigeon. Conservation Viability Assessment.* Species Survival Commission/IUCN: 12-19

FORAN, D.R., Hixon, J.E. & Brown, W.N. (1988). Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. *Nucleic Acids Research* **16**(13): 5841-5862.

FORBES, S.H. & Allendorf, F.W. (1991a). Association between mitochondrial and nuclear genotypes in cutthroat trout hybrid swarms. *Evolution* **45**(6): 1332-1349.

FORBES, S.H. & Allendorf, F.W. (1991b). Mitochondrial genotypes have no detectable effects on meristic traits in cutthroat trout hybrid swarms. *Evolution* **45**(6): 1350-1359.

FORD, J.K.B. & Fisher, H.D. (1982). Killer whale (*Orcinus orca*) dialects as an indicator of stocks in British Columbia. *Report for the International Whaling Commission* **32**: 671-679.

FOREY, P.L. (1988). Golden jubilee for the coelacanth *Latimeria chalumnae. Nature,* **336**: 727-732.

FOSBROOKE, H.A. (1963). The Stomoxys plague in Ngorongoro, 1962. *East African Widlife Journal* **I**: 124-126.

FRANCISCO, J.F., Brown, G.G. & Simpson, M.V. (1979). Further studies on types A and B rat mtDNAs: cleavage maps and evidence for cytoplasmic inheritance in mammals. *Plasmid* **2**: 426-436.

FRESCO, J.R. & Alberts, B.M. (1960). The accommodation of noncomplementary bases in helical polyribonucleotides and deoxyribonucleic acids. *Proceedings of the National Academy of Science, USA* **46**: 311-321.

FRIEDLANDER, M. (1980). Monospermic fertilisation in *Chrysopa cornia* (Neuroptera: Chryopidae): Behaviour of the fertilising spermatozoa prior to syngamy. *International Journal of Insect Morphology and Embryology.* **9**: 53-57

FRIES, R., Eggen, A. & Womack, J.E. (1993). The Bovine Genome Map. *Mammalian Genome.* **4**: 405-428

FRITZSCH, B. (1987). The inner ear of the coelacanth fish *Latimeria* has tetrapod affinities. *Nature,* **327**: 153-154.

FU, Y-H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M., Holden, J.J.A., Fenwick, R.G.Jr., Warren, S.T., Oostra, B.A., Nelson. D.L. & Caskey, C.T. (1992). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **67**: 1047-1058,

FUENTES, J.J., Banchs, I., Volpini, V. & Estivill, X. (1993). Genetic variation of microsatellite markers D1S117, D6S89, D11S35, APOC2, and D21S168 in the Spanish population. *International Journal of Legal Medicine* **105**: 271-277.

GABRIEL, K.R. & Sokal, R.R. (1969). A new statistical approach to geographic variation analysis. *Systematic Zoology* **18**: 259-270.

GADELATA, G., Pepe, G., De Candia, G., Quagliariello, C., Sbisá, E., Saccone, C. (1989). The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. *Journal of Molecular Evolution*, **28**: 497-516.

GARESSE, R. (1988). Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations. *Genetics*, **118**: 649-663.

GARTEN, C.T. (1976). Relationships between aggressive behaviour and genetic heterozygosity in the oldfield mouse, *Peromyscus polionotus. Evolution* **30**: 59-72.

GEBHARD, W . & Zachau, H.C. (1983). Simple DNA sequences and dispersed repetitive elements in the vicinity of mouse immunoglobin K light chain genes. *Journal of Molecular Biology* **170** (2): 567-573.

GEORGE, M, Jr. & Ryder, O.A. (1986). Mitochondrial DNA evolution in the genus *Equus. Molecular Biology and Evolution* 3(6): 535-546.

GEORGES, M., Lequarre, A.S., Castelli, M., Hanset, R. & Vassart, G. (1988). DNA fingerprinting in domestic animals using four different minisatellite probes. *Cytogenetics and Cell Genetics* **47**: 127-131.

GIBBS, H.L., Weatherhead, P.J., Boag, P.T., White, B.N. Tabak, L.M. & Hoysak, D.J. (1990). Realised reproductive success of polygynous red-winged blackbirds revealed by DNA markers. *Science* **250**: 1394-1397.

GILBERT, D.A., Lehman, N., O'Brien, S.J. & Wayne, R.K. (1990). Genetic fingerprinting reflects population differentiation in the Channel Island Fox. *Nature* **344**: 764-767.

GILBERT, D.A., Packer, C., Pusey, A.E., Stephens, J.C. & O'Brien, S.J. (1991). Analytical DNA Fingerprinting in Lions: Parentage, Genetic Diversity, and Kinship. *Journal of Heredity* **82**: 378-386.

GILES, R.E., Blane, H., Cann, H.M. & Wallace, D.C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, **77**: 6715-6718.

GILL, P., Lygo, J.E., Fowler, S.J. & Werrett, D.J. (1987). An evaluation of DNA fingerprinting for forensic purposes. *Electrophoresis* **8**:38-44

GILL, P. & Werrett,D.J. (1987). Exclusion of a man charged with murder by DNA fingerprinting. *Forensic Science International* **35**:145-148

GILLESPIE, J.H. (1984). The molecular clock may be an episodic clock. *Proceedings of the National Academy of Sciences USA*, **81**: 8009-8013.

GILPIN, M.E. (1987). Spatial structure and population vulnerability. In: Soulé, M.E. (ed.)*Viable Populations for Conservation*. Cambridge University Press, Cambridge. pp 125-139.

GJETVAJ, B., Cook, D.I. & Zouros, E. (1992). Repeated sequences and large-scale size variation of mitochondrial DNA: a common feature among scallops (Bivalvia: Pectinidae). *Molecular Biology and Evolution*, **9**(1): 106-124.

GONZALEZ-ZULUETA, M., Ruppert, J.M., Tokino, K., Tsai, Y.C., Spruck, C.H., Miyao, N., Nichols, P.W., Hermann, G.G., Horn, T., Steven, K., Summerhayes, I.C., Sidransky, D. & Jones. P.A. (1993). Microsatellite instability in bladder cancer. *Cancer Research* **53**: 5602-5623.

GOODBOURNE, S.E.Y., Higgs, D.R., Clegg, J.B. & Weatherall, D.J. (1983). Molecular basis of length polymorphism in the human ζ-globin gene complex. *Proceedings of the National Academy of Science USA*. **80**: 5022-5025.

GOTTELLI, D., Sillero-Zubiri, C., Applebaum, G.D., Rot, M.S., Girman, D.J., Garcia-Moreno, J., Ostrander, E.A. & Wayne , R.K. (1994). Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis Simensis. Molecular Ecology* in press

GRACHEV, M.A., Slobobyanynk, S.J., Kholodilov, N.G., Fyodorov, S.P. Belikov, S.I., Sherbakov, D.Yu., Sideleva, V.G., Zubin, A.A. & Kharchenko, V.V. (1992). Comparative study of two protein coding regions of mitochondrial DNA from three endemic sculpins (Cottidei) of Lake Baikal. *Journal of Molecular Evolution* **34**(1): 85-90.

GREEN, P.M., Bentley, D.R., Mibashan, R.S., Nilsson, I.M. & Gianelli, F. (1989). Molecular pathology of haemophilia B. *EMBO*, **8**: 1067-1072.

GREENBERG, B. D., Newbold, J.E. & Sugino, A. (1983). Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene*, **21**: 33-49.

GRESSON, R.A.R. (1940). Presence of the sperm middle-piece in the fertilised egg of the mouse (*Mus musculus*). *Nature*, **145**: 425

GRETH, A., Sunnucks, P., Vassart, M. & Stanley, H. (1992). Genetic management of an Arabian Oryx (*Oryx leucoryx*) population without known pedigree. In: Spitz, Janeau, Gonzalez, Aulagnier (eds.) "Ongulés / Ungulates 91". Proceedings of the International Symposium "Ongulés / Ungulates 91". S.F.E.P.M. - I.R.G.M. Paris

GYLLENSTEN, U., Wharton, D. & Wilson, A.C. (1987). Maternal inheritance of mitochondrial DNA during back-crossing of two species of mice. *Journal of Heredity* **76**(5): 321-324.

GYLLENSTEN, U., Wharton, D., Josefsson, A. & Wilson, A. C. (1991). Paternal inheritance of mitochondrial DNA in mice. *Nature*, **352**: 255-257.

GYLLENSTEN, U.B., Jakobsson, S. & Temrin, H. (1990). No evidence for illegitimate young in monogamous and polygynous warblers. *Nature* **343**: 168-170.

HAIGH, J. & Maynard-Smith, J. (1972). Population size and protein variation in man. *Genetic Research* **19**: 73-89.

HALE, L.R. & Singh, R.S. (1987). Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **4**(6): 622-637.

HAMADA, H. & Kakunaga, T. (1982). Potential Z DNA forming sequences are highly dispersed in the human genome. *Nature* **298**: 396-398.

HAMADA, H., Petrino, M.G. & Kakunaga, T. (1982). A novel repeated element with Z DNA forming potential is widely found in evolutionary diverse eukaryote genomes. *Proceedings of the National Academy of Science USA* **79**: 6465-6469.

HAMADA, H., Seidman, M., Howard, B.H. & Gorman, C.M. (1984). Enhanced gene expression by the poly (dT-dG) poly (dC-dA) sequence. *Molecular and Cell Biology* **4**: 2622-2630.

HAMAGUCHI, K., Itô, Y. & Takenaka, O. (1993). GT dinucleotide repeat polymorphisms in a polygynous ant *Leptothorax spinosior*, and their use for measurement of relatedness. *Naturwissenschaften* **80**: 179-181.

HAMILTON, W.D. (1964). The genetic evolution of social behaviour. I and II. *Journal of Theoretical Biology* **7**: 1-52.

HANOTTE, O., Bruford, M.W. & Burke, T. (1992). Multilocus DNA fingerprints in gallinaceous birds: general approach and problems. *Heredity* **68**: 481-494.

HANSKI, I. (1982). Dynamics of regional distribution: the core and satellite species hypothesis. *Oikos* **38**: 210-221.

HANSKI, I. (1985). Single-species spatial dynamics may contribute to long-term rarity and commonness. *Ecology* **66**: 335-343.

HANSKI, J. (1991). Single Species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnaean Society*. **42**: 17-35.

HARDING, S.P. (1986). *Aspects of the Ecology and Social Organisation of the Muntjac Deer* (Muntiacus reevesi). D.Phil. thesis University of Oxford.

HARTL, D.L. & Clark, A.G. (1989). *Principles of Population Genetics*. Sinauer Associates, Sunderland MA. USA.

HARRIS, S., Creswell, W.J., Forde, P.G., Trewhella, W.J., Woollard, T. & Wray, S. (1990). Home-range analysis using radio-tracking data - a review of problems and techniques particularly as applied to the study of mammals. *Mammal Review* **20**: 97-123.

HASTINGS, A. (1982). Dynamics of a single species in a spatially varying environment: the stabilising role of high dispersal rates. *Journal of Mathematical Biology* **16**: 49-55.

HASTINGS, Duke of Bedford (1949).*The Years of Transition* . Andrew Dakers Ltd. London.

HAUGE, X.Y. & Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Human Molecular Genetics* **2**: 441-415.

HAYASHI, J.I., Yonekawa, H., Gotoh, J., Watanabe, J. & Tagashira, Y. (1978). Strictly maternal inheritance in rat mitochondrial DNA. *Biochemistry and Biophysics Research Communications*. **83**: 1032-1038.

HEILIG, R., Muraskowsky, R. & Mandel, J.L (1982). The ovalbumin gene family: the 5′ end region of the X and Y genes. *Journal of Molecular Biology* **156**: 1019.

HIGGS, D.R., Goodbourne, S.E.Y., Wainscott, J.S., Glegg. J.B. & Weatherall, D.J. (1981). Highly variably regions of DNA flank the human globin genes. *Nucleic Acids Research* **9**: 4213-4224.

HILLEL, J., Plotzt, Y., Haberfield, A., Lavi, U., Cahaner, A. & Jeffreys, A.J. (1989). DNA Fingerprints in Poultry. *Animal Genetics* **20**: 145-155.

HILLEL. J., Schaap, T., Haberfeld, A., Jeffreys, A.J., Plotzky, Y., Cahaner, A. & Lavi, U. (1990). DNA fingerprints applied to gene introgression in breeding programs. *Genetics* **124**: 783-789.

HIXON, J.E. & Brown, W. M. (1986). A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution and phylogenetic implications. *Molecular Biology and Evolution* **3**: 1-18.

HOELZEL A.R. & Dover G.A. (1991). Genetic differentiation between sympatric killer whale populations. *Heredity* **66**:191-195.

HOELZEL A.R. & Osborne, R.W. (1986). Killer whale call characteristics; implications for co-operative foraging strategies. In: Kirkevold, B.C. & Lockard J.S. (eds.) *Behavioural Biology of Killer Whales*. A.R. Liss, New York. pp 373-406.

HOELZEL, A.R. & Green, A. (1992). Analysis of population-level variation by sequencing PCR-amplified DNA. In: A.R. Hoelzel (ed.) *Molecular Genetic Analysis of Populations*. IRL Press at Oxford University Press. pp 159-186. (D. Rickwood, B.D. Hames. The Practical Approach Series).

HOELZEL, A.R., Hancock, J. M. & Dover, G.A. (1991). Evolution of Cetacean Mitochondrial D-Loop Region. *Molecular Biology and Evolution* **8**(3): 475-493.

HOLT, R.D. (1985). Population dynamics in two-patch environments: some anomalous consequences of an optimal habitat distribution. *Theoretical Population Biology* **28**: 181-208.

HORAI, S., Gojobori, T. & Matsunaga, E. (1986). Distinct clustering of mitochondrial DNA types among Japanese, Caucasians and Negroes. *Japanese Journal of Genetics*, **61**: 271-275.

HOUSTON, D. (1974). The northern Yellowstone elk. Parts II and III. History and demography. *Yellowstone National Park Report*. [Cited in Cameron & Vyse 1978]

HUANG, T., Cottingham Jr., R., Ledbetter, D. & Zoghbi, H. (1992). Genetic mapping of four dinucleotide repeat loci , DXS453, DXS458, DXS454 AND DXS424 on the X chromosome using multiplex polymerase chain reaction. *Genomics* **13**: 375-380.

HÜBSCHER, U., Kuenzle, C.C. & Spadari, S. (1979). DNA polymerase $\gamma$ is the only polymerase present in the mitochondria and is responsible for mtDNA replication not repair. *Proceedings of the National Academy of Science USA* **76**(5): 2316-2320.

HUGHES, C.R. & Queller, D.C. (1993). Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology* **2**: 131-137.

HYNES, R.A., Duke, E.J. & Joyce, P. (1989). Mitochondrial DNA as a genetic marker for brown trout, *Salmo trutta* L., populations. *Journal of Fish Biology*, **35**: 687-701.

IRWIN, D.M., Kocher, T.D. & Wilson, A.C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of Molecular Evolution*, **32**: 128-144.

JACQUARD, D.D. (1974). *The Genetic structure of Populations*. Springer Verlag, New York.

JARMAN, A.P. & Wells, R.A. (1989). Hypervariable minisatellites: recombinators or innocent bystanders. *Trends in Genetics* **5**(11): 367-371.

JARNE, P., Delay, B., Bellec, C., Roizes, G. & Cuny, G. (1990). DNA fingerprinting in schistosome-vector snails. *Biochemical Genetics* **28**: 577-583.

JARNE, P., Delay, B., Bellec, C., Roizes, G. & Cuny, G. (1992). Analysis of mating systems in the schistosome-vector hermaphrodite snail *Bulinus globosus* by DNA fingerprinting. *Heredity* **68**: 141-146.

JEANG, K.T. & Hayward, G.S. (1983). A cytomegalovirus DNA sequence containing tracts of tandemly repeated CA dinucleotides hybridises to highly repetitive dispersed elements in mammalian cell genomes. *Molecular and Cell Biology* **3**: 1389-1402.

JEFFREYS, A.J. (1979). DNA sequence variants in the $^{G}\gamma$-, $^{A}\gamma$-, $\delta$- and $\beta$-globin genes of man. *Cell* **18**: 1-10

JEFFREYS, A.J., Brookfield, J.F.Y. & Semeonoff, R. (1985[c]). positive identification of an immigration test-case using human DNA fingerprints. *Nature* **317**: 818-819.

JEFFREYS, A.J. & Morton, D.B. (1987). DNA fingerprints of dogs and cats. *Animal Genetics* **18**: 1-15.

JEFFREYS, A.J., Turner, M. & Debenham, P. (1991). The efficiency of multilocus DNA finger print probes for individualisation and establishment of family relationships, determined from extensive casework. *American Journal of Human Genetics.* **48**: 824-840.

JEFFREYS, A.J., Wilson, V., Kelly, R., Taylor, B.A. & Bulfield, G. (1987). Mouse DNA "fingerprints": analysis of chromosome localisation and germline stability of hypervariable loci in recombinant inbred strains. *Nucleic Acids Research* **15**: 2823-2836.

JEFFREYS, A.J., Wilson, V. & Thein, S.L. (1985a). Hypervariable 'minisatellite' regions in human DNA. *Nature.* **314**: 67-74.

JEFFREYS, A.J., Wilson, V. & Thein, S.L. (1985b). Individual-specific DNA 'fingerprints' of human DNA. *Nature* **316**: 76-79.

JEFFREYS, A.J., Wilson, V. & Thein, S.L., Weatherall, D.J. & Ponder, B.A.J. (1986[a]). DNA "Fingerprints" and Segregation Analysis of Multiple Markers in Human Pedigrees. *American Journal of Human genetics.* **39**: 11-24.

JEFFREYS, A.J., Wilson, V., Wong, Z., Royale, N., Patel, I., Kelly, R. & Clarkson, R. (1986[b]). Highly variable minisatellites and DNA fingerprints. *Biochemical Society Symposium Series* **53**: 165-180.

JEWELL, P.A. & Bullock, D.J. (1991). Chinese muntjac *Muntiacus reevesi.* In Corbet, G.B. & Harris, S. (eds.) *The Handbook of British Mammals.* Blackwell Scientific Publications, London. pp 526-552.

JOHANSSEN, S., Guddal, P.E. & Johanssen, T. (1990). Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua. Nucleic Acids Research* **18**(3): 411-419.

JOHNSON, D. & Morgan, A.R. (1978). Unique structure formed by pyrimidine•purine DNAs which may be four stranded. *Proceedings of the National Academy of Science USA* **75**: 1637-1641.

JOHNSTON, F.P., Church, R.B. & Lin, C.C. (1981). Chromosome rearrangement between the Indian muntjac and Chinese muntjac is accompanied by a deletion of middle repetitive DNA. *Canadian Journal of Biochemistry* **60**: 497-506.

KARLIN, S. (1969). *Equilibrium Behavior of Population Genetic Models with Non-Random Mating.* Gordon & Breach, New York.

KEANE, B., Waser, P.M., Creel, S.R., Creel, N.M., Elliot, L.F. & Minchella, D.J. (1994). Subordinate reproduction in dwarf mongooses. *Animal Behaviour* **47**: 65-75.

KEDES, L.H. (1979). Histone genes and Histone messengers. *Annual Review of Biochemistry* **48**: 837-870.

KEMPENAERS, B., Verheyen, G.R., Van den Broeck, M., Burke, T., Van Broeckhoven, C & Dhondt, A.A. (1992). Extra-pair results from female preference for high-quality males in the blue tit. *Nature* **357**: 494-496.

KENDAL, M.G. (1973). *Time-Series.* Griffin, London.

KESSLER, L.G. & Avise, J.C. (1985). A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. *Molecular Biology and Evolution,* **2**(2): 109-125.

KIMURA, M. & Crow, J.F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**: 725-738.

KIMURA, M . & Ohta, T. (1978). Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences USA,* **75**: 2868-2872.

KOCHER, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X. & Wilson, A.C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences USA,* **86**: 6196-6200

KOEHLER, C.M., Lindberg, G.L., Brown, D.R., Beitz, D.C., Freeman, A.E., Mayfield, J.E. & Myers, A.M. (1991). Replacement of bovine mitochondrial DNA by a sequence variant within one generation. *Genetics,* **129**: 247-255.

KOEHN, R.K., Turano, F.J., Mitton, J.B. (1973). Population genetics of marine pelecypods II. Genetic differences in microhabitats of *Modiolus demissus. Evolution* **27**: 100-105.

KONDO, R., Satta, Y., Matsura, E.T., Ishiwa, H., Takahata, N. & Chigusa, S. I. (1990). Incomplete maternal transmission of mitochondrial DNA in *Drosophila. Genetics,* **126**: 657-663.

KOOREY, D.J., Bishop, G.A. & McCaughan, W. (1993). Allele non-amplification: a source of confusion in linkage studies employing microsatellite polymorphisms. *Human Molecular Genetics* **2**(3): 289-191.

KUNKEL, T.A. (1985). the mutational specificity of DNA polymerase-beta during *in vitro* DNA synthesis: production of frame shift, base substitution and deletion mutations. *Journal of Biological Chemistry* **260**: 5787-5796.

KURNIT, D.M. (1979). Satellite DNA and Heterochromatin variants: the case for unequal mitotic crossing over. *Human Genetics* **47**: 169-186.

KWAITKOWSKI, D., Henske, E., Weimer, K., Ozelius, L., Gusella, J. & Haines, J. (1992). Construction of a GT polymorphism map of human 9q. *Genomics* **12**: 229-240.

LAGERCRANTZ, U., Ellegren, H, & Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research* **21**(5): 1111-1115.

LAGIOS, M.D. (1982). *Latimeria* and the Chondrichthyse as sister taxa: a rebutal to recent attempts of refutation. *Copeia* **1982**: 942-948.

LAIKRE, L., Tegelström, H. & Gelter H.P. (1992). DNA fingerprints of captive wolves (*Canis lupus*). *Hereditas* **117**: 293-296.

LAMB, T. & Avise, J.C. (1986). Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: The influence of mating behaviour. *Proceedings of the National Academy of Sciences USA*, **83**: 2526-2530.

LANDE, R. (1988). Genetics and demography in biological conservation. *Science* **241**: 1455-1460.

LANDER, E. S. (1989). DNA fingerprinting on trial. *Nature* **339**: 501-505.

LANSMAN, R.A. Avise, J.C., Aquadro, C.F., Shapira, J.F. & Daniel, S.W. (1983). Extensive genetic variation in mitochondrial DNA's among geographic populations of the deer mouse *Peromyscus maniculatus*. *Evolution* **37**(1): 1-16.

LANSMAN, R.A., Shade, R.O., Shapira, J.F. & Avise, J.C. (1981). The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations III. Techniques and potential applications. *Journal of Molecular Evolution*, **17**: 214-226.

LÁZARO, C., Gaona, A., Ravella, A., Volpini, V., Casals, T., Fuentes, J.J. & Estivill, X. (1993). Novel alleles, hemizygosity and deletions at an Alu-repeat within the neurofibromatosis type 1 (NF1) gene. *Human Molecular Genetics* **2**: 725-730.

LEDIG, F.T., Guries, R.P. & Bonefield, B.A. (1983). The relation of growth to heterozygosity in pitch pine. *Evolution* **37**: 1227-1238.

LEE, C., Sasi, R. & Lin, C.C. (1993). Interstitial localisation of telomeric DNA sequences in the Indian muntjac chromosomes: further evidence for tandem chromosome fusions in the karyotypic evolution of the Asian muntjacs. *Cytogenetics and Cell Genetics* **63**: 156-159.

LEHMAN, N., Clarkson, P., Mech, D.L., Meier, T.J. & Wayne, R.K. (1992). A study of the genetic relationships within and between wolf packs using DNA fingerprinting and mitochondrial DNA. *Behavioural Ecology and Sociobiology* **30**: 83-94.

LEHMAN, N., Eisenhawer, A., Hansen, K., Mech, D.L., Peterson, R.O., Gogan, P.J.P. & Wayne, R.K. (1991). Introgression of coyote mitochondrial DNA into sympatric north American gray wolf populations. *Evolution*, **45**(1): 104-119.

LESICA, P., Leary, R.F., Allendorf, F.W. & Bilderback, D.E. (1988). Lack of genetic diversity within and among populations of an endangered plant, *Howellia aquatilis*. *Conservation Biology* **2**: 275-282.

LEVER, C. (1977). *The Naturalised Animals of the British Isles*. Hutchinson. London.

LEVER, C. (1978). *The Naturalised Mammals of the World*. Longman. London.

LEVIN, S.A. (1974). Dispersal and population interactions. *American Naturalist* **108**: 207-228.

LEVINSON, G., Marsh, L., Epplen, J.T. & Gutman, G.A. (1985). Cross-hybridising snake satellite, Drosophila, and mouse DNA sequences may have arisen independently. *Molecular Biology and Evolution* **2**: 494-504.

LEVINSON, G. & Gutman, G.A. (1987a). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* **4**(3): 203-221.

LEVINSON, G & Gutman, G.A. (1987b). High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Research* **15**: 5323-5338.

LEWIN, R. (1989). Limits to DNA fingerprinting. *Science* **243**: 1549-1551

LEWONTIN. R.C. (1974). *The Genetic Basis of Evolutionary Change.* Columbia University Press, N.Y.

LEWONTIN, R.C & Cockerham, C.C. (1959). The goodness-of-fit test for detecting natural selection in random mating populations. *Evolution* **13** : 561-564

LIMA-de-FARIA, A., Arnason, U., Widegren, B., Essen-Moller, J., Isaksson, M., Olsson, E. & Jaworska, H. (1984). Conservation of repetitive DNA sequences in deer species studied by Southern blot transfer. *Journal of Molecular Evolution* **20**: 17-24.

LI, W-H. (1986). Evolutionary change of restriction cleavage sites and phylogenetic inference. *Genetics*, **113**: 187-213.

LITT, M. & Luty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**: 397-401.

LITT, M. (1991). PCR of TG microsatellites. In McPherson, M.C., Quirke, P. & Taylor, G. (eds.) *PCR: A Practical Approach.* Oxford University Press. pp 85-89.

LIU, J., Lissens, W., Devroey, P., Van Steirteghem, A. & Liebaers, I. (1992). Efficiency and accuracy of PCR assay for cystic fibrosis allele Delta F508 in single cell. *Lancet* **339**(8803). 1190-1192.

LIVNEH, Z. (1983). Directed mutagenesis method for analysis of mutagen specificity: application to ultraviolet-induced mutagenesis. *Proceedings of the National Academy of Science USA* **80**: 4639-4643.

LOMNICKI, A (1980). Regulation of population density due to individual differences and patchy environment. *Oikos* **35**: 185-193.

LOUDON, A.S.I. (1987). The influence of forest habitat structure on growth, body size and reproduction in roe deer (*Capreolus capreolus* L.). In: Wemmer, C.M. (ed.) *Symposium of Biology and Management of the Cervidae.* Smithsonian Institution Press, Washington D.C. pp 559-568.

LØVTRUP, S. (1977). *The phylogeny of vertebrata.* Wiley, London.

LUTY, J.A., Guo, Z., Wilard, H.F., Ledbetter, D.H., Ledbetter, S. & Litt. M. (1990). Five polymorphic microsatellite VNTRs on the human X chromosome. *American Journal of Human Genetics* **46**: 776-783.

LYNCH, M. (1988). Estimating relatedness by DNA fingerprinting. *Molecular Biology and Evolution.* **5**: 584-599.

LYNCH, M. (1990). The Similarity Index and DNA Fingerprinting. *Molecular Biology and Evolution* 7(5): 478-484.

LYNCH, M. (1991). Analysis of Population Genetic structure by DNA Fingerprinting. In: T. Burke, G. Dolf, A.J. Jeffreys and R. Wolff (eds.) *'DNA Fingerprinting: Approaches and Applications.' Birkhaüser* Verlag Basel, Switzerland.

LYNCH, M. & Crease, T.J. (1990). The analysis of population survey data on DNA sequence variation, *Molecular biology and Evolution*, **7**: 377-394.

MacHUGH, D.E., Loftus, R.T., Bradley, D.G., Sharp, P.M. & Cunningham, D.G. (1994). Microsatellite DNA variation within and among European Cattle Breeds. *Proceeding of the Royal Society*, in Press.

MacRAE, A.F. & Anderson, W.W. (1988). Evidence for non-neutrality of mitochondrial DNA haplotypes in *Drosophila pseudoobscura*. *Genetics*, **120**: 485-494.

MAHER, D.S. (1990). The Florida panther and private lands. *Conservation Biology* **4**: 167-170.

MANTEL, N. (1967). The detection of disease clustering and a generalised regression approach. *Cancer Research* **27**: 209-220.

MARINELLI, L., Messier, F. & Plante, Y. (1992). Use of DNA fingerprinting to determine parentage in muskrats (*Ondatra zibethicus*). *Journal of Heredity* **83**: 3560360.

MARTIN, A.P., Naylor, G.J.P. & Palumbi, S.R. (1992). Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature*, **357**: 153-155.

MAUGHAN, E. & Williams, J.R.B. (1967). Haemoglobin types in deer. *Nature* **215**: 404-405.

MAX, E.E. (1984). New twist to DNA methylation. *Nature* **310**: 100.

MAY, R.M. (1979). When to be incestuous. *Nature* **279**: 192-194.

MAYHOOK, A.G., Rinaldi, A. & Jacobs, H.T. (1992). Replication origins and pause sites in sea urchin mitochondrial DNA. *Proceedings of the Royal Society of London*, **248**: 85-94.

MAYNARD SMITH, J. (1989). *Evolutionary Genetics.* Oxford University Press, Oxford.

McCABE, P.C. (1990). Production of Single Stranded DNA by asymmetric PCR. In: M.A. Innis, D.H. Gelfand, J.J. Sninsky, T.J. White. (editors). *PCR Protocols*. Academic Press. New York. pp 76-83.

MENG, A., Carter, R.E. & Parkin, D.T. (1990). The variability of DNA fingerprints in three species of swan. *Heredity* **64**: 73-80.

MERRIWETHER, D.A., Clark, A.G., Ballinger, S.W., Schurr, T.G., Soodyall, H., Jenkins, T., Sherry, S.T. & Wallace, D.C. (1991). The structure of human mitochondrial DNA variation. *Journal of Molecular Evolution*, **33**: 543-555.

MEYER, A., Kocher, T.D., Basasibwaki, P. & Wilson, A.C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, **347**: 550-553.

MEYER, W., Kock, A., Nieman, C., Beyermann, B., Epplen, J.T. & Börner, Th. (1991). Differentiation of species and strains among filamentous fungi by DNA fingerprinting. *Current Genetics* **19**: 239-242.

MEYER, W., Lieckfeldt, E., Wösremeyer, J. & Börner, Th. (1992). DNA Fingerprinting for differentiating aggressivity groups of the rape seed pathogen *Leptosphaeria maculans Mycological Research*. **96**(8): 651-657.

MEYER, A. & Wilson, A.C. (1990). Origin of Tetrapods inferred from their mitochondrial DNA affiliation to lungfish. *Journal of Molecular Evolution*, **31**: 359-364.

MIESFELD, R., Krystal, M & Arnheim N. (1981). A member of a new repeat sequence family which is conserved throughout eukaryotic evolution is found between the human delta and beta globin genes. *Nucleic Acids Research* **9**: 5931-5947.

MIGNOTTE, F., Gueride, M., Champagne, A-M. & Mounolou, J-C. (1990). Direct repeats in the non-coding region of rabbit mitochondrial DNA. *European Journal of Biochemistry*, **194**: 561-571.

MITCHISON, H.M., Thompson, A.D., Mulley, J.C., Kozman, H.M., Richards, R.I., Callen, D.F., Stallings, R.L., Doggett, N.A., Attwood, J., McKay, T.R. Sutherland, G.R. & Gardiner, M.R. (1993). Fine genetic mapping of the Batten disease locus (CLN3) by haplotype analysis and demonstration of allelic association with chromosome 16p microsatellite. *Genomics* **16**: 455-460.

MITTON, J.B. (1993). Theory and data pertinent to the relationship between heterozygosity and fitness. In: Thornhill, N.W. (ed.) *The Natural History of Inbreeding and Outbreeding*. The University of Chicago Press, London.

MITTON, J.B. & Koehn, R.K. (1975). Genetic organisation and adaptive response of allozymes to ecological variables in *Fundulus heteroclitus*. *Genetics* **79**: 97-111.

MITTON, J.B. & Pierce, B.A. (1980). The distribution of individual heterozygosity in natural populations. *Genetics* **95**: 1043-1054.

MIYAMOTO, M.M., Kraus, F. & Ryder, O.A. (1990). Phylogeny and evolution of antlered deer determined from mitochondrial DNA sequences. *Proceedings of the National Academy of Sciences USA*,, **87**: 6127-6131.

MOORE, S.S., Sargeant, L.L., King, T.J., Mattick, J.S., Georges, M. & Hetzel, D.J.S. (1991). The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, **10**: 654-660.

MORAN, P.A.P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika*, **37**: 17-23.

MORITZ, C. & Brown, W.M. (1987). Tandem duplications in animal mitochondrial DNAs: Variation in incidence and gene content among lizards. *Proceedings of the National Academy of Sciences USA*, **84**: 7183-7187.

MORITZ, C., Dowling, T.E. & Brown, W.M. (1987). Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual Review of Ecological Systematics*, **18**: 269-292.

MORK, J. & Sundnes, G. (1985). O-group cod (*Gadus morhua*) in captivity: differential survival of certain genotypes. *Helgolander Meeresuntersuchungen* **39**: 63-70.

MORRAL, M., Nunes, V., Casals, T., Chillon, M., Gimenez, J., Bertranpetit, J. & Estivill, X. (1993a). Microsatellite haplotypes for cystic fibrosis: mutation frameworks and evolutionary tracers. *Human Molecular Genetics* **2**: 1015-1022.

MORRAL, M., Nunes, V., Casals, T., Cobos, N., Asensio, O., Dapena, J. & Estivill, X. (1993b). Uniparental inheritance of microsatellite alleles of the cystic fibrosis gene (CFTR): identification of a 50kb deletion. *Human Molecular Genetics* **2**: 677-681.

MORTON, D.B., Yaxley, R.E., Patel, I., Jeffreys, A.J., Howes, S.J. & Debenham, P.G. (1987). Use of DNA fingerprint analysis in identification of the sire. *The Veterinary Record*. **121**: 592-593.

MULLUS, K.B. & Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Methods in Enzymology,* **155**: 335-350.

MURRAY, J., Stine, O.C. & Johnson, M.S. (1991). The evolution of mitochondrial DNA in *Partula*. *Heredity* **66**: 93-104.

MYERS, J.H. (1976). Distribution and dispersal in populations capable of resource depletion. *Oecologia* (Berlin) **23**: 255-269.

NEI, M. (1965). Variation and covariation of gene frequencies in subdivided populations. *Evolution* **19**: 256-258.

NEI, M. (1972). Genetic distance between populations. *American Naturalist* **106**: 283-292.

NEI, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA,* **70**: 3321-3323.

NEI, M. (1977). F-statistics and the analysis of gene diversity in subdivided populations. *Annals of Human Genetics* **41**: 225-233.

NEI, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

NEI, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York.

NEI, M. & Li, W-H. (1976). The transient distribution of allele frequencies under mutation pressure. *Genetic Research* **28**: 205-215.

NEI, M. & Li, W-H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA,* **76**: 5269-5273.

NEI, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck event and genetic variability in populations. *Evolution* **29**: 1

NEI, M. & Tajima, F. (1983). Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics,* **105**: 207-217.

NIGRO, L. & Prout, T. (1990). Is there selection on RFLP differences in mitochondrial DNA? *Genetics,* **125**: 551-555.

NIKAIDO, T., Nakai. S. & Honjo, T. (1981). Switch region of immunoglobin Cμ gene is composed of simple tandem repetitive sequences. *Nature* **292**: 845-848.

NORDHIEM, A & Rich, A (1983). The sequence (dC-dA)n•(dT-dG) forms left-handed Z DNA in negatively supercoiled plasmids. *Proceedings of the National Academy of Science USA* **80**: 1821-1825.

O'BRIEN, S.J., Martenson, J.S., Packer, C., Herbst, L., de Vos, V., Joslin, P., Ott-Joslin, J., Wildt, D.E. & Bush, M. (1987). Biochemical genetic variation in geographic isolates of African and Asiatic lions. *National Geographic Research* **3**: 114-124.

O'BRIEN, S.J., Roelke, M.E., Marker, L. Newman, A., Winkler, C.W., Meltzer, D., Colly, L., Everman, J., Bush, M. & Wildt, D.E. (1985). A genetic basis for species vulnerability in the cheetah. *Science* **27**: 1428-1434.

O'BRIEN, S.J., Roelke, M.E., Yuhki, N., Richards,, K.W., Johnson, W.E., Franklin, W.L., Anderson, A.E., Bass, O.L., Belden, R.C. & Martenson, J.S. (1990). Genetic introgression within the Florida panther (*Felis concolor coryi*). *National Geographic Resources* **6**: 485-494.

O'BRIEN, S.J., Wildt, D.E., Goldman, D., Meril, C.R. & Bush, M. (1983). The cheetah is depauperate in genetic variation. *Science* **221**: 459-462.

ODELBERG, S.J., Demers, D.B., Westin, E.H. & Hossaini. (1988). Establishing paternity using minisatellite DNA probes when the putative father is unavailable for testing. *Journal of Forensic Sciences* **33**: 921-928.

ODEN, N.L. (1984). Assessing the significance of a spatial correlogram. *Geographical Analysis* **16**(1): 1-16.

ODEN, N.L. & Sokal, R.R. (1986). Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology*. **35**(4): 608-617.

OHNO, S. (1970). *Evolution by gene duplication*. Springer, New York.

OHNO, S. & Epplen, J.T. (1983). The primitive code and repeats of base oligomers as the primordial protein-encoding sequence. *Proceedings of the National Academy of Science USA* **80**: 3391-3395.

OHTA, T., & Kimura, M. (1973). A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genetic Research **22**: 201-204

OLIVO, P.D., Van de Walle, M.J., Laipis, P.J. & Hauswirth, W.W. (1983). Nucleotide sequence evidence for rapid genotype shifts in the bovine mitochondrial DNA D-loop. *Nature*, **306**: 400-402.

OSTRANDER, E.A., Sprague, G.F. & Rine, J. (1993). Identification and characterisation of dinucleotide repeat (CA)n markers for genetic mapping in dog. *Genomics*, **16**: 207-213.

OWEN, J.E., Schultz, D.W., Taylor, A. & Smith, G.R. (1983). Nucleotide sequence of the lysozyme gene of bacteriophage T4: analysis of mutations involving repeated sequences. *Journal of Molecular Biology* **165**: 229-248.

PÄÄBO, S., Gifford, J.A. & Wilson, A.C. (1988). Mitochondrial DNA sequence from a 7000-year old brain. *Nucleic Acids Research* **16**(20): 9775-9788.

PACKER, C. (1979). Inter-troop transfer and inbreeding avoidance in *Papio anubis. Animal Behaviour* **27**: 1-36.

PACKER, C., Gilbert, D.A., Pusey, A.E. & O'Brien, S.J. (1991a). Molecular genetic analysis of kinship and co-operation in African Lions. *Nature* **351:** 562-565

PACKER, C., Pusey, A.E., Rowley, H., Gilbert D.A., Martenson, J.S. & O'Brien, S.J. (1991b). Case study of a population bottleneck: lions of the Ngorongoro Crater. *Conservation Biology* **5**: 1-12.

PALVA, T.K. & Palva, E.T. (1985). Rapid isolation of animal mitochondrial DNA by alkaline extraction. *FEBS*, **192**(2): 267-270.

PARDUE, M.L., Lowenhaupt, K., Rich, A & Nordheim, A (1987). (dC-dA)n (dG-dT)n sequences have evolutionarily conserved chromosomal locations in Drosophila with implications for roles in chromosome structure and function. *EMBO Journal* **6**:1781-1789.

PEMBERTON. J.M., Albon, S.D., .Guinness, F.E., Clutton-Brock T.H. & Dover G.A. (1992). Behavioural estimates of male mating success tested by DNA fingerprinting in a polygynous mammal. *Behavioural Ecology* **3**(1): 66-75.

PESOLE, G., Sbisá, E., Preparata, G. & Saccone, C. (1992). The evolution of the mitochondrial D-loop region and the origin of modern man. *Molecular Biology and Evolution*, **9**(4): 587-598.

PHILLIPS, H.A., Harris, P., Richards, R.I., Sutherland, G.R. & Mulley, J.C. (1991b). Dinucleotide repeat polymorphism at the D16S164, D16S168, D16S186 loci at 16q21-q22.1. *Nucleic Acids Research* **19**: 6964.

PHILLIPS, H.A., Hyland, V.J., Holman, K., Callen, D.F., Richards, R.I. & Mulley, J.C. (1991a). Dinucleotide repeat polymorphism at D16S287. *Nucleic Acids Research* **19**: 6664.

PICKVANCE, T.J. & Chard, J.S.R. (1960). Midland mammals survey 1. Feral muntjac deer (*Muntiacus* spp.) in the West Midlands, with special reference to Warwickshire. *Proceedings of the Birmingham Natural History and Philosophical Society* 19: 1-8,

PIERCE, B.A. & Mitton, J.B. (1982). Allozyme heterozygosity and growth in the tiger salamander, *Amblystoma tigrinum. Journal of Heredity* **73**: 250-253.

PROUDFOOT, N.J. Gil, A. & Maniatis, T. (1982). The structure of the human Zeta-Globin gene and a closely linked, nearly identical psuedogene. *Cell* **31**: 553-563.

PUTMAN, R. (1988). *The Natural History of Deer*. Christopher Helm, Kent.

QUELLER, D.C. & Goodnight, K.F. (1989). Estimating relatedness using genetic markers. *Evolution* **43**: 258

QUINN, T.W., Quinn, J.S., Cooke. F. & White B.N. (1987). DNA marker analysis detects multiple maternity and paternity in single broods of the lesser snow goose. *Nature* **326**: 392-394.

RALLS, K. & Ballou, J. (1982a). Effects of inbreeding on infant mortality in captive primates. *International Journal of Primatology* **3**: 491-505

RALLS, K. & Ballou, J. (1982b). Effects of inbreeding on juvenile mortality in some mammal species. *Laboratory Animals* **16**: 159-166.

RALLS, K. & Ballou, J. (1983). Extinction: lesson from zoos. In: Schonewald-Cox, C.M., Chambers, S.M., MacBride, B. & Thomas, L. (eds.). *Genetics and Conservation: A Reference for Managing Wild Animal and plant Populations.* Benjamin-Cummings, Menlo Park, California. pp 164-184.

RALLS, K., Brugger, K. & Ballou, J. (1979). Inbreeding and juvenile mortality in small populations of ungulates. *Science* **206**: 1101-1103.

RALLS, K., Harvey, P. H. & Lyles, A.M. (1986). Inbreeding in Natural Populations of Birds and Mammals. In Soulé, M.E. (ed.) *Conservation Biology: The Science of Scarcity and Diversity.* Sinauer. Massachusetts. pp 35-57.

RAND, D.M., Harrison, R.G. (1986). Mitochondrial DNA transmission genetics in crickets. *Genetics,* **114**: 955-970.

RATCLIFFE, P.R. (1991). Sika deer *Cervus nippon*. In: Corbet G.B. & Harris, S. (eds.). *The Handbook of British Animals*, 3rd edn. Blackwell Scientific Publications, London. pp 504-508.

RAZIN, A. & Riggs, A.D. (1980). DNA methylation and gene function. *Science* **210**: 604-610.

RICHARDS, R.I., Shen, Y., Holman, K., Kozman, H., Hyland, V.J., Mulley, J.C. & Sutherland, G.R. (1991). Fragile X syndrome: diagnosis using highly polymorphic microsatellite markers. *American Journal of Human Genetics* **48**: 1051-1057.

RIEGSTRA, R. (1993). *The use of DNA fingerprinting in parentage analysis and population viability assessment in Muntiacus reevesi.* M.Sc. Thesis, Durrell Institute of Conservation and ecology, University of Kent. U.K.

ROELKE, M.E., Martenson, J.S. & O'Brien, S.J. (1993). The consequences of a demographic reduction and genetic depletion in the endangered Florida panther. *Current Biology* **3**(6): 340-350.

ROEWER, L., Nagy, M., Schmidt, P., Epplen, J.T. & Herzog-Schöder, G (1993). Microsatellite and HLA class II oligonucleotide typing in a population of Yanomama indians. In: Pena S.D.J., Chakraborty, T., Epplen, J.T. & Jeffreys, A.J. (eds.) *Fingerprinting: State of the Science.* Birkhäuser Verlag, Basel. pp 221-230.

ROGERS, J (1983). CACA sequences - the ends and the means? *Nature* **305**: 101-102.

ROGSTAD, S.H., Patton, J.C. & Sahaal, B.A. (1988). A human minisatellite probe reveals RFLPs among individuals of two angiosperms. *Nucleic Acids Research* **16**(23): 11378.

ROLF, D.A. (1974). Spatial heterogeneity and the persistence of populations. *Oecologia,* (Berlin) **15**: 245-258.

ROMER, A.S. (1966). *Vertebrate Palaeontology.* 3rd Edition. University of Chicago Press. Chicago.

ROSSI, R.E., Mulla, D.J., Journel, A.G. & Franz, E.H. (1992). Geostatistical tools for modelling and interpreting ecological spatial dependence. *Ecological Monographs* **62**: 277-314.

ROWE, B.A., Ma, D-P., Wilson, R.K. & Wong, J.F.-H. (1985). The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *Journal of Biochemistry,* **260**(17): 9759-9774.

ROY, M.S., Geffen, E., Smith, D., Ostrander, E.A. & Wayne R.K. (1994). Patterns of differentiation and hybridisation in North American wolf-like canids revealed by analysis of microsatellite loci. *Molecular Biology and Evolution* **11**(4): 553-570.

RUSSELL, D.W., Smith, M., Cox, D., Williamson, V.M. & Young, E.T. (1983). DNA sequences of two yeast promoter-up mutants. *Nature* **304**: 652-654.

RYDER, O.A., Chemnick, L.G., Schaffer, S.F. & Shima, A.L. (1989). Individual DNA fingerprints from Galapagos tortoises. *International Zoo Yearbook* **28**: 84-87.

RYMAN, N., Beckman, G., Brunn-Petersen, G. & Reuterwall, C. (1977). variability of red cell enzymes and genetic implications of management policies in Scandinavian moose (*Alces alces*). *Hereditas* **85**: 157-162.

RYSKOV, A.P., Jincharadze, A.G., Prosnyak, M.I., Ivanov, P.L. & Limborska, S.A. (1988). M13 phage DNA as a universal marker for DNA fingerprinting of animals, plants and microorganisms. *FEBS Letters* **233**(2): 388-392.

SACCONE, C., Pesole, G. & Sbisá, E. (1991). The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *Journal of Molecular Evolution,* **33**: 83-91.

SAHA, B.K., Shields, J.J., Miller, R.D., Hansen, T.H. & Shreffler, D.C. (1993). A highly polymorphic microsatellite in the class II Eb gene allows tracing of major histocompatability complex evolution in mouse. *Proceedings of the National Academy of Science USA,* **90**: 5312-5316. [Erratum in Proc. Nat. Acad. Sci. USA **90**: 8757].

SAMBROOK, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular cloning. A Laboratory Manual .* (2nd. Edn.). Cold Harbour Laboratory, New York.

SANGER, F., Nicklen, S. & Coulsden, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Science USA*, **74**: 5463-5467.

SCHALLER, G.B. (1975). *The Serengeti Lion.* University of Chicago Press, Chicago.

SCHERTHAN, H. (1990). Localisation of the repetitive telomeric sequence $(TTAGGG)_n$ in two muntjac species and implications for their karyotypic evolution. *Cytogenetics and Cell Genetics* **53**: 115-117.

SCHLÖTTERER, C., Amos, W. & Tautz, D (1991). Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, **354**: 63-65.

SCHLÖTTERER, C. & Tautz D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**(2): 211-215.

SCHMID, C.W. & Shen, C-K.J. (1985). The evolution of interspersed repetitive DNA sequences in mammals and other vertebrates. In: MacIntyre, R.J. (ed.)*Molecular Evolutionary Genetics*. Plenum Press, New York. pp 325-358.

SCHMIDTKE, J., Brennecke, H., Schmid, M., Neitzel, H. & Sperling, K. (1981). Evolution of muntjac DNA. *Chromosoma* **84**: 187-193.

SEAL, U.S. & Lacy, R. (1989). Florida panther population viability analysis. Report to the U.S. Fish and Wildlife Service. Captive Breeding Specialist Group, Species Survival Commission, IUCN, Apple Valley, Minnesota.

SEDEROFF, R.R. (1984). Structural variation in mitochondrial DNA. *Advances in Genetics*, **22**: 2-82.

SEEMANOVA, E. (1971). A study of children of incestuous mating. *Human Heredity* **21**: 108-128.

SERRADILLA, J.M. & Ayala, F.J. (1983). Alloprocoptic selection: A mode of natural selection promoting polymorphism. *Proceedings of the National Academy of Science* USA **80**: 2022-2025.

SHAFFER, M.L. (1981). Minimum population sizes for species conservation. *BioScience* **31**: 131-134.

SHENG, H. (1992). Reeves' muntjac *Muntiacus reevesi*. In: Sheng, H (ed.). *The Deer of China.* East China Normal University Press, Shanghai. pp 126-144.

SHI, L. & Pathak, S. (1981). Gametogenesis in a Male Indian Muntjac x Chinese Muntjac hybrid. *Cytogenetics and Cell Genetics* **30**: 152-156.

SHI, L., Ye, Y. & Duan, X. (1980). Comparative Cytogenetic Studies on the Red Muntjac, Chinese Muntjac, and their F1 Hybrids. *Cytogenetics and Cell Genetics* **26**: 22-27.

SINGH, R.S. & Hale, L.R. (1990). Are mitochondrial DNA variants selectively non-neutral? *Genetics*, **124**: 995-997.

SINGH, S.M. & Zouros, E. (1978). Genetic variation associated with growth rate in the American oyster (*Crassostrea virginica*). *Evolution* **32**: 342-353.

SJÖGREN, P. & Wyöni, P-I. (1992). Conservation genetics and detection of rare alleles in finite populations. *Conservation Biology* **8**(1): 267-270.

SLAGBOOM, P.E., Mullaart, E., Droog. S. & Vijg, J. (1991). Somatic mutations and cellular ageing: two-dimensional DNA typing of rat fibroblast clones. *Mutation Research* **256**: 311-321.

SLATKIN, M. (1987). Gene flow and geographic structure of natural populations. *Science* **236**: 787-792.

SLATKIN, M. & Arter, H.E. (1991). Spatial autocorrelation methods in population genetics. *The American Naturalist*. **138**: 499-517.

SLIGHTOM, J.L., Blechl, A.E. & Smithies, O. (1980). Human foetal $^G\gamma$ and $^A\gamma$ globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* **21**: 627-638.

SMEETS, H.J.M., Brunner, H.G., Ropers, H.H. & Wieringa, B. (1989). Use of variable simple sequence motifs as genetic markers: Application to study of myotonic dystrophy. *Human Genetics* **83**: 245-251.

SMITH, G.P. (1973). Unequal cross over and the evolution of multigene families. *Quantitative Biology* **38**: 507-513.

SMITH, G.P. (1976). Evolution of repeated DNA sequences by unequal crossing over. *Science* **191**: 528-535.

SMITH, R.H. (1979). On selection for inbreeding in polygynous animals. *Heredity* **43**: 205-211.

SMITH, M.H., Garten, C.T. & Ramsev, P.R. (1975). Genetic heterozygosity and population dynamics of small mammals. In: Market, C.L. (ed.). *Isozymes IV: Genetics and Evolution*. Academic Press New York.

SNYDER, M., Fraser, A.R., LaRoche, J., Gartner-Kepkay, K.E .& Zouros, E. (1987). Atypical mitochondrial DNA from the deep-sea scallop *Placopecten magellanicus*. *Proceedings of the National Academy of Sciences USA*, **84**: 7595-7599.

SOKAL, R.R. (1979). Ecological parameters inferred from spatial correlograms. In G.P. Patil & M.L. Rosenzweig (eds.), *Contemporary Quantitative Ecology and Related Ecometrics*. International Cooperative Publishing House, Fairland, MD. pp167-196.

SOKAL, R.R. & Oden, N.L. (1978a). Spatial autocorrelation in biology 1 Methodology. *Biological Journal of the Linnean Society*. **10**: 199-228.

SOKAL, R.R. & Oden, N.L. (1978b). Spatial autocorrelation in biology 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society*. **10**: 229-249.

SOKAL, R.R. & Rohlf, F.J. (1981). *Biometry*. W.H. Freeman and Co. San Francisco.

SOKAL R.R. & Wartenberg, D.E. (1981). Space and population structure, in: Griffith, D, & McKinnin, R. (eds.) *Dynamic Spatial Models*. Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands.

SOLIGNAC, M., Genermont, J., Monnerot, M. & Mounolou, J-C. (1984). Genetics of mitochondria in *Drosophila* : mtDNA inheritance in heteroplasmic strains of *D. mauritiana*. *Molecular Genetics*, **197**: 183-188.

SOLIGNAC, M., Monnerot, M. & Mounolou, J-C. (1986). Concerted evolution of sequence repeats in *Drosophila* mitochondrial DNA. *Journal of Molecular Evolution*, **24**: 53-60.

SOPER, E.A. (1969). *Muntjac*. Longmans, Green & Co. Ltd., London.

SORGE, J. & Hughs, S.H. (1982). Polypurine tract adjacent to the U3 region of the Rous sarcoma virus genome provides a cis-acting function. *Journal of Virology* **43**:n 482-488.

SOUTHERN, E.I. (1972). Repetitive DNA in mammals. In Pfeiffer R.A. (ed.) *Modern Aspects of Cytogenetics: Constitutive Heterochromatin in Man*. Schattauer, Stuttgart FRG. pp 19-28.

SOUTHERN, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**: 503-517.

SOUTHERN, S.O., Southern, P.J. & Dizon, A.E. (1988). Molecular characterisation of a cloned dolphin mitochondrial genome. *Journal of Molecular Evolution*, **28**: 32-42.

STALLINGS, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E. & Moyzis, R.K. (1991). Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics* **10**: 807-815.

STEPHAN, W. (1989). Tandem-repetitive noncoding DNA: forms and forces. *Molecular Biology and Evolution* **6**(2): 198-212.

STRAUS, N.A. & Birnboim, H.C. (1976). Polypyrimidine sequences found in eukaryotic DNA have been conserved during evolution. *Acta Biochemica et Biophysica* **454**: 419-428.

STREISINGER, G. & Owen, J. (1985). Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics* **104**: 633-659.

STRINGER, J.R. (1982). DNA sequence homology and chromosomal deletion at a site of SV40 DNA integration. *Nature* **296**:363.

STURMBAUER, C. & Meyer A. (1992). Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature*, **358**; 578-581.

STRYER, L. (1975). *Biochemistry*. W.H. Freeman & Co. San Francisco.

SWARBRICK, P.A., Buchanan, F.C. & Crawford, A.M. (1990). Ovine dinucleotide repeat polymorphism at the MAF23 locus. *Animal Genetics* **21**:191.

SWARBRICK, P.A., Buchanan, F.C. & Crawford, A.M. (1991a). Ovine dinucleotide repeat polymorphism at the MAF35 locus. *Animal Genetics* **22**: 369-370.

SWARBRICK, P.A., Buchanan, F.C. & Crawford, A.M. (1991b). Ovine dinucleotide repeat polymorphism at the MAF36 locus. *Animal Genetics* **22**: 377-378.

SWARBRICK, P.A., Buchanan, F.C. & Crawford, A.M. (1991c). Ovine dinucleotide repeat polymorphism at the MAF64 locus. *Animal Genetics* **22**: 375-376.

SWARBRICK, P.A., Buchanan, F.C. & Crawford, A.M. (1992). Ovine dinucleotide repeat polymorphism at the MAF50 locus. *Animal Genetics* **23**: 187

SWARBRICK, P.A. & Crawford, A.M. (1992). Ovine dinucleotide repeat polymorphism at the MAF109 locus. *Animal Genetics* **23**: 84.

SWARBRICK. P.A., Dietz, A.B., Womack, J.E. & Crawford, A.M. (1992). Ovine and bovine dinucleotide repeat polymorphisms at the MAF46 locus. *Animal Genetics* **23** 182.

SWOFFORD, D. L. (1990). PAUP: Phylogenetic Analysis Using Parsimony, Version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign, Illinois.

SWOFFORD, D.L. & Selander, R.B. (1981). BIOSYS-1: A FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity* **72**: 281-283. (Release 7.1)

TAMURA, K. & Aotsuka, T. (1988). Rapid isolation of animal mitochondrial DNA by the alkaline lysis procedure. *Biochemical Genetics*, **26**(11/12): 815-819.

TAUTZ. D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**: 6464-6471.

TAUTZ, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryote genomes. *Nucleic Acids Research* **12**: 4127-4138.

TAUTZ, D., Trick, M. & Dover, G.A. (1986). Cryptic simplicity in DNA is a major source of genetic variation. *Nature* **322**: 652-656.

TAYLOR A.C., Sherwin, B.W. & Wayne R.K. (1994). Genetic variation of simple sequence loci in bottlenecked species: the decline of the northern hairy-nosed wombat (*Lasiorhinus krefftii*). *Molecular Ecology*, in press.

TAYLOR, L.R. & Taylor, A.J (1977). Aggregation, migration and population mechanics. *Nature* **265**: 415-421.

TEE, L.A., Rowe, J.J. & Pepper, H.W. (1985). *Mammal Bird Damage Questionnaire 1983*. Forestry Commission Research and Development Paper 137. Forestry Commission, Edinburgh.

TEGELSTROM, H. (1987). Genetic variability in mitochondrial DNA in a regional population of the great tit (*Parus major*). *Biochemical Genetics*, **25**(1/2): 95-110.

TEMPLETON, A.R. (1983). Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution*, **37**(2): 221-244.

THOMAS, W.K., Pääbo, S., Villablanca, F.X. & Wilson, A.C. (1990). Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. *Journal of Molecular Evolution*, **31**: 101-112.

TODD, J.A., Aitman, T.J., Cornall, R.J., Ghosh, S., Hall, J.R., Hearne, C.M., Knight, A.M., Love, J.M., McAleer, M.A., Prins, J.B., Rodrigues, N., Lathrop, M., Pressey, A., DeLarato, N.H., Peterson, L.B. & Wicker, L.S. (1991) Genetic analysis of autoimmune type I diabetes melitus in mice. *Nature* **351**: 542-547.

TOPAL, M.D. & Fresco, J.R. (1976). Complementary base paring and the origin of substitution mutations. *Nature* **263**: 285-289.

TORRONI, A., Schurr, T.G., Yang, C.-C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K.M. & Wallace, D.C. (1992). Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics*, **130**: 153-162.

TRIGGS, S.J., Williams M.J.U., Marshall, S.J. & Chambers G.K. (1992). Genetic structure of the blue duck (Hymenolaimus malacorhynchos) populations revealed by DNA fingerprinting. *The Auk* **109**(1): 80-89.

TZAGOLOFF, A., Myers, A.M. (1986). Genetics of mitochondrial biogenesis. *Annual Review of Biochemistry*, **55**: 249-285.

UPHOLT, W.B. (1977). Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. *Nucleic Acids Research*, **4**(5): 1257-1265.

UPHOLT, W.B. & Dawid, I.B. (1977). Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D-loop region. *Cell*, **11**: 571-583.

URSPRING, H & Schabtach, E. (1965). Fertilisation in tunicates: loss of the paternal mitochondrion prior to sperm entry. *Journal of Experimental Zoology*, **159**: 379-384.

VALDES, A.M., Slatkin, M. & Freimer, N.B. (1993). Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* **133**: 737-749.

VANCE, R.R. (1984). The effect of dispersal on population stability in one species, discrete-space population growth models. *American Naturalist* **123**: 230-254.

VASSART, G., Georges, M., Monsieur, R., Brocas. H., Lequarre, A.S. & Christophe, D. (1987). A sequence of M13 phage detects hypervariable minisatellites in human and animal DNA. *Science* **235**: 683-684.

VAWTER, L. & Brown, W.M. (1986). Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. *Science,* **234**: 194-195.

VIGILANT, L., Pennington, R., Harpending, H., Kocher, T.D. & Wilson, A.C. (1989). Mitochondrial DNA sequences in single hairs from a southern African population. *Proceedings of the National Academy of Sciences USA,* **86**: 9350-9354.

WAHLUND, S. (1928). Zusammensetzung von Populationen und Korrelationsercheinungen von Standpunkt der Vererbungslehre aus betrachtet. *Hereditas,* **11**: 65-106.

WAINSCOAT. J. (1987). Out of the garden of Eden. *Nature,* **325**: 13

WALBERG, M.W. & Clayton, D.A. (1981). Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Research* **9**: 5411-5421.

WALLACE, D.C. (1982). Structure and evolution of organelle genomes. *Microbiological Reviews* **46**: 208-240.

WALLACE, D.C. (1992). Mitochondrial genetics: A paradigm for ageing and degenerative diseases? *Science,* **256**: 628-632.

WALLACE, R. (1923). *Farm Livestock of Great Britain.* (5th edn.). Oliver & Boyd Edinburgh.

WALMSLEY, R.M., Wilkinson, B.M. & Kong, T.H. (1989). Genetic fingerprinting for yeasts. *Bio/Technol.* **7** : 1168-1170.

WAYNE, R.K. Lehman, N., Girman, D., Gogan, P.J.P., Gilbert, D.A., Hansen, K., Peterson, R.O., Seal, U.S., Eisenhawer, A., Mech, L.D. & Krumenaker, R.J. (1991a). Conservation genetics of the Isle Royale Gray Wolf. *Conservation Biology* **5**(1): 41-51.

WAYNE, R.K., George, S.B., Gilbert, D., Collins, P.W., Kovach, S.D., Girman, D. & Lehman, N. (1991b). A morphologic and genetic study of the island fox *Urocyon littoralis. Evolution* **45**(8): 1849-1868.

WAYNE, R.K., Van Valkenburgh, B., Kat, P.W., Fuller, T.K., Johnson, W.E. & O'Brien, S.J. (1989). Genetic and Morphological divergence among sympatric canids. *Journal of Heredity* **80**: 447-454.

WEBER, J.L. (1990a). Informativeness of human $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphisms. *Genomics* **7**: 524-530.

WEBER, J.L. (1990b). Human DNA polymorphisms based on length variations in simple sequence tandem repeats. In Tilghman, S. & Davies, K. (eds.) *Genome Analysis Series. Genetic and Physical Mapping.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Volume 1: 159-181.

WEBER, J.L. (1990c). Human DNA polymorphisms and methods of analysis. *Current Opinions in Biotechnology* **1**: 166-171.

WEBER, J.L. & May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the Polymerase Chain Reaction. *American Journal of Human Genetics.* **44**: 228-297.

WEBER, J.L., Polymeropoulos, M.H., May, P.E., Kwitek, A.E., Xiao, H., McPherson, J.D. & Wasmuth, J.J. (1991). Mapping of human chromosome 5 microsatellite DNA polymorphisms. *Genomics* **11**: 695-700.

WEISSENBACH, J (1993). Microsatellite polymorphisms and the genetic linkage map of the human genome. *Current Opinions in Genetic Development* **3**: 414-417.

WEISSENBACH, J., Gyapay, G., Dib, C., Vignal, A., Morisette, J., Millasseau, P., Vaysseix, G. & Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature* **359**: 794.

WELTER C., Dooly, S. & Blin, N (1989). A rapid protocol for the purification of mitochondrial DNA suitable for studying restriction fragment length polymorphisms. *Gene,* **83**: 169-172.

WEMMER, C.M. (1987). Symposium of Biology andManagement of the Cervidae. Smithsonian Institute Press. Washington D.C. pp 110-121.

WESTNEAT, D.F. (1990). Genetic parentage in the indigo bunting: a study using DNA fingerprinting. *Behavioural Ecology and Sociobiology* **27**: 67-76.
WETTON, J.H. Carter, R.E., Parkin, D.T. & Walters, D. (1987). Demographic study of a wild house sparrow populations by DNA fingerprinting. *Nature* **327**: 147-149.
WHITEHEAD, G.K. (1950). *Deer and their Management*. Country Life Ltd. London.
WHITEHEAD, G.K. (1964). *The Deer of Great Britain and Ireland*. Routledge & Kegan Paul.
WHITEHEAD, G.K. (1972). *Deer of the World*. Constable, London.
WIER, B.S. & Cockerham, C.C. (1984). Estimating F statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.
WIESNER, R.J., Kurowski, T.T. & Zak, R. (1992). Regulation by thyroid hormone of nuclear and mitochondrial genes encoding subunits of cytochrome-c oxidase in rat liver and skeletal muscle. *Molecular Endocrinology* **6**(9): 1458-1467.
WILDT, D.E., Bush, M., Goodrowe, K.L., Packer, C., Pusey, A.E., Brown, J.L., Joslin, P. & O'Brien, S.J. (1987). Reproductive and genetic consequences of founding isolated lion populations. *Nature* **329**: 328-331.
WILKINSON, G.S. & Chapman, A.M. (1990). Length and sequence variation in evening bat D-loop mtDNA: *Genetics*, **128**: 607-617.
WILLIAMS, G.C. (1966). *Adaptation and Natural Selection*. Princetown University Press, Princetown N.J.
WILSON, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D. & Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnaean Society*. **26**:357-400.
WILSON, A.C. & Sarich, V.M. (1967). Immunological time scale for evolution. *Science*, **158**: 1200-1203
WILSON, A.C., Zimmer, E.A., Prager, E.M. & Kocher, T.D. (1989). Restriction mapping in the molecular systematics of mammals: a retrospective salute. In: Fernholm, B., Bremer, K. & Jörnvall, H. (eds.) *The Hierarchy of Life*. Elsevier, Amsterdam. pp 407-419.
WINSHIP, P.R. (1989). An improved method for directly sequencing PCR amplified material using DMSO. *Nucleic Acids Research*, **17**: 1266.
WINTERO, A.K., Fredholm, M, & Thomsen, P.D. (1992). Variable $(dG\text{-}dT)_n \bullet (dC\text{-}dA)_n$ sequences in the porcine genome. *Genomics* **12**: 281-288.
WONG, A.K., Yee, H.A., Van de Sande, J.H. & Rattner, J.B. (1990). Distribution of CT rich tracts inversed in invertebrate chromosomes. *Chromosoma (Berlin)* **99**: 344-351.
WONG, Z., Wilson, V., Jeffreys, A.J. & Thein, S.L. (1986). Cloning a selected fragment from a human DNA 'fingerprint': isolation of an extremely polymorphic minisatellite. *Nucleic Acids Research*. **14**: 4605-4616.
WRIGHT, S. (1923). Mendelian analysis of the pure breeds of livestock, II. The Duchess family of Shorthorns as bred by Thomas Bates. *Journal of Heredity* **14**: 405-422.
WRIGHT, S. (1931). Evolution in Mendelian Populations. *Genetics* **16** : 97-159.
WRIGHT, S. (1943). Isolation by Distance. *Genetics* **28**: 114-138.
WRIGHT, S. (1951). The Genetic Structure of Populations. *Annals of Eugenics* **15**: 323-354.
WRIGHT, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**: 395-420.
WRIGHT, S. (1969). *Evolution and the Genetics of Populations*. Vol. 2: *The Theory of Gene Frequencies*. University of Chicago Press. Chicago: Illinois.
WRIGHT, S. (1982). The shifting balance theory and microevolution. *Annual Review of Genetics*. **16**: 1-19.
WRISCHNIK, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N. & Wilson, A.C. (1987). Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. *Nucleic Acids Research*, **15**(2): 529-542.
WURSTER, D.H. & Atkin, N.B. (1972). Muntjac Chromosomes: a new karyotype for Muntiacus muntjac. *Experientia* **28**: 972-972.
WURSTER, D.H. & Benirschke. (1967). Chromosome studies in some deer, the springbok and the pronghorn, with notes on placentation in deer. *Cytologia* **32**: 273-285
WYMAN, A. & White, R (1980). A highly polymorphic locus in human DNA. *Proceedings of the National Academy of Science USA*. **77**: 67-54-6758.

YAP, E.P.H & McGee, J.O. (1991). Short PCR product fields improved by lower denaturation temperatures. *Nucleic Acids Research* **19**: 1713.

YOKOI, T., Odaira, T., Nata, M. & Sagisaka, K. (1990). Investigation of paternity establishing without the putative father using hypervariable DNA probes. *Japanese Journal of Human Genetics* **35**: 235-244.

YOUNG, J.Z. (1977). *The Life of Vertebrates*. 2nd. edn. Oxford University Press, Oxford.

YOUNG, S. (1990) Naked mole rats keep it in the family. *New Scientist* **12**: 38.

YU, L.C., Lowensteiner, D., Wong, E.F., Sawada, I., Mazrimas, J. & Schmid, C. (1986). Localisation and characterisation of recombinant DNA clones derived from the highly repetitive DNA sequences in the Indian muntjac cells: their presence in the Chinese muntjac. *Chromosoma* **93**: 521-528.

YUHKI, N. & O'Brien, S.J. (1990). DNA variation in the mammalian major histocompatibility complex reflects genomic diversity and population history. *Proceedings of the National Academy of Science USA.* **87**: 836-840.

ZHANG, L., Cui, X., Schmitt, K., Herbert, R., Navidi, W. & Arnheim, N. (1992). Whole genomic amplification from a single cell: implications for genetic analysis. *Proceedings of the National Academy of Science USA.* **89**: 5847-5851.

ZOUROS, E., Freeman, K.R., Ball, A.O. & Pogson, G.H. (1992). Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature*, **359**: 412-414.

ZOUROS, E., Smith, S.M. & Miles, H.E. (1980). Growth rate in oysters: an overdominant phenotype and its possible explanation. *Evolution* **34**: 856-867.

ZULIANI, G. & Hobbs. H.H. (1990). A high frequency of length polymorphisms in repeated sequences adjacent to Alu sequences. *American Journal of Human Genetics* **46**: 936-969.

# APPENDIX A

## SAMPLE SITES & INDIVIDUALS

| Location | Grid Ref. | Individual number | Sex | Location | Grid Ref. | Individual number | Sex |
|---|---|---|---|---|---|---|---|
| WAVERLY | | | | SALCEY FOREST | | | |
| Waverly & West Wd | SP 3570 | Wav 01 | M | Salcey Forest | SP 7981 | Sal 01 | M |
| Waverly & West Wd | SP 3570 | Wav 02 | M | Salcey Forest | SP 7981 | Sal 02 | M |
| Waverly & West Wd | SP 3570 | Wav 03 | M | Salcey Forest | SP 7981 | Sal 03 | M |
| Waverly & West Wd | SP 3570 | Wav 04 | F | Salcey Forest | SP 7981 | Sal 04 | M |
| Waverly & West Wd | SP 3570 | Wav 05 | F | Salcey Forest | SP 7981 | Sal 05 | M |
| TUSMORE | | | | Salcey Forest | SP 7981 | Sal 06 | M |
| Tusmore Park | SP563088 | Tus 01 | F | Salcey Forest | SP 7981 | Sal 07 | M |
| Tusmore Park | SP 5630 | Tus 02 | M | Salcey Forest | SP 7981 | Sal 08 | M |
| Tusmore Park | SP 5630 | Tus 03 | M | Salcey Forest | SP 7981 | Sal 09 | F |
| Tusmore Park | SP 5630 | Tus 04 | F | Salcey Forest | SP 7981 | Sal 10 | F |
| Tusmore Park | SP 5630 | Tus 05 | F | Salcey Forest | SP 7981 | Sal 11 | F |
| Tusmore Park | SP 5630 | Tus 06 | F | Salcey Forest | SP 7981 | Sal 12 | F |
| Tusmore Park | SP 5630 | Tus 07 | M | Horton | SP 8254 | Htn 01 | ? |
| Tusmore Park | SP 5630 | Tus 08 | M | Horton | SP 8254 | Htn 02 | ? |
| Bicester Oxfd. | SP 5-2- | Bst 01 | F | WOBURN ABBEY | | | |
| Bicester Oxfd. | SP 5-2- | Bst 02 | M | Woburn Abbey | SP 9632 | Wob 01 | M |
| Bicester Oxfd. | SP 5-2- | Bst 03 | M | Woburn Abbey | SP 9632 | Wob 02 | M |
| Willaston | SP 6132 | Wil 01 | M | Woburn Abbey | SP 9632 | Wob 03 | M |
| Willaston | SP 6132 | Wil 02 | M | Woburn Abbey | SP 9632 | Wob 04 | M |
| Willaston | SP 6132 | Wil 03 | M | Woburn Abbey | SP 9632 | Wob 05 | M |
| Willaston | SP 6132 | Wil 04 | M | Woburn Abbey | SP 9632 | Wob 06 | F |
| Willaston | SP 6132 | Wil 05 | F | Woburn Abbey | SP 9632 | Wob 07 | F |
| Willaston | SP 6132 | Wil 06 | F | Woburn Abbey | SP 9632 | Wob 08 | F |
| Willaston | SP 6132 | Wil 07 | F | Woburn Abbey | SP 9632 | Wob 09 | F |
| SOUTHBLOCK | | | | Woburn Abbey | SP 9632 | Wob 10 | F |
| South Block | SP 6542 | Sbl 01 | M | Woburn Abbey | SP 9632 | Wob 11 | F |
| South Block | SP 6542 | Sbl 02 | M | Woburn Abbey | SP 6932 | Wob 12 | ? |
| South Block | SP 6542 | Sbl 03 | M | Woburn Abbey | SP 6932 | Wob 13 | M |
| South Block | SP 6542 | Sbl 04 | M | Woburn Abbey | SP 6932 | Wob 14 | M |
| South Block | SP 6542 | Sbl 05 | M | Woburn Abbey | SP 6932 | Wob 15 | ? |
| South Block | SP 6542 | Sbl 06 | F | Woburn Abbey | SP 6932 | Wob 16 | M |
| South Block | SP 6542 | Sbl 07 | M | Woburn Abbey | SP 6932 | Wob 17 | M |
| South Block | SP 6542 | Sbl 08 | F | Woburn Abbey | SP 6932 | Wob 18 | ? |
| South Block | SP 6542 | Sbl 09 | F | Woburn Abbey | SP 6932 | Wob 19 | ? |
| Hazelborough Wood | SP 6543 | Hzl 01 | M | Woburn Abbey | SP 6932 | Wob 20 | F |
| Hazelborough Wood | SP 6543 | Hzl 02 | M | Woburn Abbey | SP 6932 | Wob 21 | M |
| Hazelborough Wood | SP 6543 | Hzl 03 | M | Woburn Abbey | SP 6932 | Wob 22 | M |
| Hazelborough Wood | SP 6543 | Hzl 04 | M | Woburn Abbey | SP 6932 | Wob 23 | ? |
| Hazelborough Wood | SP 6543 | Hzl 05 | F | Woburn Abbey | SP 6932 | Wob 24 | F |
| Hazelborough Wood | SP 6543 | Hzl 06 | F | HARRY'S PARK | | | |
| Bucknell | SP 6545 | Bkl 01 | M | Harry's Park | SP 9487 | Hpk 01 | M |
| Bucknell | SP 6545 | Bkl 02 | M | Harry's Park | SP 9487 | Hpk 02 | M |
| Bucknell | SP 6545 | Bkl 03 | M | Harry's Park | SP 9487 | Hpk 03 | M |
| Bucknell | SP 6545 | Bkl 04 | F | Fermyn Woods | SP966857 | Fmw 01 | M |
| Bucknell | SP 6545 | Bkl 05 | F | Fermyn Woods | SP980840 | Fmw 02 | F |
| Bucknell | SP 6545 | Bkl 06 | F | Bathurst Estate | SO 9800 | Bsh 01 | M |
| Bucknell | SP 6545 | Bkl 07 | F | Wadenhoe | SP 9984 | Wdn 01 | F |
| Silverstone | SP 6644 | Sst 01 | M | | | | |
| Silverstone | SP 6644 | Sst 02 | M | | | | |
| Silverstone | SP 6644 | Sst 03 | M | | | | |

| Location | Grid Ref. | Individual number | Sex |
|---|---|---|---|
| BOWDEN PARK | | | |
| Bowden Park | ST 9368 | Bdp 01 | M |
| Bowden Park | ST 9368 | Bdp 02 | F |
| Bowden Park | ST 9368 | Bdp 03 | ? |
| Bowood Estate | ST 9769 | Bwe 01 | M |
| Bowood Estate | ST 9769 | Bwe 02 | F |
| Bowood Estate | ST 9769 | Bwe 03 | ? |
| Bowood Estate | ST 9769 | Bwe 04 | ? |
| Bowood Estate | ST 9769 | Bwe 05 | M |
| Spye Park | ST 9567 | Spp 01 | M |
| Spye Park | ST 9567 | Spp 02 | M |
| Spye Park | ST 9567 | Spp 03 | M |
| Spye Park | ST 9567 | Spp 04 | M |
| Spye Park | ST 9567 | Spp 05 | M |
| ELING COMMON | | | |
| Eling Common | SU 5274 | Elg 01 | M |
| Eling Common | SU 5274 | Elg 02 | M |
| Eling Common | SU 5274 | Elg 03 | M |
| Eling Common | SU 5274 | Elg 04 | M |
| Eling Common | SU 5274 | Elg 05 | M |
| Eling Common | SU 5274 | Elg 06 | M |
| Eling Common | SU 5274 | Elg 07 | F |
| Eling Common | SU 5274 | Elg 08 | F |
| Eling Common | SU 5274 | Elg 09 | F |
| Eling Common | SU 5274 | Elg 10 | F |
| Eling Common | SU 5274 | Elg 11 | F |
| Eling Common | SU 5274 | Elg 12 | F |
| Eling Common | SU 5274 | Elg 13 | F |
| Eling Common | SU 5274 | Elg 14 | F |
| Eling Common | SU 5274 | Elg 15 | F |
| Eling Common | SU 5274 | Elg 16 | ? |
| Eling Common | SU 5274 | Elg 17 | ? |
| Copyhold Farm | SU 4972 | Cpy 01 | M |
| Box Wood | SU 5273 | Bxw 01 | F |
| Common Firs | SU 5273 | Cmf 01 | M |
| Hermitage | SU 5-7- | Hmt 01 | M |
| Hermitage | SU 5-7- | Hmt 02 | M |
| Hermitage | SU 5-7- | Hmt 03 | M |
| Hermitage | SU 5-7- | Hmt 04 | M |
| Hermitage | SU 5-7- | Hmt 05 | F |
| Hermitage | SU 5-7- | Hmt 06 | F |
| Hermitage | SU 5-7- | Hmt 07 | ? |
| Hermitage | SU 5-7- | Hmt 08 | ? |
| Hermitage | SU 5-7- | Hmt 09 | M |
| Hermitage | SU 5-7- | Hmt 10 | ? |
| WORMSLEY ESTATE | | | |
| Wormsley Estate | SU 7394 | Wst 01 | M |
| Wormsley Estate | SU 7394 | Wst 02 | M |
| Wormsley Estate | SU 7394 | Wst 03 | M |
| Wormsley Estate | SU 7394 | Wst 04 | M |
| Wormsley Estate | SU 7394 | Wst 05 | M |
| Wormsley Estate | SU 7394 | Wst 06 | M |
| Wormsley Estate | SU 7394 | Wst 07 | M |
| Wormsley Estate | SU 7394 | Wst 08 | M |
| Wormsley Estate | SU 7394 | Wst 09 | M |
| Wormsley Estate | SU 7394 | Wst 10 | F |
| Wormsley Estate | SU 7394 | Wst 11 | F |
| Wormsley Estate | SU 7394 | Wst 12 | ? |
| Wormsley Estate | SU 7394 | Wst 13 | M |
| Wormsley Estate | SU 7394 | Wst 14 | F |
| OLD WARDEN | | | |
| Old Warden | TL 1343 | Owd 01 | M |
| Old Warden | TL 1343 | Owd 02 | M |
| Old Warden | TL 1343 | Owd 03 | M |
| Old Warden | TL 1343 | Owd 04 | M |
| Old Warden | TL 1343 | Owd 05 | M |
| Old Warden | TL 1343 | Owd 06 | F |
| Old Warden | TL 1343 | Owd 07 | F |
| Old Warden | TL 1343 | Owd 08 | F |
| Old Warden | TL 1343 | Owd 09 | F |
| Old Warden | TL 1343 | Owd 10 | F |
| Old Warden | TL 1343 | Owd 11 | F |
| Old Warden | TL 1343 | Owd 12 | F |
| Old Warden | TL 1343 | Owd 13 | F |
| Old Warden | TL 1343 | Owd 14 | F |
| Old Warden | TL 1343 | Owd 15 | M |
| Old Warden | TL 1343 | Owd 16 | M |
| Old Warden | TL 1343 | Owd 17 | M |
| Old Warden | TL 1343 | Owd 18 | M |
| Old Warden | TL 1343 | Owd 19 | M |
| Old Warden | TL 1343 | Owd 20 | M |
| NORTHILL | | | |
| Northill | TL 1446 | Nth 01 | M |
| Northill | TL 1446 | Nth 02 | M |
| Northill | TL 1446 | Nth 03 | M |
| Northill | TL 1446 | Nth 04 | M |
| Northill | TL 1446 | Nth 05 | F |
| Northill | TL 1446 | Nth 06 | F |
| Northill | TL 1446 | Nth 07 | F |
| Northill | TL 1446 | Nth 08 | F |
| Northill | TL 1446 | Nth 09 | F |
| Northill | TL 1446 | Nth 10 | F |
| Northill | TL 1446 | Nth 11 | M |
| Northill | TL 1446 | Nth 12 | M |
| Northill | TL 1446 | Nth 13 | M |
| Northill | TL 1446 | Nth 14 | M |
| Northill | TL 1446 | Nth 15 | F |
| Northill | TL 1446 | Nth 16 | M |
| Northill | TL 1446 | Nth 17 | M |
| Northill | TL 1446 | Nth 18 | M |
| Homewood Northill | TL 1446 | Hmw 01 | M |
| Homewood Northill | TL 1446 | Hmw 02 | M |
| Homewood Northill | TL 1446 | Hmw 03 | F |
| Homewood Northill | TL 1446 | Hmw 04 | F |
| Homewood Northill | TL 1446 | Hmw 05 | F |
| Homewood Northill | TL 1446 | Hmw 06 | F |
| Homewood Northill | TL 1446 | Hmw 07 | F |

| Location | Grid Ref. | Individual | Sex | Location | Grid Ref. | Individual | Sex |
|---|---|---|---|---|---|---|---|
| MINSBURY HILL | | | | Newton Wood | Tl 2222 | Ntw 02 | M |
| Minsbury Hill | TL 1527 | Min 01 | M | Newton Wood | Tl 2222 | Ntw 03 | M |
| Minsbury Hill | TL 1527 | Min 02 | M | Newton Wood | Tl 2222 | Ntw 04 | F |
| Minsbury Hill | TL 1527 | Min 03 | F | Newton Wood | Tl 2222 | Ntw 05 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 01 | M | Newton Wood | Tl 2222 | Ntw 06 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 02 | M | Newton Wood | Tl 2222 | Ntw 07 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 03 | F | Newton Wood | Tl 2222 | Ntw 08 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 04 | M | Newton Wood | Tl 2222 | Ntw 09 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 05 | F | Newton Wood | Tl 2222 | Ntw 10 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 06 | M | Newton Wood | Tl 2222 | Ntw 11 | M |
| Wain Wood Hitchin | TL 1725 | Wwd 07 | M | Newton Wood | Tl 2222 | Ntw 12 | M |
| Wain Wood Hitchin | TL 1725 | Wwd 08 | F | Newton Wood | Tl 2222 | Ntw 13 | F |
| Wain Wood Hitchin | TL 1725 | Wwd 09 | M | Newton Wood | Tl 2222 | Ntw 14 | M |
| Wain Wood Hitchin | TL 1725 | Wwd 10 | F | Watery Grove | TL 2322 | Wtg 01 | M |
| Westwood Hitchin | TL 1625 | Wwh 01 | M | Watery Grove | TL 2322 | Wtg 02 | M |
| Westwood Hitchin | TL 1625 | Wwh 02 | F | HEXTON | | | |
| Westwood Hitchin | TL 1625 | Wwh 03 | M | Hexton Estate | TL 3010 | Hex 01 | M |
| Westwood Hitchin | TL 1625 | Wwh 04 | F | Hexton Estate | TL 3010 | Hex 02 | M |
| BROCKET PARK | | | | Hexton Estate | TL 3010 | Hex 03 | M |
| Brocket Park | TL 2113 | Bkp 01 | F | Hexton Estate | TL 3010 | Hex 04 | M |
| Brocket Park | TL 2113 | Bkp 02 | F | Hexton Estate | TL 3010 | Hex 05 | F |
| Brocket Park | TL 2113 | Bkp 03 | F | Hexton Estate | TL 3010 | Hex 06 | F |
| Brocket Park | TL 2113 | Bkp 04 | F | Hexton Estate | TL 3010 | Hex 07 | F |
| Brocket Park | TL 2113 | Bkp 05 | F | Hexton Estate | TL 3010 | Hex 08 | M |
| Brocket Park | TL 2113 | Bkp 06 | F | Hexton Estate | TL 3010 | Hex 09 | M |
| Brocket Park | TL 2113 | Bkp 07 | M | Hexton Estate | TL 3010 | Hex 10 | M |
| Brocket Park | TL 2113 | Bkp 08 | M | Hexton Estate | TL 3010 | Hex 11 | M |
| Brocket Park | TL 2113 | Bkp 09 | M | Hexton Estate | TL 3010 | Hex 12 | F |
| Brocket Park | TL 2113 | Bkp 10 | M | Hexton Estate | TL 3010 | Hex 13 | F |
| Brocket Park | TL 2113 | Bkp 11 | M | Hexton Estate | TL 3010 | Hex 14 | F |
| Brocket Park | TL 2113 | Bkp 12 | M | Hexton Estate | TL 3010 | Hex 15 | M |
| Brocket Park | TL 2113 | Bkp 13 | M | Hexton Estate | TL 3010 | Hex 16 | M |
| Brocket Estate | TL 2113 | Bkt 01 | F | Hexton Estate | TL 3010 | Hex 17 | M |
| Brocket Estate | TL 2113 | Bkt 02 | ? | Hexton Estate | TL 3010 | Hex 18 | M |
| Brocket Estate | TL 2113 | Bkt 03 | ? | Hexton Estate | TL 3010 | Hex 19 | F |
| Brocket Estate | TL 2113 | Bkt 04 | ? | Hexton Estate | TL 3010 | Hex 20 | M |
| Brocket Estate | TL 2113 | Bkt 05 | ? | Hexton Estate | TL 3010 | Hex 21 | F |
| Brocket Estate | TL 2113 | Bkt 06 | ? | Hexton Estate | TL 3010 | Hex 22 | F |
| Brocket Estate | TL 2113 | Bkt 07 | ? | Hexton Estate | TL 3010 | Hex 23 | M |
| Brocket Estate | TL 2113 | Bkt 08 | ? | Hexton Estate | TL 3010 | Hex 24 | F |
| Brocket Estate | TL 2113 | Bkt 09 | ? | Hexton Estate | TL 3010 | Hex 25 | M |
| KNEBWORTH | | | | Hexton Estate | TL 3010 | Hex 26 | ? |
| Knebworth House | TL 2320 | Knb 01 | M | Hexton Estate | TL 3010 | Hex 27 | ? |
| Knebworth House | TL 2320 | Knb 02 | F | Hexton Estate | TL 3010 | Hex 28 | F |
| Knebworth House | TL 2320 | Knb 03 | M | Hexton Estate | TL 3010 | Hex 29 | F |
| Knebworth House | TL 2320 | Knb 04 | ? | BUNTINGFORD | | | |
| Knebworth House | TL 2320 | Knb 05 | M | Buntingford | TL 3629 | Btf 01 | M |
| Knebworth House | TL 2320 | Knb 06 | M | Buntingford | TL 3629 | Btf 02 | M |
| Knebworth House | TL 2320 | Knb 07 | F | Buntingford | TL 3629 | Btf 03 | M |
| Knebworth House | TL 2320 | Knb 08 | M | Buntingford | TL 3629 | Btf 04 | M |
| Knebworth House | TL 2320 | Knb 09 | F | Buntingford | TL 3629 | Btf 05 | F |
| Knebworth House | TL 2320 | Knb 10 | F | Buntingford | TL 3629 | Btf 06 | F |
| Knebworth House | TL 2320 | Knb 11 | ? | Buntingford | TL 3629 | Btf 07 | F |
| Graffridge Wood | TL 1212 | Grf 01 | M | Buntingford | TL 3629 | Btf 08 | M |
| Newton Wood | Tl 2222 | Ntw 01 | M | Buntingford | TL 3629 | Btf 09 | M |

| Location | Grid Ref. | Individual | Sex | Location | Grid Ref. | Individual | Sex |
|---|---|---|---|---|---|---|---|
| Buntingford | TL 3629 | Btf 10 | M | Euston | TL 9-7- | Eus 07 | M |
| Buntingford | TL 3629 | Btf 11 | M | Euston | TL 9-7- | Eus 08 | M |
| Buntingford | TL 3629 | Btf 12 | F | Euston | TL 9-7- | Eus 09 | F |
| Buntingford | TL 3629 | Btf 13 | M | Euston | TL 9-7- | Eus 10 | M |
| Buntingford | TL 3629 | Btf 14 | F | Euston | TL 9-7- | Eus 11 | M |
| Buntingford | TL 3629 | Btf 15 | ? | Euston | TL 9-7- | Eus 12 | M |
| Buntingford | TL 3629 | Btf 16 | M | Euston | TL 9-7- | Eus 13 | F |
| Buntingford | TL 3629 | Btf 17 | M | Euston | TL 9-7- | Eus 14 | F |
| Buntingford | TL 3629 | Btf 18 | F | Euston | TL 9-7- | Eus 15 | F |
| KING'S FOREST | | | | Euston | TL 9-7- | Eus 16 | F |
| Kings Forest | TL 8-7- | Kf 01 | M | Euston | TL 9-7- | Eus 17 | M |
| Kings Forest | TL 8-7- | Kf 02 | M | Euston | TL 9-7- | Eus 18 | M |
| Kings Forest | TL 8-7- | Kf 03 | M | Euston | TL 9-7- | Eus 19 | F |
| Kings Forest | TL 8-7- | Kf 04 | M | Euston | TL 9-7- | Eus 20 | F |
| Kings Forest | TL 8-7- | Kf 05 | M | Euston | TL 9-7- | Eus 21 | M |
| Kings Forest | TL 8-7- | Kf 06 | M | Euston | TL 9-7- | Eus 22 | M |
| Kings Forest | TL 8-7- | Kf 07 | M | Euston | TL 9-7- | Eus 23 | M |
| Kings Forest | TL 8-7- | Kf 08 | F | Euston | TL 9-7- | Eus 24 | M |
| Kings Forest | TL 8-7- | Kf 09 | F | Euston | TL 9-7- | Eus 25 | M |
| Kings Forest | TL 8-7- | Kf 10 | F | TAIWAN | | | |
| Kings Forest | TL 8-7- | Kf 11 | ? | Taiwan | | Twn 01 | F |
| Kings Forest | TL 8-7- | Kf 12 | ? | Taiwan | | Twn 02 | F |
| Kings Forest | TL 8-7- | Kf 13 | M | Taiwan | | Twn 03 | F |
| Kings Forest | TL835693 | Kf 14 | M | Taiwan | | Twn 04 | F |
| Kings Forest | TL 8-7- | Kf 15 | M | Taiwan | | Twn 05 | M |
| Kings Forest | TL 8171 | Kf 16 | M | Taiwan | | Twn 06 | M |
| Kings Forest | TL 8-7- | Kf 17 | M | Taiwan | | Twn 07 | F |
| Kings Forest | TL 8-7- | Kf 18 | M | Taiwan | | Twn 08 | M |
| Kings Forest | TL 8-7- | Kf 19 | F | Taiwan | | Twn 09 | F |
| Kings Forest | TL 8-7- | Kf 20 | M | Taiwan | | Twn 10 | M |
| Kings Forest | TL 8-7- | Kf 21 | F | Taiwan | | Twn 11 | F |
| Kings Forest | TL 8-7- | Kf 22 | M | Taiwan | | Twn 12 | M |
| Kings Forest | TL 8-7- | Kf 23 | M | Taiwan | | Twn 13 | M |
| Kings Forest | TL 8-7- | Kf 24 | F | Taiwan | | Twn 14 | F |
| Kings Forest | TL 8-7- | Kf 25 | M | Taiwan | | Twn 15 | M |
| Kings Forest | TL 8-7- | Kf 26 | F | Taiwan | | Twn 16 | ? |
| Kings Forest | TL 8-7- | Kf 27 | F | Taiwan | | Twn 17 | F |
| Kings Forest | TL 8-7- | Kf 28 | M | Taiwan | | Twn 18 | F |
| Kings Forest | TL 8-7- | Kf 29 | F | Taiwan | | Twn 19 | ? |
| Kings Forest | TL 8-7- | Kf 30 | F | Taiwan | | Twn 20 | M |
| Kings Forest | TL 8-7- | Kf 31 | M | Taiwan | | Twn 21 | ? |
| Kings Forest | TL 8-7- | Kf 32 | F | Taiwan | | Twn 22 | M |
| Kings Forest | TL 8-7- | Kf 33 | F | Taiwan | | Twn 23 | F |
| Kings Forest | TL 8-7- | Kf 34 | ? | Taiwan | | Twn 24 | M |
| Kings Forest | TL 8-7- | Kf 35 | M | Taiwan | | Twn 25 | M |
| EUSTON | | | | | | | |
| Euston | TL 9-7- | Eus 01 | M | | | | |
| Euston | TL 9-7- | Eus 02 | M | | | | |
| Euston | TL 9-7- | Eus 03 | M | | | | |
| Euston | TL 9-7- | Eus 04 | M | | | | |
| Euston | TL 9-7- | Eus 05 | M | | | | |
| Euston | TL 9-7- | Eus 06 | M | | | | |

# APPENDIX B

## Fingerprint Data

### Mean Band Sharing Coefficients Within and Between All Populations - Jeffreys' Probe 33·6

Knebworth: mean band sharing coefficient = 0·15 ± 0·09 (±59%)

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.17 | 0.2 | 0.17 | 0.17 | 0.11 | 0.21 | 0.05 | 0.22 | 0.17 | 0.18 | 0.11 | 0.23 |
| B | | | 0.23 | 0 | 0.06 | 0.18 | 0.23 | 0.05 | 0.06 | 0 | 0.13 | 0.06 | 0 |
| C | | | | 0.28 | 0.17 | 0.06 | 0.05 | 0.05 | 0.13 | 0.05 | 0.11 | 0.16 | 0.11 |
| D | | | | | 0.18 | 0.12 | 0.11 | 0.05 | 0.3 | 0.06 | 0.13 | 0.05 | 0.12 |
| E | | | | | | 0.25 | 0.28 | 0.16 | 0.24 | 0.18 | 0 | 0.05 | 0.12 |
| F | | | | | | | 0.23 | 0.27 | 0.18 | 0.12 | 0.06 | 0.05 | 0 |
| G | | | | | | | | 0.35 | 0.27 | 0.17 | 0.23 | 0.21 | 0.22 |
| H | | | | | | | | | 0.21 | 0.27 | 0.11 | 0.15 | 0.31 |
| I | | | | | | | | | | 0.06 | 0.18 | 0.16 | 0.18 |
| J | | | | | | | | | | | 0.26 | 0.11 | 0.24 |
| K | | | | | | | | | | | | 0.23 | 0.43 |
| L | | | | | | | | | | | | | 0.11 |

Brocket Park: mean band sharing coefficient = 0·331 ± 0·124 (±37%)

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.34 | 0.23 | 0.3 | 0.26 | 0.17 | 0.17 | 0.16 | 0.3 | 0.16 | 0.2 | 0.25 |
| B | | | 0.5 | 0.46 | 0.46 | 0.43 | 0.31 | 0.3 | 0.4 | 0.35 | 0.29 | 0.31 |
| C | | | | 0.46 | 0.34 | 0.5 | 0.37 | 0.24 | 0.26 | 0.29 | 0.11 | 0.21 |
| D | | | | | 0.42 | 0.4 | 0.2 | 0.25 | 0.21 | 0.25 | 0.06 | 0.22 |
| E | | | | | | 0.45 | 0.28 | 0.38 | 0.18 | 0.21 | 0.16 | 0.12 |
| F | | | | | | | 0.37 | 0.42 | 0.66 | 0.35 | 0.35 | 0.2 |
| G | | | | | | | | 0.36 | 0.53 | 0.47 | 0.53 | 0.41 |
| H | | | | | | | | | 0.45 | 0.34 | 0.34 | 0.4 |
| I | | | | | | | | | | 0.5 | 0.37 | 0.37 |
| J | | | | | | | | | | | 0.5 | 0.45 |
| K | | | | | | | | | | | | 0.52 |

Buntingford: mean band sharing coefficient = 0·253 ± 0·129 (±51%)

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.05 | 0.11 | 0.22 | 0.1 | 0.16 | 0.2 | 0.25 | 0.16 | 0.13 | 0.38 | 0.23 |
| B | | | 0.31 | 0.48 | 0.36 | 0.47 | 0.24 | 0.25 | 0.27 | 0.21 | 0.16 | 0.22 |
| C | | | | 0.36 | 0.42 | 0.35 | 0.16 | 0.35 | 0.24 | 0.21 | 0.12 | 0.18 |
| D | | | | | 0.53 | 0.42 | 0.22 | 0.22 | 0.12 | 0.37 | 0.31 | 0.26 |
| E | | | | | | 0.63 | 0.21 | 0.2 | 0.17 | 0.41 | 0.23 | 0.36 |
| F | | | | | | | 0.21 | 0.28 | 0.12 | 0.28 | 0.18 | 0.31 |
| G | | | | | | | | 0.13 | 0.16 | 0.13 | 0.05 | 0.05 |
| H | | | | | | | | | 0.11 | 0.18 | 0.22 | 0.07 |
| I | | | | | | | | | | 0.22 | 0.18 | 0.38 |
| J | | | | | | | | | | | 0.37 | 0.54 |
| K | | | | | | | | | | | | 0.51 |

## King's Forest: mean band sharing coefficient = 0·161 ± 0·141 (±88%)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.29 | 0.08 | 0.24 | 0.1 | 0.17 | 0 | 0.16 | 0 | 0.16 | 0.08 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0.24 |
| B | | | 0.21 | 0.28 | 0.48 | 0.14 | 0.28 | 0.21 | 0.06 | 0.07 | 0.13 | 0.13 | 0.06 | 0 | 0.13 | 0 | 0.12 | 0.07 |
| C | | | | 0.23 | 0.27 | 0.08 | 0.08 | 0.15 | 0 | 0.15 | 0.15 | 0.14 | 0.21 | 0 | 0.3 | 0.15 | 0.07 | 0.08 |
| D | | | | | 0.29 | 0.08 | 0 | 0.08 | 0.07 | 0.08 | 0.15 | 0.14 | 0.29 | 0.17 | 0.37 | 0.04 | 0 | 0.15 |
| E | | | | | | 0.1 | 0.27 | 0.09 | 0.08 | 0.09 | 0.17 | 0 | 0.08 | 0.11 | 0.26 | 0.18 | 0.08 | 0.18 |
| F | | | | | | | 0.24 | 0.16 | 0.07 | 0 | 0 | 0.15 | 0.07 | 0 | 0 | 0.08 | 0.07 | 0.02 |
| G | | | | | | | | 0.22 | 0.21 | 0.08 | 0.15 | 0.15 | 0.07 | 0.09 | 0 | 0.08 | 0.07 | 0.15 |
| H | | | | | | | | | 0.07 | 0.23 | 0.22 | 0.29 | 0.29 | 0.17 | 0.07 | 0.15 | 0.13 | 0.38 |
| I | | | | | | | | | | 0.14 | 0.13 | 0.19 | 0.13 | 0.08 | 0.07 | 0.14 | 0.12 | 0 |
| J | | | | | | | | | | | 0.59 | 0.43 | 0.36 | 0.09 | 0 | 0.08 | 0.13 | 0.23 |
| K | | | | | | | | | | | | 0.41 | 0.41 | 0.08 | 0 | 0.07 | 0.19 | 0.15 |
| L | | | | | | | | | | | | | 0.53 | 0.08 | 0 | 0.14 | 0.19 | 0.29 |
| M | | | | | | | | | | | | | | 0.16 | 0.41 | 0.36 | 0.38 | 0.36 |
| N | | | | | | | | | | | | | | | 0.25 | 0.26 | 0.15 | 0.17 |
| O | | | | | | | | | | | | | | | | 0.52 | 0.91 | 0.22 |
| P | | | | | | | | | | | | | | | | | 0.4 | 0.54 |
| Q | | | | | | | | | | | | | | | | | | 0.27 |

## Mean band sharing coefficient, Knb/Bkp = 0·129 ± 0·078 (±60%)

Knb Bkp

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.16 | 0.29 | 0.35 | 0.18 | 0.27 | 0.23 | 0.23 | 0.34 | 0.25 | 0.11 | 0.16 | 0.26 |
| B | 0.11 | 0.18 | 0.31 | 0.13 | 0.34 | 0.18 | 0.18 | 0.18 | 0.13 | 0.11 | 0.11 | 0.13 |
| C | 0.15 | 0.28 | 0.11 | 0.12 | 0.26 | 0.17 | 0.17 | 0.22 | 0.12 | 0.16 | 0.32 | 0.31 |
| D | 0.11 | 0.18 | 0.12 | 0.13 | 0.11 | 0.12 | 0.18 | 0.06 | 0.06 | 0.11 | 0.11 | 0.13 |
| E | 0.05 | 0.18 | 0.06 | 0 | 0.05 | 0.12 | 0.18 | 0.24 | 0.13 | 0.17 | 0.23 | 0.13 |
| F | 0.17 | 0.12 | 0.06 | 0.13 | 0.11 | 0.12 | 0.06 | 0.12 | 0.06 | 0.05 | 0.05 | 0.13 |
| G | 0.26 | 0.05 | 0.17 | 0.12 | 0.1 | 0.11 | 0.05 | 0.11 | 0.06 | 0.11 | 0.05 | 0.06 |
| H | 0.1 | 0 | 0.05 | 0.05 | 0.1 | 0 | 0.05 | 0.05 | 0.05 | 0.15 | 0.1 | 0.23 |
| I | 0.05 | 0.18 | 0.12 | 0.13 | 0.16 | 0.24 | 0.18 | 0.11 | 0.25 | 0.05 | 0 | 0.06 |
| J | 0.05 | 0.06 | 0 | 0.06 | 0.11 | 0.12 | 0.06 | 0.18 | 0.13 | 0.23 | 0.05 | 0.06 |
| K | 0.05 | 0.13 | 0.19 | 0.06 | 0.11 | 0.19 | 0.06 | 0.06 | 0.06 | 0 | 0.06 | 0.07 |
| L | 0.05 | 0.05 | 0.05 | 0.12 | 0.26 | 0.11 | 0.05 | 0.11 | 0.06 | 0.21 | 0 | 0.06 |
| M | 0 | 0.06 | 0.12 | 0.13 | 0.16 | 0.18 | 0.18 | 0.17 | 0.13 | 0.23 | 0.17 | 0.13 |

## Mean band sharing coefficient, Knb/Btf = 0·138 ± 0·081 (±59%)

Knb Btf

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.21 | 0.15 | 0.05 | 0.06 | 0.16 | 0.11 | 0.1 | 0.14 | 0.11 | 0.07 | 0.17 | 0.12 |
| B | 0 | 0.05 | 0.06 | 0.06 | 0.23 | 0.18 | 0.11 | 0.07 | 0 | 0 | 0 | 0.06 |
| C | 0.2 | 0.11 | 0.05 | 0.11 | 0.05 | 0.05 | 0.1 | 0.13 | 0.22 | 0.2 | 0.23 | 0.23 |
| D | 0.22 | 0.11 | 0.06 | 0.12 | 0.11 | 0.18 | 0.05 | 0.05 | 0.06 | 0.07 | 0.25 | 0.19 |
| E | 0.16 | 0.16 | 0.12 | 0.12 | 0.06 | 0.12 | 0 | 0.22 | 0.25 | 0.15 | 0.25 | 0.38 |
| F | 0 | 0.21 | 0.18 | 0.06 | 0.11 | 0.12 | 0.05 | 0.3 | 0.19 | 0.07 | 0.12 | 0.06 |
| G | 0.05 | 0.05 | 0.22 | 0.11 | 0.21 | 0.11 | 0.1 | 0.06 | 0.11 | 0.2 | 0.05 | 0.17 |
| H | 0 | 0.19 | 0.15 | 0.21 | 0.2 | 0.26 | 0.24 | 0.12 | 0.05 | 0.25 | 0.11 | 0.11 |
| I | 0.16 | 0.16 | 0.17 | 0.18 | 0.28 | 0.23 | 0.16 | 0.21 | 0.12 | 0.21 | 0.24 | 0.31 |
| J | 0.16 | 0.1 | 0.12 | 0.06 | 0 | 0.06 | 0.22 | 0.11 | 0.25 | 0.07 | 0.18 | 0.13 |
| K | 0.11 | 0.33 | 0.12 | 0.19 | 0.36 | 0.25 | 0.23 | 0.15 | 0.06 | 0.15 | 0.13 | 0.13 |
| L | 0.15 | 0.05 | 0 | 0.05 | 0.05 | 0.05 | 0.05 | 0.13 | 0.17 | 0.13 | 0.11 | 0.11 |
| M | 0.16 | 0.37 | 0.06 | 0.24 | 0.23 | 0.17 | 0.27 | 0.07 | 0.18 | 0.14 | 0.18 | 0.12 |

Mean band sharing coefficient, Knb/Kf = 0·094 ± 0·089 (±95%)

Knb

Kf

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.25 | 0 | 0.08 | 0.16 | 0.1 | 0 | 0 | 0 | 0 | 0.02 | 0.08 | 0.08 | 0.07 | 0.18 | 0.23 | 0.24 | 0.14 | 0.24 |
| B | 0 | 0.17 | 0 | 0 | 0 | 0.1 | 0 | 0.1 | 0.08 | 0.1 | 0 | 0.09 | 0 | 0 | 0 | 0.1 | 0.24 | 0.19 |
| C | 0 | 0.15 | 0.09 | 0.17 | 0.11 | 0.09 | 0.26 | 0 | 0.15 | 0.09 | 0.08 | 0.15 | 0.24 | 0.1 | 0.17 | 0.26 | 0.15 | 0.26 |
| D | 0.09 | 0.08 | 0 | 0.17 | 0 | 0.18 | 0 | 0.17 | 0.08 | 0.81 | 0 | 0 | 0 | 0.1 | 0.08 | 0.09 | 0.07 | 0 |
| E | 0.1 | 0 | 0.25 | 0.25 | 0.1 | 0.09 | 0 | 0.08 | 0.22 | 0.24 | 0.24 | 0.25 | 0.15 | 0.1 | 0.08 | 0.08 | 0.07 | 0.08 |
| F | 0.17 | 0 | 0.08 | 0.08 | 0 | 0 | 0 | 0 | 0.07 | 0.08 | 0.08 | 0 | 0 | 0.1 | 0.08 | 0.08 | 0 | 0 |
| G | 0.16 | 0.21 | 0.08 | 0.15 | 0.1 | 0.08 | 0.15 | 0 | 0.07 | 0.08 | 0.07 | 0.07 | 0 | 0 | 0.07 | 0 | 0.13 | 0 |
| H | 0 | 0.19 | 0.19 | 0.14 | 0.08 | 0.07 | 0.14 | 0.07 | 0 | 0.14 | 0.14 | 0.13 | 0.13 | 0.16 | 0.07 | 0.07 | 0.06 | 0.07 |
| I | 0.09 | 0 | 0.09 | 0.09 | 0 | 0.09 | 0 | 0 | 0.15 | 0 | 0.08 | 0.08 | 0 | 0 | 0.08 | 0.17 | 0.07 | 0.09 |
| J | 0.16 | 0 | 0.15 | 0.15 | 0.09 | 0 | 0.15 | 0.08 | 0.07 | 0.15 | 0.22 | 0.07 | 0.14 | 0.09 | 0.07 | 0.33 | 0.07 | 0.23 |
| K | 0.09 | 0.07 | 0 | 0 | 0 | 0 | 0.17 | 0.17 | 0.15 | 0.08 | 0.08 | 0.08 | 0 | 0 | 0 | 0.08 | 0.14 | 0.08 |
| L | 0.19 | 0.08 | 0.18 | 0.09 | 0.11 | 0 | 0.18 | 0.18 | 0.16 | 0.18 | 0.17 | 0.08 | 0.08 | 0 | 0 | 0 | 0 | 0 |
| M | 0.09 | 0 | 0.17 | 0.08 | 0.1 | 0.17 | 0.17 | 0.25 | 0.07 | 0 | 0.16 | 0.08 | 0.08 | 0.1 | 0.16 | 0.17 | 0 | 0.17 |

Mean band sharing coefficient, Bkp/Btf = 0·144 ± 0·075 (±53%)

Bkp

Btf

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.1 | 0 | 0.16 | 0.17 | 0.27 | 0.11 | 0.1 | 0.13 | 0.17 | 0.06 | 0.17 | 0.17 |
| B | 0.16 | 0.21 | 0.06 | 0.12 | 0.11 | 0.12 | 0.16 | 0.05 | 0.12 | 0.05 | 0.19 | 0.32 |
| C | 0.27 | 0.16 | 0.12 | 0.12 | 0.17 | 0.12 | 0.27 | 0.11 | 0.12 | 0.05 | 0.12 | 0.19 |
| D | 0.17 | 0.23 | 0.06 | 0.06 | 0.18 | 0.13 | 0.17 | 0.11 | 0.2 | 0 | 0.13 | 0.13 |
| E | 0.2 | 0.2 | 0.05 | 0.23 | 0.21 | 0.22 | 0.15 | 0.06 | 0.06 | 0.06 | 0.23 | 0.29 |
| F | 0.27 | 0.27 | 0.18 | 0.12 | 0.17 | 0.12 | 0.22 | 0.3 | 0.31 | 0 | 0.18 | 0.32 |
| G | 0.08 | 0.11 | 0 | 0.18 | 0.23 | 0.18 | 0.16 | 0.22 | 0.25 | 0 | 0.06 | 0.06 |
| H | 0.1 | 0.15 | 0.06 | 0.18 | 0.23 | 0.17 | 0.11 | 0.21 | 0.18 | 0.07 | 0.06 | 0.06 |
| I | 0.11 | 0.17 | 0.06 | 0.2 | 0.12 | 0.06 | 0.11 | 0.16 | 0.26 | 0 | 0.2 | 0.14 |
| J | 0.05 | 0.1 | 0.05 | 0.06 | 0.05 | 0.17 | 0.21 | 0 | 0.23 | 0.13 | 0.23 | 0.18 |
| K | 0.15 | 0.15 | 0.06 | 0.12 | 0.16 | 0.17 | 0.1 | 0.14 | 0.35 | 0.14 | 0.17 | 0.24 |
| L | 0.06 | 0.17 | 0 | 0.2 | 0.13 | 0.13 | 0.18 | 0.16 | 0.2 | 0.16 | 0.07 | 0.07 |

Mean band sharing coefficient, Bkp/Kf = 0·161 ± 0·105 (±65%)

Bkp

Kf

| A | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| B | 0.07 | 0.26 | 0.14 | 0.14 | 0.08 | 0.15 | 0.14 | 0 | 0.13 | 0.02 | 0.21 | 0.02 | 0.2 | 0.16 | 0.14 | 0.07 | 0.25 | 0.07 |
| C | 0.26 | 0 | 0.08 | 0.08 | 0 | 0.09 | 0.08 | 0.08 | 0.22 | 0.25 | 0.32 | 0.38 | 0.31 | 0.19 | 0.16 | 0.17 | 0.21 | 0.17 |
| D | 0.17 | 0.12 | 0.25 | 0.08 | 0.2 | 0.09 | 0.08 | 0 | 0.07 | 0.17 | 0.08 | 0.23 | 0.15 | 0.19 | 0.32 | 0.25 | 0.21 | 0.25 |
| E | 0.29 | 0.08 | 0 | 0.09 | 0.11 | 0 | 0.18 | 0.18 | 0.16 | 0.18 | 0.17 | 0.17 | 0.08 | 0.11 | 0.09 | 0.09 | 0.08 | 0.18 |
| F | 0.17 | 0 | 0.08 | 0.25 | 0 | 0 | 0.24 | 0 | 0 | 0.17 | 0.16 | 0.15 | 0.23 | 0.1 | 0.16 | 0.17 | 0.07 | 0.08 |
| G | 0.25 | 0.21 | 0.16 | 0.08 | 0.19 | 0.17 | 0.16 | 0.08 | 0.07 | 0 | 0.08 | 0.22 | 0.15 | 0.18 | 0.23 | 0.24 | 0.14 | 0.16 |
| H | 0.25 | 0.07 | 0.24 | 0.08 | 0 | 0 | 0.08 | 0.16 | 0.07 | 0.32 | 0.15 | 0.44 | 0.3 | 0.09 | 0.23 | 0.16 | 0.07 | 0.16 |
| I | 0.23 | 0.13 | 0.15 | 0.22 | 0.26 | 0 | 0.22 | 0.15 | 0 | 0.22 | 0.36 | 0.15 | 0.28 | 0.17 | 0.21 | 0.3 | 0.32 | 0.22 |
| J | 0.09 | 0 | 0.17 | 0.09 | 0 | 0 | 0.09 | 0.09 | 0.08 | 0.26 | 0.42 | 0.32 | 0.16 | 0.2 | 0.17 | 0.26 | 0.07 | 0.09 |
| K | 0.15 | 0.06 | 0.29 | 0.14 | 0.17 | 0.15 | 0.21 | 0.21 | 0.06 | 0.21 | 0.41 | 0.33 | 0.4 | 0.24 | 0.28 | 0.43 | 0.19 | 0.29 |
| L | 0 | 0.07 | 0.32 | 0.24 | 0.19 | 0.08 | 0.16 | 0.16 | 0.07 | 0.16 | 0.15 | 0.3 | 0.44 | 0.27 | 0.38 | 0.4 | 0.14 | 0.24 |
| M | 0.11 | 0 | 0.1 | 0 | 0 | 0 | 0 | 0.1 | 0 | 0.4 | 0.29 | 0.27 | 0.36 | 0.24 | 0.1 | 0.2 | 0.26 | 0.2 |

Mean bandsharing coefficient, Btf/Kf = 0·128 ± 0·093 (±73%)

Btf Kf

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.17 | 0.07 | 0.17 | 0.08 | 0.2 | 0.26 | 0.13 | 0 | 0.3 | 0.08 | 0 | 0.15 | 0.15 | 0.1 | 0.16 | 0.08 | 0 | 0.25 |
| B | 0.23 | 0.07 | 0.15 | 0.3 | 0.09 | 0 | 0.07 | 0.22 | 0.07 | 0.07 | 0.07 | 0.07 | 0.21 | 0.25 | 0.29 | 0.22 | 0 | 0.07 |
| C | 0.26 | 0.15 | 0.33 | 0.33 | 0.1 | 0.17 | 0.08 | 0.08 | 0 | 0 | 0 | 0.15 | 0.23 | 0 | 0.16 | 0.25 | 0 | 0.17 |
| D | 0.1 | 0.08 | 0.1 | 0.19 | 0 | 0 | 0 | 0.1 | 0 | 0.29 | 0.17 | 0.35 | 0.35 | 0.17 | 0.18 | 0.19 | 0.09 | 0.19 |
| E | 0.18 | 0.08 | 0 | 0.26 | 0 | 0.09 | 0 | 0 | 0 | 0.17 | 0.08 | 0.16 | 0.16 | 0.1 | 0 | 0.09 | 0.15 | 0 |
| F | 0.08 | 0.07 | 0.08 | 0.24 | 0 | 0.08 | 0.08 | 0.08 | 0.14 | 0.16 | 0.15 | 0.3 | 0.4 | 0.09 | 0.15 | 0.24 | 0.14 | 0.16 |
| G | 0.25 | 0.29 | 0.16 | 0.08 | 0.29 | 0.17 | 0.16 | 0 | 0.07 | 0.08 | 0.08 | 0 | 0.07 | 0.18 | 0.08 | 0.16 | 0.14 | 0.24 |
| H | 0.05 | 0 | 0.1 | 0.2 | 0 | 0 | 0 | 0 | 0.09 | 0 | 0 | 0.09 | 0.18 | 0.12 | 0.1 | 0.1 | 0 | 0.1 |
| I | 0.17 | 0.07 | 0.25 | 0.17 | 0.2 | 0.09 | 0.17 | 0.08 | 0.15 | 0.08 | 0.08 | 0.15 | 0.15 | 0 | 0.32 | 0.25 | 0 | 0.08 |
| J | 0.08 | 0 | 0 | 0.1 | 0 | 0.1 | 0.09 | 0 | 0.08 | 0.09 | 0.09 | 0.08 | 0.25 | 0.21 | 0.17 | 0.18 | 0.15 | 0.27 |
| K | 0.19 | 0.08 | 0.18 | 0.27 | 0.11 | 0.1 | 0.08 | 0.09 | 0.16 | 0.18 | 0.17 | 0.25 | 0.17 | 0 | 0.17 | 0.18 | 0.08 | 0.18 |
| L | 0.09 | 0.08 | 0.17 | 0.35 | 0.21 | 0.18 | 0.09 | 0 | 0.23 | 0.09 | 0.17 | 0.32 | 0.24 | 0 | 0.25 | 0.26 | 0.15 | 0.17 |

## Mean Band Sharing Coefficients Within and Between All Populations - Jeffreys' Probe 33·15

Knebworth: mean bandsharing coefficient = 0·187± 0·081 (±43%)

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.25 | 0.18 | 0.15 | 0.21 | 0.2 | 0.18 | 0.33 | 0.2 | 0.12 | 0.2 | 0.12 | 0.27 |
| B | | | 0.15 | 0.04 | 0.17 | 0.22 | 0.03 | 0.15 | 0.13 | 0.09 | 0.13 | 0.08 | 0.25 |
| C | | | | 0.3 | 0.15 | 0.09 | 0.19 | 0.12 | 0.17 | 0.18 | 0.16 | 0.22 | 0.17 |
| D | | | | | 0.16 | 0.22 | 0.34 | 0.04 | 0.04 | 0.15 | 0.25 | 0.15 | 0.04 |
| E | | | | | | 0.3 | 0.26 | 0.26 | 0.25 | 0.17 | 0.21 | 0.17 | 0.16 |
| F | | | | | | | 0.17 | 0.18 | 0.2 | 0.1 | 0.21 | 0.21 | 0.2 |
| G | | | | | | | | 0.27 | 0.26 | 0.18 | 0.13 | 0.13 | 0.17 |
| H | | | | | | | | | 0.17 | 0.23 | 0.13 | 0.18 | 0.34 |
| I | | | | | | | | | | 0.1 | 0.1 | 0.2 | 0.1 |
| K | | | | | | | | | | | 0.37 | 0.16 | 0.25 |
| J | | | | | | | | | | | | 0.36 | 0.4 |
| L | | | | | | | | | | | | | 0.35 |

Brocket Park: mean bandsharing coefficient = 0·301± 0·110 (±36%)

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.36 | 0.45 | 0.28 | 0.32 | 0.29 | 0.25 | 0.32 | 0.22 | 0.15 | 0.2 | 0.3 |
| B | | | 0.38 | 0.47 | 0.36 | 0.43 | 0.23 | 0.25 | 0.25 | 0.11 | 0.22 | 0.34 |
| C | | | | 0.4 | 0.4 | 0.46 | 0.32 | 0.61 | 0.35 | 0.38 | 0.26 | 0.43 |
| D | | | | | 0.48 | 0.43 | 0.2 | 0.18 | 0.16 | 0.15 | 0.2 | 0.41 |
| E | | | | | | 0.16 | 0.17 | 0.26 | 0.19 | 0.24 | 0.22 | 0.47 |
| F | | | | | | | 0.31 | 0.31 | 0.28 | 0.21 | 0.1 | 0.21 |
| G | | | | | | | | 0.3 | 0.36 | 0.22 | 0.11 | 0.27 |
| H | | | | | | | | | 0.48 | 0.46 | 0.29 | 0.4 |
| I | | | | | | | | | | 0.37 | 0.23 | 0.24 |
| J | | | | | | | | | | | 0.22 | 0.34 |
| K | | | | | | | | | | | | 0.43 |

Buntingford: mean bandsharing coefficient = 0·237 ± 0·076 (±32%)

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.27 | 0.25 | 0.19 | 0.28 | 0.28 | 0.23 | 0.29 | 0.26 | 0.22 | 0.22 | 0.32 |
| B | | | 0.39 | 0.27 | 0.32 | 0.29 | 0.35 | 0.15 | 0.27 | 0.33 | 0.19 | 0.22 |
| C | | | | 0.15 | 0.23 | 0.3 | 0.35 | 0.17 | 0.14 | 0.21 | 0.14 | 0.06 |
| D | | | | | 0.24 | 0.28 | 0.2 | 0.21 | 0.14 | 0.06 | 0.21 | 0.17 |
| E | | | | | | 0.3 | 0.21 | 0.23 | 0.2 | 0.24 | 0.19 | 0.19 |
| F | | | | | | | 0.36 | 0.23 | 0.24 | 0.21 | 0.24 | 0.15 |
| G | | | | | | | | 0.22 | 0.19 | 0.2 | 0.15 | 0.07 |
| H | | | | | | | | | 0.33 | 0.28 | 0.32 | 0.16 |
| I | | | | | | | | | | 0.37 | 0.29 | 0.25 |
| J | | | | | | | | | | | 0.32 | 0.21 |
| K | | | | | | | | | | | | 0.44 |

King's Forest: mean bandsharing coefficient = 0·255 ± 0·096 (±37%)

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | | 0.53 | 0.21 | 0.11 | 0.21 | 0.22 | 0.32 | 0.27 | 0.35 | 0.34 | 0.37 | 0.25 | 0.19 | 0.31 | 0.27 | 0.34 |
| B | | | 0.29 | 0.11 | 0.11 | 0.15 | 0.33 | 0.25 | 0.25 | 0.2 | 0.28 | 0.33 | 0.19 | 0.28 | 0.28 | 0.22 |
| C | | | | 0.27 | 0.18 | 0.19 | 0.18 | 0.14 | 0.22 | 0.32 | 0.35 | 0.26 | 0.23 | 0.28 | 0.24 | 0.22 |
| D | | | | | 0.19 | 0.24 | 0.15 | 0.22 | 0.18 | 0.1 | 0.19 | 0.2 | 0.16 | 0.21 | 0.13 | 0.16 |
| E | | | | | | 0.3 | 0.29 | 0.21 | 0.22 | 0.16 | 0.25 | 0.18 | 0.31 | 0.12 | 0.04 | 0.17 |
| F | | | | | | | 0.26 | 0.36 | 0.14 | 0.2 | 0.29 | 0.3 | 0.28 | 0.2 | 0.25 | 0.2 |
| G | | | | | | | | 0.39 | 0.25 | 0.28 | 0.28 | 0.3 | 0.27 | 0.12 | 0.04 | 0.13 |
| H | | | | | | | | | 0.21 | 0.23 | 0.24 | 0.29 | 0.18 | 0.11 | 0.39 | 0.17 |
| I | | | | | | | | | | 0.27 | 0.31 | 0.25 | 0.23 | 0.32 | 0.2 | 0.3 |
| J | | | | | | | | | | | 0.42 | 0.28 | 0.29 | 0.17 | 0.31 | 0.14 |
| K | | | | | | | | | | | | 0.51 | 0.41 | 0.51 | 0.35 | 0.29 |
| L | | | | | | | | | | | | | 0.4 | 0.45 | 0.33 | 0.41 |
| M | | | | | | | | | | | | | | 0.34 | 0.26 | 0.19 |
| N | | | | | | | | | | | | | | | 0.36 | 0.4 |
| O | | | | | | | | | | | | | | | | 0.55 |

Mean band sharing coefficient, Knb/Bkp = 0·126 ± 0·071 (±70%)

Knb Bkp

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.27 | 0.12 | 0.08 | 0.15 | 0.17 | 0.16 | 0.16 | 0.03 | 0.13 | 0.12 | 0.04 | 0.08 |
| B | 0.21 | 0.23 | 0.31 | 0.17 | 0.19 | 0.21 | 0.22 | 0.29 | 0.19 | 0.23 | 0.26 | 0.32 |
| C | 0.3 | 0.14 | 0.13 | 0.13 | 0.09 | 0.22 | 0.13 | 0.21 | 0.15 | 0.14 | 0.09 | 0.13 |
| D | 0.14 | 0 | 0.05 | 0.04 | 0.08 | 0.09 | 0.16 | 0.04 | 0.04 | 0.15 | 0.2 | 0.05 |
| E | 0.04 | 0.22 | 0.17 | 0.16 | 0.13 | 0.17 | 0.09 | 0.16 | 0.14 | 0.13 | 0.25 | 0.17 |
| F | 0.1 | 0.16 | 0.05 | 0.05 | 0.11 | 0.15 | 0.21 | 0.09 | 0.11 | 0.11 | 0.1 | 0.05 |
| G | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.08 | 0.09 | 0.12 | 0.14 | 0.27 | 0.08 | 0.13 |
| H | 0.17 | 0.14 | 0.04 | 0.17 | 0.04 | 0.08 | 0.13 | 0.08 | 0.2 | 0.14 | 0.09 | 0.18 |
| I | 0.19 | 0.21 | 0.2 | 0.14 | 0.11 | 0.25 | 0.21 | 0.18 | 0.11 | 0.05 | 0.1 | 0.1 |
| J | 0 | 0.11 | 0.05 | 0.05 | 0.11 | 0.05 | 0.16 | 0.1 | 0.06 | 0.11 | 0.05 | 0.05 |
| K | 0.15 | 0.11 | 0.05 | 0 | 0.17 | 0.05 | 0.1 | 0.09 | 0.17 | 0.11 | 0.15 | 0 |
| L | 0.05 | 0.05 | 0.05 | 0 | 0.17 | 0.15 | 0.05 | 0.14 | 0.11 | 0.11 | 0.26 | 0.16 |
| M | 0.19 | 0.05 | 0.25 | 0.04 | 0.21 | 0.19 | 0.1 | 0.23 | 0 | 0.15 | 0.05 | 0.15 |

Mean band sharing coefficient, Knb/Btf = 0·141± 0·068 (±48%)

Knb Btf

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.33 | 0.2 | 0.15 | 0.09 | 0.14 | 0.25 | 0.03 | 0.03 | 0.15 | 0.19 | 0.11 | 0.18 |
| B | 0.28 | 0.18 | 0.06 | 0.21 | 0.18 | 0.26 | 0.21 | 0.11 | 0.16 | 0.24 | 0.2 | 0.27 |
| C | 0.18 | 0.15 | 0.2 | 0.21 | 0.23 | 0.11 | 0.14 | 0.16 | 0.04 | 0.21 | 0.16 | 0.12 |
| D | 0.07 | 0.08 | 0.07 | 0.19 | 0.12 | 0.12 | 0.16 | 0.04 | 0 | 0 | 0.13 | 0.04 |
| E | 0.14 | 0.25 | 0.16 | 0.02 | 0.14 | 0.11 | 0.1 | 0.19 | 0.16 | 0.1 | 0.11 | 0.15 |
| F | 0.32 | 0.17 | 0.11 | 0.24 | 0.17 | 0.22 | 0.12 | 0.13 | 0.04 | 0.12 | 0.18 | 0.22 |
| G | 0.07 | 0.1 | 0.06 | 0.07 | 0.15 | 0.11 | 0.07 | 0.04 | 0.08 | 0.1 | 0.04 | 0.03 |
| H | 0.14 | 0.18 | 0.2 | 0.21 | 0.15 | 0.15 | 0.11 | 0.12 | 0.16 | 0.14 | 0.16 | 0.16 |
| I | 0.1 | 0.12 | 0.33 | 0.11 | 0.34 | 0.17 | 0.2 | 0.17 | 0.09 | 0.15 | 0.13 | 0 |
| J | 0.08 | 0.13 | 0.07 | 0.08 | 0.08 | 0.13 | 0.08 | 0.13 | 0.14 | 0.12 | 0.14 | 0.13 |
| K | 0.16 | 0.08 | 0.15 | 0.2 | 0.17 | 0.02 | 0.04 | 0.13 | 0.09 | 0.12 | 0.23 | 0.13 |
| L | 0.28 | 0.25 | 0.15 | 0.12 | 0.17 | 0.08 | 0.12 | 0.22 | 0.19 | 0.2 | 0.04 | 0.04 |
| M | 0.11 | 0.24 | 0.14 | 0.15 | 0.08 | 0.12 | 0.16 | 0.21 | 0.18 | 0.15 | 0.13 | 0.08 |

Mean band sharing coefficient, Knb/Kf = 0·141± 0·069 (±49%)

Knb Kf

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.23 | 0.24 | 0.24 | 0.18 | 0.2 | 0.25 | 0.31 | 0.13 | 0.17 | 0.18 | 0.27 | 0.21 | 0.22 | 0.19 | 0.11 | 0.16 |
| B | 0.18 | 0.22 | 0.26 | 0.16 | 0.15 | 0.15 | 0.15 | 0.21 | 0.18 | 0.12 | 0.1 | 0.23 | 0.16 | 0.16 | 0.25 | 0.27 |
| C | 0.07 | 0.11 | 0.19 | 0.2 | 0.07 | 0.19 | 0.15 | 0.18 | 0.03 | 0.25 | 0.18 | 0.19 | 0.16 | 0.17 | 0.12 | 0.13 |
| D | 0.08 | 0.12 | 0.08 | 0.13 | 0.08 | 0.04 | 0.16 | 0.16 | 0.2 | 0.27 | 0.08 | 0.08 | 0.08 | 0.09 | 0.09 | 0 |
| E | 0.25 | 0.14 | 0.18 | 0.07 | 0.29 | 0.26 | 0.14 | 0.1 | 0.21 | 0.2 | 0.25 | 0.07 | 0.15 | 0.08 | 0.12 | 0.08 |
| F | 0.08 | 0.13 | 0 | 0.09 | 0.04 | 0.08 | 0.04 | 0.08 | 0.17 | 0.04 | 0.12 | 0.08 | 0.04 | 0.04 | 0.14 | 0.05 |
| G | 0.07 | 0.03 | 0.15 | 0.16 | 0.11 | 0.07 | 0.11 | 0.11 | 0.14 | 0.36 | 0.14 | 0.07 | 0.2 | 0.08 | 0.12 | 0.04 |
| H | 0.22 | 0.26 | 0.15 | 0.16 | 0.19 | 0.11 | 0.11 | 0.11 | 0.26 | 0.12 | 0.25 | 0.15 | 0.2 | 0.17 | 0.17 | 0.18 |
| I | 0.12 | 0.08 | 0.17 | 0.18 | 0.12 | 0.21 | 0.17 | 0.2 | 0.2 | 0.04 | 0.16 | 0.04 | 0.18 | 0.04 | 0.09 | 0.05 |
| J | 0.08 | 0.13 | 0.07 | 0.14 | 0.08 | 0.18 | 0.08 | 0 | 0.12 | 0.14 | 0.21 | 0.18 | 0.19 | 0.2 | 0.1 | 0.05 |
| K | 0.04 | 0.13 | 0.13 | 0.09 | 0.13 | 0.17 | 0.08 | 0.16 | 0.12 | 0.14 | 0.16 | 0.08 | 0.09 | 0.14 | 0.19 | 0.05 |
| L | 0.08 | 0.08 | 0.13 | 0.13 | 0.08 | 0.22 | 0.08 | 0.08 | 0.12 | 0.09 | 0.2 | 0.13 | 0.04 | 0.19 | 0.04 | 0.05 |
| M | 0.24 | 0.29 | 0.2 | 0.08 | 0.12 | 0.17 | 0.12 | 0.08 | 0.12 | 0.22 | 0.36 | 0.12 | 0.17 | 0.23 | 0.13 | 0.15 |

Mean band sharing coefficient, Bkp/Btf = 0·154 ± 0·065 (±42%)

Bkp Btf

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.19 | 0.16 | 0.22 | 0.19 | 0.08 | 0.12 | 0.28 | 0.13 | 0.09 | 0.19 | 0.08 | 0.17 |
| B | 0.17 | 0.17 | 0.2 | 0.16 | 0.22 | 0.13 | 0.17 | 0.23 | 0.1 | 0.12 | 0.14 | 0.14 |
| C | 0.2 | 0.17 | 0.11 | 0.08 | 0.17 | 0.31 | 0.2 | 0.27 | 0.09 | 0.16 | 0.04 | 0.13 |
| D | 0.07 | 0.16 | 0.29 | 0.11 | 0.12 | 0.17 | 0.2 | 0.26 | 0.13 | 0.26 | 0.13 | 0.13 |
| E | 0.21 | 0.04 | 0.12 | 0.08 | 0.04 | 0.09 | 0.04 | 0.19 | 0.2 | 0.12 | 0.1 | 0.09 |
| F | 0.24 | 0.25 | 0.26 | 0.15 | 0.21 | 0.21 | 0.24 | 0.31 | 0.18 | 0.27 | 0.18 | 0.17 |
| G | 0.16 | 0.26 | 0.11 | 0.2 | 0.13 | 0.22 | 0.17 | 0.09 | 0.09 | 0.24 | 0.23 | 0.13 |
| H | 0.26 | 0.16 | 0.11 | 0.07 | 0.12 | 0.16 | 0.15 | 0.25 | 0.08 | 0.26 | 0.13 | 0.21 |
| I | 0.17 | 0.23 | 0.16 | 0.17 | 0.23 | 0.09 | 0.09 | 0.1 | 0.05 | 0.17 | 0.25 | 0.2 |
| J | 0.17 | 0.08 | 0.04 | 0.16 | 0.13 | 0.18 | 0.08 | 0.14 | 0.15 | 0.12 | 0.09 | 0.14 |
| K | 0.24 | 0.12 | 0.11 | 0 | 0.08 | 0.08 | 0.2 | 0.13 | 0.04 | 0.12 | 0.13 | 0.13 |
| L | 0.16 | 0.13 | 0.11 | 0.08 | 0.08 | 0.13 | 0.17 | 0.18 | 0.19 | 0.2 | 0.09 | 0.04 |

Mean band sharing coefficient, Bkp/Kf = 0·146 ± 0·068 (±46%)

Bkp Kf

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.08 | 0.16 | 0.08 | 0.26 | 0.16 | 0.25 | 0.2 | 0.16 | 0.16 | 0.09 | 0.16 | 0.17 | 0.17 | 0.13 | 0.18 | 0.1 |
| B | 0.21 | 0.18 | 0.04 | 0.14 | 0.09 | 0.18 | 0.22 | 0.21 | 0.08 | 0.1 | 0.08 | 0.13 | 0.09 | 0.1 | 0.15 | 0.11 |
| C | 0.29 | 0.13 | 0.17 | 0.23 | 0.13 | 0.17 | 0.13 | 0.12 | 0.17 | 0.19 | 0.2 | 0.22 | 0.04 | 0.09 | 0.14 | 0.16 |
| D | 0.36 | 0.2 | 0.16 | 0.08 | 0.2 | 0.21 | 0.2 | 0.2 | 0.2 | 0.18 | 0.2 | 0.08 | 0.17 | 0.18 | 0.23 | 0.1 |
| E | 0.17 | 0.09 | 0 | 0.05 | 0.09 | 0.09 | 0.23 | 0.08 | 0.18 | 0.15 | 0.17 | 0.04 | 0.1 | 0.1 | 0.1 | 0.05 |
| F | 0.36 | 0.21 | 0.13 | 0.13 | 0.08 | 0.26 | 0.21 | 0.16 | 0.08 | 0.23 | 0.12 | 0.17 | 0.04 | 0 | 0.14 | 0.05 |
| G | 0.17 | 0.08 | 0.13 | 0.14 | 0.18 | 0.22 | 0.09 | 0.08 | 0.21 | 0.09 | 0.13 | 0.27 | 0.24 | 0.2 | 0.1 | 0.06 |
| H | 0.12 | 0.08 | 0.16 | 0.22 | 0.16 | 0.29 | 0.08 | 0.16 | 0.12 | 0.18 | 0.16 | 0.17 | 0.13 | 0.05 | 0.09 | 0.15 |
| I | 0.09 | 0.1 | 0.05 | 0.15 | 0.19 | 0.24 | 0.1 | 0.02 | 0.14 | 0.11 | 0.14 | 0.2 | 0.21 | 0.11 | 0.05 | 0.06 |
| J | 0.17 | 0.04 | 0.09 | 0.19 | 0.13 | 0.18 | 0.18 | 0.13 | 0.04 | 0.25 | 0.21 | 0.18 | 0.09 | 0.05 | 0.05 | 0.11 |
| K | 0.22 | 0.04 | 0.21 | 0.09 | 0.21 | 0.22 | 0.17 | 0.25 | 0.04 | 0.14 | 0.16 | 0.04 | 0.09 | 0 | 0.04 | 0.05 |
| L | 0.25 | 0.08 | 0.17 | 0.28 | 0.26 | 0.22 | 0.22 | 0.17 | 0.17 | 0.09 | 0.21 | 0.13 | 0.14 | 0.25 | 0.15 | 0.22 |

Mean band sharing coefficient, Btf/Kf = 0·175 ± 0·072 (±41%)

Btf Kf

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.23 | 0.17 | 0.28 | 0.18 | 0.14 | 0.21 | 0.21 | 0.1 | 0.2 | 0.17 | 0.18 | 0.21 | 0.07 | 0.15 | 0.03 | 0.2 |
| B | 0.17 | 0.21 | 0.25 | 0.11 | 0.18 | 0.33 | 0.11 | 0.14 | 0.25 | 0.23 | 0.28 | 0.26 | 0.27 | 0.2 | 0.16 | 0.17 |
| C | 0.25 | 0.2 | 0.16 | 0.14 | 0.2 | 0.13 | 0.16 | 0.25 | 0.16 | 0.28 | 0.13 | 0.13 | 0.03 | 0.11 | 0.11 | 0.07 |
| D | 0.06 | 0.07 | 0.03 | 0.11 | 0.17 | 0.1 | 0.07 | 0.06 | 0.17 | 0.11 | 0.1 | 0.1 | 0.11 | 0.15 | 0.15 | 0.04 |
| E | 0.21 | 0.14 | 0.22 | 0.27 | 0.11 | 0.18 | 0.18 | 0.21 | 0.14 | 0.16 | 0.17 | 0.18 | 0.07 | 0.2 | 0.12 | 0.17 |
| F | 0.36 | 0.22 | 0.15 | 0.12 | 0.03 | 0.19 | 0.11 | 0.21 | 0.26 | 0.28 | 0.25 | 0.19 | 0.12 | 0.12 | 0.25 | 0.18 |
| G | 0.24 | 0.17 | 0.28 | 0.15 | 0.17 | 0.18 | 0.17 | 0.24 | 0.21 | 0.3 | 0.14 | 0.29 | 0.19 | 0.07 | 0.27 | 0.21 |
| H | 0.29 | 0.32 | 0.19 | 0.12 | 0.07 | 0.23 | 0.07 | 0.07 | 0.26 | 0.2 | 0.26 | 0.19 | 0.12 | 0.29 | 0.12 | 0.14 |
| I | 0.27 | 0.12 | 0.24 | 0.12 | 0.12 | 0.12 | 0.2 | 0.11 | 0.31 | 0.21 | 0.27 | 0.08 | 0.12 | 0.17 | 0.08 | 0.05 |
| J | 0.16 | 0.21 | 0.31 | 0.18 | 0.1 | 0.24 | 0.17 | 0.16 | 0.13 | 0.15 | 0.26 | 0.24 | 0.25 | 0.26 | 0.22 | 0.2 |
| K | 0.22 | 0.23 | 0.23 | 0.16 | 0.11 | 0.24 | 0.08 | 0.19 | 0.27 | 0.04 | 0.07 | 0.16 | 0.21 | 0.17 | 0.17 | 0.09 |
| L | 0.26 | 0.27 | 0.27 | 0.2 | 0.15 | 0.35 | 0.23 | 0.18 | 0.22 | 0.16 | 0.22 | 0.11 | 0.08 | 0.08 | 0.04 | 0.09 |

# APPENDIX C

## Muntjac D-loop Data

(sizes in base pairs)

| | | |
|---|---|---|
| *Hin* fI 720 = A | *Mbo* I 590/540 = D | *Rsa* I 670/300 = F |
| *Hin* fI 590/130 = B | *Mbo* I 590/490 = E | *Rsa* I 540/300/130 = G |
| *Hin* fI 390/330 = C | | *Rsa* I 540/250/130 = H |

| Pop. | H | M | R | Sex | Pop. | H | M | R | Sex | Pop. | H | M | R | Sex | Pop. | H | M | R | Sex |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wav 01 | A | E | G | m | Sal 01 | A | E | G | m | Bdp 01 | A | E | G | m | Wst 09 | A | E | G | m |
| Wav 02 | A | E | G | m | Sal 02 | B | D | G | m | Bdp 02 | A | E | G | f | Wst 10 | A | E | G | f |
| Wav 03 | B | D | G | m | Sal 03 | B | D | G | m | Bdp 03 | A | E | G | ? | Wst 11 | A | E | G | f |
| Wav 04 | B | D | G | f | Sal 04 | B | D | G | m | Spp 01 | A | E | G | m | Wst 12 | A | E | G | ? |
| Wav 05 | B | D | G | f | Sal 05 | B | D | G | m | Spp 02 | A | E | G | m | Wst 13 | A | E | G | m |
| | | | | | Sal 06 | A | D | F | m | Spp 03 | A | E | G | m | Wst 14 | A | D | F | f |
| Bst 01 | A | D | F | f | Sal 07 | A | E | G | m | Spp 04 | A | E | G | m | | | | | |
| Bst 02 | A | D | F | m | Sal 08 | A | D | F | m | Spp 05 | A | E | G | m | Owd 01 | A | E | G | m |
| Bst 03 | A | D | F | m | Sal 09 | B | D | G | f | Bwe 01 | A | E | G | m | Owd 02 | A | E | G | m |
| Tus 01 | A | D | F | f | Sal 10 | B | D | G | f | Bwe 02 | A | E | G | f | Owd 03 | A | E | G | m |
| Tus 02 | A | D | F | m | Sal 11 | A | E | G | f | Bwe 03 | A | E | G | ? | Owd 04 | A | E | G | m |
| Tus 03 | A | D | F | m | Sal 12 | B | D | G | f | Bwe 04 | A | E | G | ? | Owd 05 | A | E | G | m |
| Tus 04 | A | D | F | f | Htn 01 | A | E | G | ? | Bwe 05 | A | E | G | m | Owd 06 | A | E | G | f |
| Tus 05 | A | D | F | f | Htn 02 | A | E | G | ? | | | | | | Owd 07 | A | E | G | f |
| Tus 06 | A | D | F | f | | | | | | Cpy 01 | A | E | G | m | Owd 08 | A | E | G | f |
| Tus 07 | A | E | G | m | Wob 01 | A | E | G | m | Bxw 01 | A | E | G | m | Owd 09 | A | E | G | f |
| Tus 08 | B | D | G | m | Wob 02 | A | E | G | m | Cmf 01 | A | E | G | m | Owd 10 | A | E | G | f |
| Wil 01 | B | D | G | m | Wob 03 | B | D | G | m | Elg 01 | A | E | G | m | Owd 11 | A | E | G | f |
| Wil 02 | A | E | G | m | Wob 04 | B | D | G | m | Elg 02 | A | E | G | m | Owd 12 | A | E | G | f |
| Wil 03 | A | D | F | m | Wob 05 | B | D | G | m | Elg 03 | A | E | G | m | Owd 13 | A | E | G | f |
| Wil 04 | A | D | F | m | Wob 06 | A | E | G | f | Elg 04 | A | E | G | m | Owd 14 | A | E | G | f |
| Wil 05 | A | D | F | f | Wob 07 | A | E | G | f | Elg 05 | C | D | F | m | Owd 15 | A | E | G | m |
| Wil 06 | A | D | F | f | Wob 08 | A | E | G | f | Elg 06 | A | E | G | m | Owd 16 | A | E | G | m |
| Wil 07 | A | D | F | f | Wob 09 | A | E | G | f | Elg 07 | C | D | F | f | Owd 17 | A | E | G | m |
| | | | | | Wob 10 | C | D | F | f | Elg 08 | A | E | G | f | Owd 18 | A | E | G | m |
| Sbl 01 | A | D | F | m | Wob 11 | C | D | F | f | Elg 09 | A | E | G | f | Owd 19 | A | E | G | m |
| Sbl 02 | A | D | F | m | Wob 12 | C | D | F | ? | Elg 10 | A | E | G | f | Owd 20 | B | D | G | m |
| Sbl 03 | A | D | F | m | Wob 13 | A | E | G | m | Elg 11 | A | E | G | f | | | | | |
| Sbl 04 | A | D | F | m | Wob 14 | C | D | F | m | Elg 12 | A | D | F | f | Nth 01 | B | D | G | m |
| Sbl 05 | B | E | G | m | Wob 15 | C | D | F | ? | Elg 13 | A | E | G | f | Nth 02 | A | E | G | m |
| Sbl 06 | A | D | F | f | Wob 16 | C | D | F | m | Elg 14 | A | E | G | f | Nth 03 | A | E | G | m |
| Sbl 07 | A | D | F | m | Wob 17 | B | D | G | m | Elg 15 | A | E | G | f | Nth 04 | A | E | G | m |
| Sbl 08 | A | D | F | f | Wob 18 | C | D | F | ? | Elg 16 | A | D | F | ? | Nth 05 | A | E | G | f |
| Sbl 09 | A | D | F | f | Wob 19 | C | D | F | ? | Elg 17 | A | E | G | ? | Nth 06 | A | E | G | f |
| Hzl 01 | B | E | G | m | Wob 20 | C | D | F | f | Hmt 01 | A | E | G | m | Nth 07 | A | E | G | f |
| Hzl 02 | A | D | F | m | Wob 21 | C | D | F | m | Hmt 02 | A | E | G | m | Nth 08 | A | E | G | f |
| Hzl 03 | B | E | G | m | Wob 22 | C | D | F | m | Hmt 03 | C | E | F | m | Nth 09 | A | E | G | f |
| Hzl 04 | B | E | G | m | Wob 23 | C | D | F | ? | Hmt 04 | A | E | G | m | Nth 10 | A | E | G | f |
| Hzl 05 | A | D | F | f | Wob 24 | C | D | F | f | Hmt 05 | A | E | G | f | Nth 11 | A | E | G | m |
| Hzl 06 | B | E | G | f | | | | | | Hmt 06 | A | E | G | f | Nth 12 | A | E | G | m |
| Bkl 01 | B | E | G | m | Hpk 01 | A | D | F | m | Hmt 07 | A | E | G | ? | Nth 13 | A | E | G | m |
| Bkl 02 | B | E | G | m | Hpk 02 | A | E | G | m | Hmt 08 | A | E | G | ? | Nth 14 | A | E | G | m |
| Bkl 03 | A | D | F | m | Hpk 03 | A | E | G | m | Hmt 09 | C | D | F | m | Nth 15 | A | E | G | f |
| Bkl 04 | A | D | F | f | Fmw 01 | A | E | G | m | Hmt 10 | A | E | G | ? | Nth 16 | A | E | G | m |
| Bkl 05 | A | E | G | f | Fmw 02 | A | E | G | f | | | | | | Nth 17 | A | E | G | m |
| Bkl 06 | A | E | G | f | Bsh 01 | A | E | G | m | Wst 01 | C | D | F | m | Nth 18 | A | E | G | m |
| Bkl 07 | A | E | G | f | Wdh 01 | A | D | F | f | Wst 02 | A | E | G | m | Hmw 01 | A | E | G | m |
| Sst 01 | B | D | G | m | | | | | | Wst 03 | A | E | G | m | Hmw 02 | A | E | G | m |
| Sst 02 | A | E | G | m | | | | | | Wst 04 | A | D | F | m | Hmw 03 | A | E | G | f |
| Sst 03 | A | D | F | m | | | | | | Wst 05 | A | E | G | m | Hmw 04 | A | E | G | f |
| | | | | | | | | | | Wst 06 | A | D | F | m | Hmw 05 | A | E | G | f |
| | | | | | | | | | | Wst 07 | A | E | G | m | Hmw 06 | A | E | G | f |
| | | | | | | | | | | Wst 08 | A | E | G | m | Hmw 07 | A | E | G | f |

| Pop. | H | M | R | Sex |
|---|---|---|---|---|
| Min 01 | A | E | G | m |
| Min 02 | A | E | G | m |
| Min 03 | A | E | G | f |
| Wwh 01 | A | E | G | m |
| Wwh 02 | A | E | G | f |
| Wwh 03 | A | E | G | m |
| Wwh 04 | A | E | G | f |
| Wwd 01 | A | E | G | m |
| Wwd 02 | A | E | G | m |
| Wwd 03 | A | E | G | f |
| Wwd 04 | A | E | G | m |
| Wwd 05 | A | E | G | f |
| Wwd 06 | A | E | G | m |
| Wwd 07 | A | E | G | m |
| Wwd 08 | A | E | G | f |
| Wwd 09 | A | E | G | m |
| Wwd 10 | A | E | G | f |
| | | | | |
| Bkt 01 | A | D | F | f |
| Bkt 02 | A | E | G | ? |
| Bkt 03 | A | E | G | ? |
| Bkt 04 | B | D | H | ? |
| Bkt 05 | A | E | G | ? |
| Bkt 06 | A | E | G | ? |
| Bkt 07 | C | D | F | ? |
| Bkt 08 | C | D | F | ? |
| Bkt 09 | B | D | H | ? |
| Bkp 01 | A | E | G | f |
| Bkp 02 | A | E | G | f |
| Bkp 03 | A | E | G | f |
| Bkp 04 | A | E | G | f |
| Bkp 05 | A | E | G | f |
| Bkp 06 | A | E | G | f |
| Bkp 07 | C | D | F | m |
| Bkp 08 | C | D | F | m |
| Bkp 09 | A | E | G | m |
| Bkp 10 | C | D | F | m |
| Bkp 11 | B | D | H | m |
| Bkp 12 | A | E | G | m |
| Bkp 13 | A | D | F | m |
| | | | | |
| Grf 01 | A | D | F | m |
| Ntw 01 | B | D | G | m |
| Ntw 02 | B | D | G | m |
| Ntw 03 | B | D | G | m |
| Ntw 04 | B | D | G | f |
| Ntw 05 | B | D | G | f |
| Ntw 06 | B | D | H | f |
| Ntw 07 | B | D | H | f |
| Ntw 08 | B | D | H | f |
| Ntw 09 | B | D | H | f |
| Ntw 10 | B | D | G | f |
| Ntw 11 | A | E | G | m |
| Ntw 12 | B | D | G | m |

| Pop. | H | M | R | Sex |
|---|---|---|---|---|
| Ntw 13 | B | D | G | f |
| Ntw 14 | B | D | G | m |
| Knb 01 | A | E | G | m |
| Knb 02 | B | D | G | f |
| Knb 03 | B | D | G | m |
| Knb 04 | A | D | F | ? |
| Knb 05 | A | D | F | m |
| Knb 06 | A | D | F | m |
| Knb 07 | B | D | G | f |
| Knb 08 | B | D | G | m |
| Knb 09 | B | D | G | f |
| Knb 10 | A | D | G | f |
| Knb 11 | B | D | G | ? |
| Wtg 01 | B | D | H | m |
| Wtg 02 | B | D | H | m |
| | | | | |
| Hex 01 | A | E | G | m |
| Hex 02 | A | E | G | m |
| Hex 03 | A | E | G | m |
| Hex 04 | A | E | G | m |
| Hex 05 | A | E | G | f |
| Hex 06 | A | E | G | f |
| Hex 07 | A | E | G | f |
| Hex 08 | A | E | G | m |
| Hex 09 | A | E | G | m |
| Hex 10 | A | D | F | m |
| Hex 11 | A | D | F | m |
| Hex 12 | A | E | G | f |
| Hex 13 | A | E | G | f |
| Hex 14 | A | D | F | f |
| Hex 15 | A | E | G | m |
| Hex 16 | A | E | G | m |
| Hex 17 | A | E | G | m |
| Hex 18 | A | E | G | m |
| Hex 19 | A | E | G | f |
| Hex 20 | B | D | H | m |
| Hex 21 | A | E | G | f |
| Hex 22 | A | E | G | f |
| Hex 23 | A | E | G | m |
| Hex 24 | A | E | G | f |
| Hex 25 | A | E | G | m |
| Hex 26 | A | E | G | ? |
| Hex 27 | A | E | G | ? |
| Hex 28 | A | E | G | f |
| Hex 29 | A | E | G | f |
| | | | | |
| Btf 01 | A | E | G | m |
| Btf 02 | A | E | G | m |
| Btf 03 | A | E | G | m |
| Btf 04 | B | D | H | m |
| Btf 05 | A | E | G | f |
| Btf 06 | B | D | H | f |
| Btf 07 | B | D | H | f |
| Btf 08 | A | E | G | m |

| Pop. | H | M | R | Sex |
|---|---|---|---|---|
| Btf 09 | A | E | G | m |
| Btf 10 | A | E | G | m |
| Btf 11 | A | E | G | m |
| Btf 12 | A | E | G | f |
| Btf 13 | B | D | H | m |
| Btf 14 | A | E | G | f |
| Btf 15 | A | E | G | ? |
| Btf 16 | B | D | H | m |
| Btf 17 | B | D | H | m |
| Btf 18 | A | E | G | f |
| | | | | |
| Kf 01 | C | D | F | m |
| Kf 02 | C | D | F | m |
| Kf 03 | C | D | F | m |
| Kf 04 | C | D | F | m |
| Kf 05 | C | D | F | m |
| Kf 06 | C | D | F | m |
| Kf 07 | C | D | F | m |
| Kf 08 | C | D | F | f |
| Kf 09 | C | D | F | f |
| Kf 10 | C | D | F | f |
| Kf 11 | C | D | F | ? |
| Kf 12 | C | D | F | ? |
| Kf 13 | C | D | F | m |
| Kf 14 | C | D | F | m |
| Kf 15 | C | D | F | m |
| Kf 16 | C | D | F | m |
| Kf 17 | C | D | F | m |
| Kf 18 | C | D | F | m |
| Kf 19 | C | D | F | f |
| Kf 20 | C | D | F | m |
| Kf 21 | A | D | G | f |
| Kf 22 | C | D | F | m |
| Kf 23 | C | D | F | m |
| Kf 24 | C | D | F | f |
| Kf 25 | C | D | F | m |
| Kf 26 | C | D | F | f |
| Kf 27 | C | D | F | f |
| Kf 28 | C | D | F | m |
| Kf 29 | C | D | F | f |
| Kf 30 | C | D | F | f |
| Kf 31 | C | D | F | m |
| Kf 32 | C | D | F | f |
| Kf 33 | C | D | F | f |
| Kf 34 | C | D | F | ? |
| Kf 35 | C | D | F | m |

| Pop. | H | M | R | Sex |
|---|---|---|---|---|
| Eus 01 | B | D | G | m |
| Eus 02 | C | D | F | m |
| Eus 03 | C | D | F | m |
| Eus 04 | B | D | G | m |
| Eus 05 | A | E | G | m |
| Eus 06 | C | D | F | m |
| Eus 07 | C | D | F | m |
| Eus 08 | C | D | F | m |
| Eus 09 | A | E | G | f |
| Eus 10 | A | E | G | m |
| Eus 11 | A | E | G | m |
| Eus 12 | C | D | F | m |
| Eus 13 | B | D | G | f |
| Eus 14 | B | D | G | f |
| Eus 15 | A | E | G | f |
| Eus 16 | A | E | G | f |
| Eus 17 | B | D | G | m |
| Eus 18 | C | D | F | m |
| Eus 19 | A | E | G | f |
| Eus 20 | A | E | G | f |
| Eus 21 | A | E | G | m |
| Eus 22 | A | E | G | m |
| Eus 23 | C | D | F | m |
| Eus 24 | B | D | G | m |
| Eus 25 | C | D | F | m |
| | | | | |
| Twn 01 | A | D | F | f |
| Twn 02 | A | D | F | f |
| Twn 03 | A | D | F | f |
| Twn 04 | A | D | F | f |
| Twn 05 | A | D | F | m |
| Twn 06 | A | D | F | m |
| Twn 07 | A | D | F | f |
| Twn 08 | A | D | F | m |
| Twn 09 | A | D | F | f |
| Twn 10 | A | D | F | m |
| Twn 11 | A | D | F | f |
| Twn 12 | A | D | F | m |
| Twn 13 | A | D | F | m |
| Twn 14 | A | D | F | f |
| Twn 15 | A | D | F | m |
| Twn 16 | A | D | F | ? |
| Twn 17 | A | D | F | f |
| Twn 18 | A | D | F | f |
| Twn 19 | A | D | F | ? |
| Twn 20 | A | D | F | m |
| Twn 21 | A | D | F | ? |
| Twn 22 | A | D | F | m |
| Twn 23 | A | D | F | f |
| Twn 24 | A | D | F | m |
| Twn 25 | A | D | F | m |

# APPENDIX $C_1$

## Input matrix for PAUP

begin data;
dimensions nta=18 nchar=225;
format symbols ="0~1";
matrix

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| WAV | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 |
| | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | | |
| TUS | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 |
| | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 |
| | 010011001 | 110011101 | 110011101 | 110011101 | 010011111 | 010011111 | 010011111 | | |
| SBL | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 |
| | 010011001 | 010011001 | 010011001 | 010011001 | 110011111 | 110011111 | 110011111 | 110011111 | 110011111 |
| | 110011111 | 110011111 | 110011101 | 110011101 | 110011101 | 110011101 | 010011111 | | |
| SAL | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 |
| | 010011001 | 010011001 | 010011001 | 010011111 | 010011111 | 010011111 | 010011111 | | |
| WOB | 010011001 | 010011001 | 010011001 | 010011101 | 010011101 | 010011101 | 110011101 | 110011101 | 110011101 |
| | 010011111 | 010011111 | 010011111 | 001011111 | 001011111 | 001011111 | 110011111 | 110011111 | 110011111 |
| | 010111001 | 010111001 | 010111001 | 110111001 | 110111001 | 110111001 | ????????? | | |
| HPK | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | | |
| BDP | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | | |
| ELG | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 010111001 | 010111001 | 010011001 | 010011001 | 110111001 | | |
| WST | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 010011001 | 010011001 | 010011001 | 010011001 | 010011001 | 010111001 | 010111001 | | |
| OWD | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 010011111 | | |
| NTH | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 010011111 | | |
| MIN | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | | |
| BKT | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 010111001 | 010111001 | 010111001 | 010111001 |
| | 010111001 | 010111001 | 001011111 | 001011111 | 001011111 | 010011001 | 010011001 | | |
| KNB | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 |
| | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 001011111 | 001011111 | 001011111 | 001011111 |
| | 001011111 | 010011001 | 010011001 | 010011001 | 110011101 | 110011101 | 010011101 | | |
| HEX | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 010011001 | 010011001 | 010011001 | 001011111 | | |
| BTF | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 001011111 |
| | 001011111 | 001011111 | 001011111 | 001011111 | 001011111 | 001011111 | 001011111 | | |
| KF | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 |
| | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 |
| | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010011101 | | |
| EUS | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 | 110011101 |
| | 110011101 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 | 010111001 |
| | 010111001 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | 010011111 | | |

;
END;

# APPENDIX D

# MUNTJAC MICROSATELLITE DATA

| | Loci & Alleles | | | | | | |
|---|---|---|---|---|---|---|---|
| Population | JP23 | JP27 | JP38 | BOVIRBP | MAF23 | MAF65 | DRBP1 |
| WAVERLY | | | | | | | |
| Wav 01 | GG | FF | BC | GN | LO | HH | AD |
| Wav 02 | FF | CC | AC | OO | JL | HH | AD |
| Wav 03 | AF | BB | CD | EO | JO | HH | AD |
| Wav 04 | GG | BC | AC | DD | IO | HH | AD |
| Wav 05 | AG | BC | AA | BN | KK | HH | AD |
| TUSMORE | | | | | | | |
| Bst 01 | GG | BC | AC | OO | KO | HH | AD |
| Bst 02 | AG | CC | AC | OO | AC | BI | AD |
| Bst 03 | AG | CC | AA | GJ | EE | HK | AD |
| Tus 01 | FF | BC | CC | EJ | OO | HH | AD |
| Tus 02 | GG | BC | BC | GJ | OO | KK | AD |
| Tus 03 | AG | BF | BC | GJ | EO | HK | AD |
| Tus 04 | GG | CC | BC | OO | NN | KK | AD |
| Tus 05 | CG | BC | AA | JO | IO | BB | AD |
| Tus 06 | AG | CF | BC | JN | OO | BB | AD |
| Tus 07 | GG | CF | AC | BG | CC | BK | AA |
| Tus 08 | GG | CC | BC | OO | LO | BK | AA |
| Wil 01 | AG | CC | BD | JM | OO | BK | AA |
| Wil 02 | AA | BC | CC | BO | OO | BB | DD |
| Wil 03 | AF | BC | BC | GJ | JN | HH | AA |
| Wil 04 | AA | BC | AC | EO | II | HH | AD |
| Wil 05 | AA | BC | AB | GO | LN | HK | AD |
| Wil 06 | BB | BB | AD | BB | HM | BB | DD |
| Wil 07 | AA | BC | BD | BJ | NN | BH | AD |
| SOUTHBLOCK | | | | | | | |
| Sbl 01 | AF | BC | AC | OO | CK | HH | AD |
| Sbl 02 | AF | CC | BB | BG | IN | KK | AD |
| Sbl 03 | AF | BC | AC | BO | LN | BB | AD |
| Sbl 04 | BF | BC | BD | BO | JO | KK | AD |
| Sbl 05 | CF | CC | BB | GN | IN | HH | AD |
| Sbl 06 | BB | CC | BC | EO | JO | BH | AD |
| Sbl 07 | AF | CC | BC | GO | JO | BB | AD |
| Sbl 08 | AA | CF | AB | OO | GL | HH | AD |
| Sbl 09 | FF | BC | BB | EG | HM | KK | AD |
| Hzl 01 | AG | BC | AA | OO | IO | BK | AD |
| Hzl 02 | AG | CC | BB | OO | HO | BK | AD |
| Hzl 03 | GG | CC | AC | OO | LN | BB | DD |
| Hzl 04 | AF | BB | BB | BO | OO | BB | AD |
| Hzl 05 | GG | CC | AB | GO | JN | BH | DD |
| Hzl 06 | AF | BC | AC | OO | JJ | BB | AD |
| Bkl 01 | GG | BC | CE | EO | LN | HH | AD |
| Bkl 02 | AG | CC | AC | EE | NN | BB | AD |
| Bkl 03 | AG | BC | BB | FF | IL | HH | AA |
| Bkl 04 | AG | BC | BD | OO | LO | HH | AA |
| Bkl 05 | AA | BC | BC | BG | LN | BB | AA |
| Bkl 06 | AA | BC | BE | GO | LO | BH | AA |
| Bkl 07 | AG | CC | BC | GO | IO | HH | AA |
| Sst 01 | AF | BC | AB | BN | OO | HK | AA |
| Sst 02 | AG | BC | CE | OO | KN | BB | AA |
| Sst 03 | AF | BC | AB | EE | IO | HK | AD |
| SALCEY FOREST | | | | | | | |
| Sal 01 | AA | BB | EE | FN | IM | BB | DD |
| Sal 02 | AG | BC | CE | EO | MM | BI | DD |
| Sal 03 | AF | CC | CC | OO | OC | KK | DD |
| Sal 04 | AG | CF | EE | EO | BE | BB | DD |
| Sal 05 | AG | BC | BC | EN | OO | BB | AD |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sal 06 | AG | BC | DD | GO | OO | HH | DD |
| Sal 07 | AF | BC | EE | OO | JO | KK | DD |
| Sal 08 | GG | BC | BB | EG | KK | BB | AD |
| Sal 09 | AA | BC | BC | BJ | NN | HH | DD |
| Sal 10 | AF | CC | BC | EJ | JN | HH | DD |
| Sal 11 | AF | BC | BE | BJ | EE | KK | AD |
| Sal 12 | GG | BC | BB | EN | NN | HH | DD |
| Htn 01 | AA | CF | CC | GJ | KK | BB | AD |
| Htn 02 | AG | CC | CD | JO | GN | BH | AD |
| WOBURN ABBEY | | | | | | | |
| Wob 01 | FF | BB | CC | EO | GK | KK | AD |
| Wob 02 | FF | BB | CE | OO | EE | BH | DD |
| Wob 03 | AC | BB | BE | OO | LL | HH | DD |
| Wob 04 | AC | BB | AE | GO | LL | BB | DD |
| Wob 05 | AF | BB | BD | GO | LL | HH | DD |
| Wob 06 | AC | BB | BB | JO | LL | BB | AD |
| Wob 07 | DF | CC | CD | GJ | LL | KK | DD |
| Wob 08 | AF | BC | BC | NN | MM | HH | AD |
| Wob 09 | CF | CC | AB | HO | MM | BB | AD |
| Wob 10 | AC | CC | CD | EE | NN | BB | AD |
| Wob 11 | AA | CE | AC | OO | KN | HH | DD |
| Wob 12 | AF | BB | BB | BG | LL | KK | AD |
| Wob 13 | AG | CC | CC | EG | CE | BK | AD |
| Wob 14 | FF | BC | CC | BE | DG | HK | DD |
| Wob 15 | CF | BB | BB | DG | EH | II | AD |
| Wob 16 | AC | BC | BB | GG | BF | BK | DD |
| Wob 17 | CF | BC | CC | BG | DF | II | DD |
| Wob 18 | AF | BC | CD | EG | MM | KK | DD |
| Wob 19 | AF | BC | BB | EO | NN | KK | DD |
| Wob 20 | AC | BC | CC | DM | OO | KK | DD |
| Wob 21 | AC | BB | BC | GG | DF | II | DD |
| Wob 22 | AF | BF | BB | BG | DH | KK | DD |
| Wob 23 | AG | BB | BC | DD | OO | KK | DD |
| Wob 24 | FF | FF | CD | DG | LN | BK | DD |
| HARRY'S PARK | | | | | | | |
| Hpk 01 | AA | CC | CD | EE | BG | KK | AD |
| Hpk 02 | AC | CC | CE | EM | MO | KK | DD |
| Hpk 03 | FF | BB | AB | BE | CN | BK | AD |
| Fmw 01 | AG | BC | BE | BE | CN | KK | DD |
| Fmw 02 | CG | BC | BE | DG | CN | KK | DD |
| Bsh 01 | AG | BC | BC | GO | CL | BB | DD |
| Wdh 01 | BF | CC | AC | BO | MM | BK | DD |
| BOWDEN PARK | | | | | | | |
| Bdp 01 | AF | CC | AA | GG | CN | HH | DD |
| Bdp 02 | AG | BB | CE | GG | NO | HK | AD |
| Bdp 03 | AG | BC | CE | BO | MO | BH | AD |
| Spp 01 | CF | CC | AE | GO | GL | HH | DD |
| Spp 02 | AF | BC | AB | GO | GL | HK | DD |
| Spp 03 | AG | BC | AE | GO | GO | KK | AD |
| Spp 04 | AG | CC | AE | GG | NN | BH | DD |
| Spp 05 | AF | BC | AE | GG | NN | BH | DD |
| Bwe 01 | AG | CC | AE | GO | EN | BB | DD |
| Bwe 02 | AG | CC | AE | GG | EM | BB | DD |
| Bwe 03 | AF | BC | AE | BB | CC | BB | DD |
| Bwe 04 | AF | BB | EE | OO | NN | HH | DD |
| Bwe 05 | AG | BC | CE | GO | CL | BB | DD |
| ELING COMMON | | | | | | | |
| Cpy 01 | AG | CF | CE | BO | DH | BB | DD |
| Bxw 01 | AG | BB | AC | EG | DN | BB | DD |
| Cmf 01 | AF | BC | CC | EM | EM | BB | AD |
| Elg 01 | AG | BB | CE | OO | NN | KK | DD |
| Elg 02 | AG | BB | EE | GO | NN | BK | DD |
| Elg 03 | AG | CF | BE | GG | EN | HK | DD |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Elg 04 | CF | BB | CE | OO | EM | KK | DD |
| Elg 05 | AG | CC | CE | OO | LN | KK | DD |
| Elg 06 | AF | BC | CC | OO | EM | HK | DD |
| Elg 07 | AG | CF | CE | OO | MM | HK | DD |
| Elg 08 | AF | BB | CC | OO | OO | KK | DD |
| Elg 09 | AG | BB | EE | GO | LL | KK | DD |
| Elg 10 | AG | BB | BE | GO | MN | BB | DD |
| Elg 11 | AG | BB | CE | OO | MO | BB | DD |
| Elg 12 | AG | BF | CC | GO | MO | KK | DD |
| Elg 13 | AF | BB | CE | OO | NN | BB | DD |
| Elg 14 | AG | BB | CE | GG | NO | KK | DD |
| Elg 15 | AG | BB | EE | GO | NN | HK | AD |
| Elg 16 | AG | BF | CC | OO | IM | BB | AD |
| Elg 17 | AG | BB | EE | GG | HL | BB | DD |
| Hmt 01 | AG | FF | BE | EG | MM | KK | DD |
| Hmt 02 | AG | BC | CE | BG | GL | KK | DD |
| Hmt 03 | AG | CF | BC | DD | NN | BH | DD |
| Hmt 04 | AF | BF | EE | EG | NO | BB | DD |
| Hmt 05 | AG | CC | BE | GO | NN | BH | AD |
| Hmt 06 | AG | BC | AE | GO | JM | KK | AD |
| Hmt 07 | AG | CC | AE | IN | NN | HK | AD |
| Hmt 08 | AG | BB | AD | GO | LL | KK | AD |
| Hmt 09 | AG | BC | BE | BO | KN | HK | AD |
| Hmt 10 | AF | BB | BB | FG | JM | KK | DD |
| | | | WORMSLEY ESTATE | | | | |
| Wst 01 | AG | CF | BB | OO | NN | KK | AD |
| Wst 02 | AG | BC | BB | AE | JM | BH | AD |
| Wst 03 | AG | BC | EE | OO | MO | BK | AD |
| Wst 04 | AG | BC | BC | AE | LN | BB | DD |
| Wst 05 | AG | CC | CE | OO | NN | BK | AD |
| Wst 06 | FF | CC | AA | EO | HH | KK | DD |
| Wst 07 | AG | BB | CD | EG | LN | KK | DD |
| Wst 08 | FF | CC | AC | AE | NN | BH | DD |
| Wst 09 | AG | BC | CC | GJ | NN | BH | DD |
| Wst 10 | AG | FF | AC | JO | DM | HH | DD |
| Wst 11 | AC | BC | BC | BB | NN | KK | DD |
| Wst 12 | AF | BC | AC | AA | IN | HK | AA |
| Wst 13 | AG | BC | BC | AJ | NN | HL | AA |
| Wst 14 | AG | BC | CE | DG | CL | BB | AD |
| | | | OLD WARDEN | | | | |
| Owd 01 | AF | CC | AC | AA | CF | BB | DD |
| Owd 02 | AG | CC | AC | GO | DF | BB | AD |
| Owd 03 | FF | BB | AC | GO | CG | BK | AD |
| Owd 04 | AG | BB | CC | EO | DG | BB | AA |
| Owd 05 | GG | BC | BC | DO | EE | BB | AD |
| Owd 06 | AG | BC | AC | EM | BE | BB | AD |
| Owd 07 | AG | BC | CE | OO | CE | BB | DD |
| Owd 08 | AG | BC | CC | DG | CG | KK | AD |
| Owd 09 | AG | BC | AC | OO | EE | BB | AD |
| Owd 10 | AG | CC | CD | EG | CG | BB | AD |
| Owd 11 | AG | BC | AA | GO | CG | BK | AD |
| Owd 12 | AG | BC | AC | DO | CE | BH | AD |
| Owd 13 | GG | BC | AD | DO | DH | BK | AD |
| Owd 14 | BG | BB | EE | OO | GG | BB | AD |
| Owd 15 | AG | BB | EE | GG | CF | BB | AD |
| Owd 16 | AF | BC | AD | OO | EG | BB | AD |
| Owd 17 | AG | CC | CC | DO | EG | BB | AD |
| Owd 18 | AG | CC | BC | DO | BF | KK | AD |
| Owd 19 | GG | CC | CC | GO | CC | BB | AD |
| Owd 20 | AG | CC | AC | EO | FF | HH | AD |
| | | | NORTHILL | | | | |
| Nth 01 | AG | CC | AC | OO | DD | KK | AA |
| Nth 02 | AG | BC | EE | EO | MO | BB | AD |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Nth 03 | AF | BC | CC | GO | DH | KK | AD |
| Nth 04 | AG | BC | BC | GE | NN | BK | AA |
| Nth 05 | CF | BC | BD | EE | OO | BH | AD |
| Nth 06 | AF | CC | CC | GG | OO | KK | AD |
| Nth 07 | FF | BC | CE | OO | OO | BB | AD |
| Nth 08 | AF | BB | CC | GO | DH | BB | AD |
| Nth 09 | AG | BC | BE | BG | EJ | HH | AD |
| Nth 10 | AG | BC | AA | OO | II | BK | AD |
| Nth 11 | CF | CC | AB | GO | DD | BK | DD |
| Nth 12 | AG | CC | CE | GG | NN | BB | DD |
| Nth 13 | AG | CC | AE | EO | IP | BK | DD |
| Nth 14 | AG | BC | AB | OO | GG | BK | AD |
| Nth 15 | AF | CC | EE | DG | EE | BK | AD |
| Nth 16 | AG | CC | BC | EG | DD | KK | AD |
| Nth 17 | AF | CC | AE | OO | KK | BK | DD |
| Nth 18 | GG | BC | BE | BG | EK | BB | DD |
| Hmw 01 | AG | BB | BC | GG | OO | BK | AA |
| Hmw 02 | AF | BC | AE | OO | EL | BK | AA |
| Hmw 03 | AG | CC | AC | GO | DF | BB | AA |
| Hmw 04 | CF | BC | CC | EO | JM | BK | AA |
| Hmw 05 | AF | CC | CE | JO | MM | KK | AA |
| Hmw 06 | AG | CC | EE | GO | EL | KK | AA |
| Hmw 07 | AG | CC | CE | JO | HL | BB | AA |
| MINSBURY HILL | | | | | | | |
| Min 01 | AG | BC | CC | MO | KL | BB | AA |
| Min 02 | AG | CC | BC | EO | MO | KK | AA |
| Min 03 | GG | BB | BC | BO | II | BB | AA |
| Wwh 01 | AG | BC | CC | EG | CC | BB | AD |
| Wwh 02 | AF | BB | CD | DN | FF | BB | AA |
| Wwh 03 | AF | CC | CE | DO | II | BH | AD |
| Wwh 04 | AC | BB | BB | GN | LL | II | AD |
| Wwd 01 | CF | CC | CD | DO | JL | HK | AD |
| Wwd 02 | CF | CC | AC | GO | KN | BB | DD |
| Wwd 03 | AG | BC | CE | EE | MO | BK | DD |
| Wwd 04 | AF | CC | AE | EO | KL | BB | AD |
| Wwd 05 | AG | CC | CC | EO | NO | KK | AD |
| Wwd 06 | AG | BC | CC | DO | GM | HH | AD |
| Wwd 07 | AG | BC | AE | EG | EJ | BK | AD |
| Wwd 08 | CC | BB | CE | EO | GL | BB | DD |
| Wwd 09 | AF | BC | BE | EO | GL | BI | AD |
| Wwd 10 | GG | BC | AE | DO | EJ | HK | AD |
| BROCKET PARK | | | | | | | |
| Bkt 01 | AF | BC | BB | DN | GM | KK | AD |
| Bkt 02 | FF | CC | CE | GG | MM | BK | AD |
| Bkt 03 | AF | BC | BD | CO | LO | BB | AD |
| Bkt 04 | AF | BC | BE | CN | LN | BB | AD |
| Bkt 05 | AF | BC | CC | GO | LO | HH | AD |
| Bkt 06 | AF | BC | CE | EO | LO | BB | AD |
| Bkt 07 | AF | BC | CC | EO | HL | BK | AD |
| Bkt 08 | AF | BC | EE | EO | GM | BK | AD |
| Bkt 09 | CF | BC | AC | AN | LO | HK | AD |
| Bkp 01 | AG | BB | BE | NN | LN | BK | DD |
| Bkp 02 | AG | CC | AC | OO | GM | HH | AD |
| Bkp 03 | AG | BC | BD | NN | GM | BB | AD |
| Bkp 04 | AF | BC | AE | GG | NP | BB | AD |
| Bkp 05 | AG | BC | BE | EO | NN | BB | AD |
| Bkp 06 | GG | BC | CC | FO | LO | HH | AD |
| Bkp 07 | CG | BB | EE | GO | GM | KK | AD |
| Bkp 08 | AF | BB | BE | GO | HM | KK | AD |
| Bkp 09 | AF | CC | AA | GO | LN | HH | DD |
| Bkp 10 | AF | BB | EE | GG | KM | KK | DD |
| Bkp 11 | AF | BC | BE | FN | II | BH | DD |
| Bkp 12 | AG | BB | AB | OO | IN | BB | DD |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Bkp 13 | AG | BB | CE | OO | MM | HK | AD |
| KNEBWORTH | | | | | | | |
| Grf 01 | AF | BC | CE | GO | KK | HK | DD |
| Ntw 01 | AF | BC | AC | DD | JL | BH | AD |
| Ntw 02 | AF | BC | CC | GO | JL | BK | AD |
| Ntw 03 | AF | BB | BD | OO | JL | BK | AD |
| Ntw 04 | AF | BC | CC | FO | IN | HK | AD |
| Ntw 05 | AF | BC | CC | GO | IN | KK | AD |
| Ntw 06 | AF | CC | BD | DO | IN | BH | AD |
| Ntw 07 | AF | BB | CC | DD | IN | HK | AD |
| Ntw 08 | AF | BB | AC | DD | IN | BK | AD |
| Ntw 09 | AF | BC | EE | DO | IN | HH | AD |
| Ntw 10 | AA | BC | AE | BO | HN | HH | AD |
| Ntw 11 | AF | BC | BE | DO | HM | II | AD |
| Ntw 12 | AC | BC | AC | EO | LO | IK | AD |
| Ntw 13 | AF | BC | BE | DG | JO | IK | AD |
| Ntw 14 | AC | BB | CE | DG | JO | HH | AA |
| Knb 01 | AF | CC | CD | GO | JO | IK | AD |
| Knb 02 | AF | BC | CE | OO | JO | BB | AD |
| Knb 03 | AF | BC | CE | EO | LO | II | AD |
| Knb 04 | AC | CC | AC | GP | JL | BI | AD |
| Knb 05 | AF | CC | CD | GO | AF | BI | AA |
| Knb 06 | AC | BC | BD | DO | AI | BI | AA |
| Knb 07 | AF | BB | CC | DN | BH | II | AD |
| Knb 08 | AC | BC | CE | GG | GG | HK | AD |
| Knb 09 | AF | BB | CC | GO | CL | BH | AA |
| Knb 10 | AC | BB | EE | EG | EE | HH | AD |
| Knb 11 | CC | BC | BC | OO | CL | IK | AD |
| Wtg 01 | AC | BB | EE | BE | CL | HK | AD |
| Wtg 02 | AF | BB | BC | EJ | CL | HH | AD |
| HEXTON | | | | | | | |
| Hex 01 | AF | BC | CC | DG | FF | KK | AD |
| Hex 02 | AF | CC | AA | EG | EL | BB | AA |
| Hex 03 | AF | CC | CD | EG | AC | BH | DD |
| Hex 04 | AF | BC | AE | EE | CG | BH | AA |
| Hex 05 | AG | BC | AC | EE | DE | BB | AD |
| Hex 06 | GG | BC | BD | DG | GI | BB | AA |
| Hex 07 | AG | BC | CE | NN | BE | HH | AD |
| Hex 08 | AF | CC | CC | DO | CC | HH | AD |
| Hex 09 | FF | CC | CE | AG | CF | BH | AD |
| Hex 10 | AG | CC | BE | EG | BE | BB | AD |
| Hex 11 | AG | CC | CE | EG | AC | KK | AA |
| Hex 12 | AA | CC | CE | DE | CF | BB | AD |
| Hex 13 | AG | CC | BD | EG | AC | HK | AA |
| Hex 14 | AG | BC | CC | EG | CC | BK | DD |
| Hex 15 | AA | BC | AB | EG | AC | BK | AD |
| Hex 16 | AA | BB | CE | DG | BG | BB | DD |
| Hex 17 | AA | CC | AB | DG | AD | BB | AA |
| Hex 18 | AA | BC | DE | DG | AD | BB | AA |
| Hex 19 | AG | CC | CC | EJ | GG | BH | DD |
| Hex 20 | AF | BC | EE | GO | DG | BB | DD |
| Hex 21 | AG | BB | AA | GG | AC | BB | AD |
| Hex 22 | AF | BB | BC | GO | EH | KK | AD |
| Hex 23 | AF | BC | AC | BB | GG | BB | AA |
| Hex 24 | AG | BC | CE | GG | DD | BB | AD |
| Hex 25 | AG | BC | CE | EG | DI | BK | AD |
| Hex 26 | AG | BC | BC | DG | DG | HK | AD |
| Hex 27 | AA | CC | AB | DD | BB | BK | AD |
| Hex 28 | AA | BC | AE | DG | AC | BB | AD |
| Hex 29 | AG | BC | AC | EG | AF | KK | AD |
| BUNTINGFORD | | | | | | | |
| Btf 01 | AA | BC | CE | DO | BF | BH | AA |
| Btf 02 | AA | BC | CC | GO | DF | BH | AA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Btf 03 | GG | CC | AC | GO | ED | BB | AD |
| Btf 04 | AA | BB | BE | BN | GI | BK | AD |
| Btf 05 | AF | BC | AC | BG | EG | AA | DD |
| Btf 06 | AE | BC | AB | BO | ED | BK | AA |
| Btf 07 | GG | BC | AD | GG | ED | BB | AA |
| Btf 08 | AF | BC | CE | EO | AF | KK | AD |
| Btf 09 | AF | BB | CE | GO | BF | BK | AD |
| Btf 10 | AG | BC | AB | OO | CK | BH | AD |
| Btf 11 | AA | BB | AE | EJ | KO | HK | AA |
| Btf 12 | FF | CC | EE | DE | FL | BH | AD |
| Btf 13 | AA | CC | EE | FG | FN | HH | DD |
| Btf 14 | AF | BC | CE | BE | GR | II | AD |
| Btf 15 | AE | BC | CD | BG | HM | HK | AD |
| Btf 16 | GG | BC | CD | OO | DL | BB | AD |
| Btf 17 | AE | BC | CE | BE | NO | BH | AA |
| Btf 18 | AG | BC | CD | HH | KO | HK | AD |
| KING'S FOREST | | | | | | | |
| Kf 01 | AF | BB | CD | FF | NO | HK | AD |
| Kf 02 | AA | BB | CC | CF | KN | KK | AA |
| Kf 03 | AE | BC | CC | EE | GN | KK | DD |
| Kf 04 | AE | BB | BB | FF | KN | BK | AD |
| Kf 05 | AF | BF | CC | DO | KN | HH | AD |
| Kf 06 | AF | BB | CD | DO | GN | KK | AD |
| Kf 07 | CF | CC | CC | EO | JO | KK | AD |
| Kf 08 | AA | BC | CC | DG | OO | KK | DD |
| Kf 09 | DE | BC | BC | EM | OO | BK | AD |
| Kf 10 | AF | BC | BB | BN | IO | KK | AD |
| Kf 11 | AG | BB | CD | HH | OO | BK | AD |
| Kf 12 | AF | BC | CD | GG | KN | KK | AD |
| Kf 13 | AA | BB | CD | DN | KN | KK | DD |
| Kf 14 | AG | BB | CC | BH | KO | KK | AD |
| Kf 15 | CF | BF | CD | EE | KP | KK | AD |
| Kf 16 | AF | BB | BC | OO | KP | BK | AD |
| Kf 17 | AE | BC | BC | EE | IN | BH | AD |
| Kf 18 | AF | BB | BD | EE | KN | BK | DD |
| Kf 19 | AG | BB | BC | OO | JM | KK | DD |
| Kf 20 | AF | BF | CD | EE | LO | KK | AD |
| Kf 21 | AE | BB | CC | OO | LO | HK | AA |
| Kf 22 | BF | BF | BC | EE | KN | KK | DD |
| Kf 23 | BF | BC | CC | EO | NR | HK | DD |
| Kf 24 | AG | BC | CE | DM | OO | KK | AD |
| Kf 25 | CF | BC | BE | BB | KN | BK | AD |
| Kf 26 | AF | BB | BC | OO | OO | KK | AD |
| Kf 27 | AE | BF | CD | BO | NN | KK | AD |
| Kf 28 | CF | BB | CE | EE | LO | KK | AD |
| Kf 29 | CF | BF | BC | DG | OO | KK | AD |
| Kf 30 | AF | BB | CC | BD | KO | BH | AA |
| Kf 31 | BF | BB | CD | BD | JN | KK | AD |
| Kf 32 | FF | BC | BC | BH | KO | HK | AD |
| Kf 33 | AF | BB | CD | DH | JN | KK | AD |
| Kf 34 | AG | BB | CC | DG | JN | BK | AD |
| Kf 35 | AF | BB | AC | GG | MM | KK | AD |
| EUSTON | | | | | | | |
| Eus 01 | AC | CF | CC | DO | KN | HK | AD |
| Eus 02 | AC | CC | BB | OO | LN | BK | AD |
| Eus 03 | AC | CF | BB | DO | LN | BK | AD |
| Eus 04 | AG | CF | CC | DO | MO | BK | AA |
| Eus 05 | AF | CC | CE | DE | JO | II | AA |
| Eus 06 | CF | BB | CC | FM | NN | KK | AA |
| Eus 07 | AA | BB | CC | DG | NO | BK | AA |
| Eus 08 | AF | BB | CC | DN | LO | BB | AA |
| Eus 09 | AF | AB | CC | DN | MM | BK | AA |
| Eus 10 | CC | CF | AC | DG | LL | BK | AA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Eus 11 | AG | BC | CC | DG | LL | HH | AA |
| Eus 12 | CF | BB | CC | GN | JJ | BK | AA |
| Eus 13 | AF | BC | CC | BG | MM | BB | AA |
| Eus 14 | AF | BC | BC | DG | LR | KK | AA |
| Eus 15 | AG | CC | BC | FG | JR | KK | AA |
| Eus 16 | AG | BF | AC | EO | MN | KK | AA |
| Eus 17 | CF | BB | BE | EE | NR | KK | AA |
| Eus 18 | AG | BF | CC | OO | LN | KK | AA |
| Eus 19 | CF | BC | AC | EN | MM | BK | AA |
| Eus 20 | AF | BB | BE | OO | KN | HH | AA |
| Eus 21 | AG | BB | CC | EO | MM | HH | AA |
| Eus 22 | AG | CC | CC | EO | JL | BK | AA |
| Eus 23 | AG | BB | CC | DG | NN | HK | AA |
| Eus 24 | AF | BC | CC | OO | GI | HK | AA |
| Eus 25 | AG | BC | AE | DD | GI | BK | AA |
| TAIWAN | | | | | | | |
| Twn 01 | CC | CF | BB | BB | GI | BH | AA |
| Twn 02 | CC | CF | DD | BB | KN | BH | AA |
| Twn 03 | AF | CF | BB | BK | II | EH | AA |
| Twn 04 | CC | FF | BB | BB | GL | II | AA |
| Twn 05 | AC | FF | BB | BB | KM | BI | AA |
| Twn 06 | AC | FF | BD | BN | II | BI | AA |
| Twn 07 | CC | CF | BB | BB | II | BI | AA |
| Twn 08 | AC | CF | BB | BB | FG | BH | AA |
| Twn 09 | CC | FF | BB | BB | IN | II | AA |
| Twn 10 | CF | FF | BB | CO | JJ | BI | AA |
| Twn 11 | BF | CC | CE | BO | GI | II | AA |
| Twn 12 | BG | CF | BB | BO | GJ | II | AA |
| Twn 13 | CC | FF | BB | BB | FI | II | AA |
| Twn 14 | AF | CF | BE | BB | IK | HH | AA |
| Twn 15 | AF | CE | BB | BB | FI | AH | AA |
| Twn 16 | AC | CC | BE | BO | LL | BH | AD |
| Twn 17 | AA | CF | BB | BO | FG | HH | AD |
| Twn 18 | AC | CE | CE | BN | IK | BI | AD |
| Twn 19 | AC | CF | BB | BN | GK | II | AD |
| Twn 20 | CC | CF | BE | BN | IN | II | AD |
| Twn 21 | CC | FF | BB | BB | IN | BI | AD |
| Twn 22 | CC | CF | BB | BB | HI | BI | AD |
| Twn 23 | AC | CC | BB | BN | IL | II | AD |
| Twn 24 | AF | CF | AB | BB | II | BI | AD |
| Twn 25 | CF | CF | BB | BB | IL | BI | AD |