A MORPHOMETRIC ANALYSIS OF THE *CULICOIDES PULICARIS* SPECIES COMPLEX (DIPTERA: CERATOPOGONIDAE)

not part of registered title

BY

Richard Paul Lane, B.Sc. (London), A.R.C.S.

A thesis submitted for the degree of Doctor of Philosophy of the University of London and for the Diploma of Imperial College

Department of Zoology and Applied Entomology,
Imperial College,
London, S.W.7.

August 1979

ABSTRACT

This study assesses the value of currently available multivariate morphometric techniques in the analysis of the Culicoides pulicaris complex. This midge complex is typical of species groups which are difficult to separate into discrete clusters (species). Initially, emphasis is given to the study of eight nominal taxa in Britain: C. delta Edwards, fagineus Edwards, grisescens Edwards, impunctatus Goetghebuer, lupicaris Downes & Kettle, newsteadi Austen, pulicaris Linnaeus and punctatus Meigen. Subsequently, material from other parts of the Palaearctic Region is included.

Morphological characters of adults are tested to evaluate the nature and extent of variation. Size is rejected as unreliable, since both intraspecific and seasonal variation is excessive. Allometry of size in legs, antennae and palps is studied in large homogeneous samples of three species and the implications for taxonomy discussed.

A new system for coding wing pattern, utilising pattern elements, is developed and compared to a mechanical scanning method. The former, based on only 13 characters, is preferable, on practical and theoretical grounds, to the scanning method involving 420 characters.

In constructing a classification, two points are considered. Firstly, whether a large number of characters is required for a reliable classification and secondly, whether the recognised species are homogeneous. Using subsets from a total of 72 characters, selected by inspecting inter-character correlations, loadings on principal components, or traditional use, approximately three quarters are found to be superfluous. Using individual specimens as operational taxonomic units to test the homogeneity of species, lupicaris is rejected and another, sp. A, is recognised as new.

Percentiles about the means of each species are incorporated into canonical variate diagrams, for the accurate identification of additional specimens.

A system of classification is developed, in which species are considered as sets with indistinct boundaries. Under these conditions, transition from membership to non-membership of each set is gradual rather than abrupt. The relationship of these findings with current species concepts is discussed.

TABLE OF CONTENTS

	PAGE
ABSTRACT	1
TABLE OF CONTENTS	2
ACKNOWLEDGEMENTS	5
1. INTRODUCTION	
1.1. The genus Culicoides Latreille (Diptera: Ceratopogonidae)	7
1.2. The phylogenetic relationships of Culicoides to other Diptera	8
1.3. Importance of Ceratopogonidae as pests and vectors of parasitic organisms	9
1.4. Species concepts as applied to Diptera	18
1.5. The taxonomic problem of species complexes in biting flies	25
1.6. Species complexes in the Ceratopogonidae	32
2. OBJECTIVES OF THIS STUDY	
2.1. Objectives	34
2.2. Rationale and limitation of the approach	34
3. MORPHOLOGY	
3.1. Aspects of the term 'character'	37
3.1. The anatomy of Culicoides	39
4. TAXONOMIC REVIEW OF THE CULICOIDES PULICARIS COMPLEX	
4.1. Definition of higher taxa	56
4.2. Formal description and nomenclature	61
4.3. Geographical distribution	74
4.4. Immature stages	75
4.5. Biosystematic data on the C. pulicaris complex	76
4.6. Applied importance of the pulicaris group	78
5. ENTOMOLOGICAL METHODS	
5.1. General methods and techniques	81
5.2. Restriction of study to females	84

PAGE

6. STATISTICAL METHODS

6.1. Introduction 85

6.2. Character types 85

6.3. Primary data matrices 86

6.4. Standardisation 86

6.5. Association matrix 87

6.6. Cluster analysis 90

6.7. Ordination methods 91

6.8. Discrimination 96

7. BIOMETRIC STUDY OF TAXONOMIC CHARACTERS AND THEIR VARIABILITY

7.1. Review of previous work 99

7.2. Seasonal variation in size 101

7.3. Geographical variation 108

7.4. Variation in segments of the antennae 110

7.5. Allometry of size 119

8. ESTABLISHMENT OF NEW CHARACTERS

8.1. Cibarium 142

8.2. Chaetotaxy 145

8.3. Wing pattern 148

9. NUMERICAL CLASSIFICATION OF THE C. PULICARIS COMPLEX

9.1. Introduction 183

9.2. Selecting characters from a single body region 186

9.3. Removal of the influence of size from the analysis 196

9.4. Reduction of characters by objective means 213

9.5. Selection of variables by subjective means 251

9.6. Summary 267

10. DISCRIMINATION AND IDENTIFICATION

10.1. Introduction 279

10.2. Discrimination of species 283

10.3. Identification using the results of canonical variate analysis 294

11. GENERAL DISCUSSION		PAGE
11.1.	Evaluation of numerical methods	300
11.2.	Taxonomy of the C. pulicaris complex	311
11.3.	Species groups in Culicoides	312
11.4.	An alternative approach - species as non-disjoint sets	313
SUMMARY		327
REFERENCES		331
ADDENDA		357
APPENDIX		359

ACKNOWLEDGEMENTS

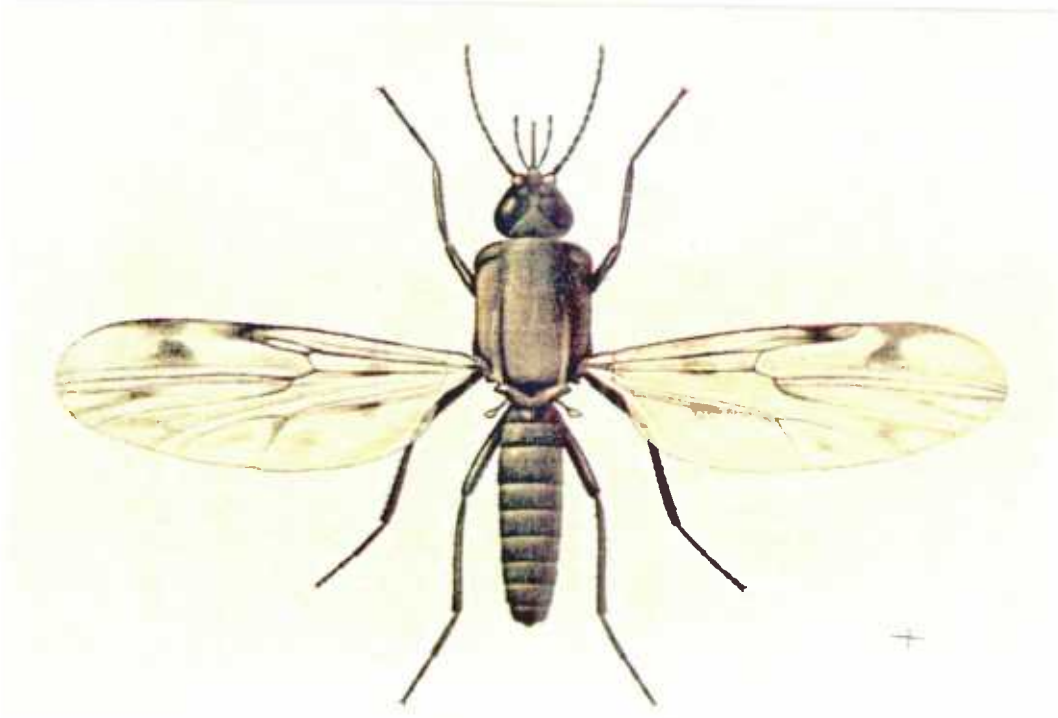
This study was carried out in the Department of Entomology, British Museum (Natural History), London. I would like to thank the Keeper, Dr. P. Freeman, not only for suggesting the topic, but also for his continual support throughout the duration of the study.

Sincere thanks are due to my Director of Studies, Mr. R. G. Davies, Department of Zoology and Applied Entomology, Imperial College. His encouragement and informative discussions throughout the course of my work have been invaluable. Thanks also go to my Supervisor, Dr. G. B. White, formerly of the British Museum (Natural History), who gave much helpful advice and sufficient rope - but never quite enough to hang myself!!

Within the Museum, colleagues in the Medical Insects Section offered much assistance, by way of constructive criticism and technical assistance. I am indebted to Dr. M Hills and Miss K. Shaw, of the Biometrics Section, for guiding me through statistical theory and for help with computer programming.

I wish to thank the following specialists for the loan or donation of specimens: Mr. J. Boorman, Animal Virus Research Institute, Pirbright; Dr. A. Campbell, University of Edinburgh; Dr. J. Clastrier, Muséum Nationale d'Histoire Naturelle, Paris; Dr. R. M. Gornostaeva, Martsinovsky Institute of Parasitology and Tropical Medicine, Moscow; Prof. A. V. Gutsevich, Zoological Institute, Leningrad; Dr. S. Kitaoka, National Institute of Animal Health, Tokyo; Mr. E. C. Pelham-Clinton, Royal Scottish Museum, Edinburgh; Prof. Y. Wada, Nagasaki School of Medicine, Nagasaki; Dr. W. Wirth, U.S. National Museum, Washington.

Finally, I must thank my wife, Maureen, who despite her advanced pregnancy, still managed to sit close enough to the typewriter to produce this thesis!



FRONTISPIECE

Culicoides pulicaris (Linnaeus).

(Drawing by Terzi from Edwards, Oldroyd & Smart, British Bloodsucking Flies).

Section 1. INTRODUCTION

1.1. THE GENUS CULICOIDES LATREILLE (DIPTERA: CERATOPOGONIDAE).

The genus Culicoides is composed of approximately 1000 species of small biting midges belonging to the family Ceratopogonidae. Culicoides have a world wide distribution, except for New Zealand and Southern Chile. Most species are small, with a wing length averaging 1.5mm. and reaching a maximum of approximately 2mm. The wings are often characteristically patterned.

The females have well developed biting mouthparts in most species. The males do not bite. Some species having females with atrophied mouthparts, associated with an autogenous life cycle, have been found in the Canadian Arctic (Downes, 1970). In most species, the females need a blood meal for the development of the eggs, although some, as noted above, are completely autogenous. Others will lay a first batch of eggs autogenously and require a blood meal for maturation of subsequent batches (e.g., C. impunctatus, one of the species studied here (Service, 1968)).

The majority of species are crepuscular and biting activity continues into the night. Others however, bite during the day. Their attack is usually inhibited by wind speeds of greater than 3 metres per second. Those species living in exposed environments, such as salt marshes, will fly during higher wind speeds. Host specificity is not well understood for most species, although it appears that the majority of species have preference for one primary host and a range of secondary hosts (Kettle, 1962). The tendency of adults to collect in large numbers and bite unceasingly makes them extremely troublesome to man and domestic animals. Their unpleasant attention to man has led to the coinage of a number of vernacular names. In Britain they are called midges; in the West Indies, southern U.S.A. and Australia they are somewhat ambiguously called sandflies (a term usually reserved elsewhere for Phlebotominae);in French Canada as brûlots; as arabis (midi) and muchits (Blavais) in France; jejenes in Cuba; punkies or no-see-ums in U.S.A.; maruins in Amazon basin; nonos in Tahiti; makanagi or nukaga in Japan; mout-mout in French Guinea.

Mating may take place on the ground (e.g., C. melleus,(Linley & Adams, 1972)) or in swarms (Downes, 1955). The eggs are elongate

ovoids in shape, and are laid on moist substrates. Immersion in water does not impede their development. The eggs are usually hatched after 5 - 9 days, halophilic species often after much less (2 - 3 days) and in one species, C. grisescens, a period of 205 - 223 days elapse before hatching (Parker, 1950). The latter case probably represents a period of hibernation. In temperate regions, over-wintering takes place as a fourth instar larva.

The larvae are apneustic, swimming in a characteristic eel-like manner. Edwards (1939) quotes a description of a larva and pupa of what is apparently Culicoides, given by a Rev. W. Derham, as long ago as 1712. In more recent times, the larvae and other immature stages have been described; West African Culicoides by Carter, Ingram and Macfie (1920, 1921); British Culicoides by Hill (1947), Kettle & Lawson (1952); European species by Lenz (1934), Mayer (1934a, 1934b) and Thienemann (1954). The larvae may be generally considered aquatic and are found in a wide range of habitats: pools, algal mats, wet soils in marshes and bogs, riverbanks, tree-holes, rotting vegetable matter, animal dung, and sand and mud periodically soaked by sea water. To date, only one species has been found in flowing water (Fredeen, 1969).

The pupae have well developed pupal horns for respiration. Duration of the pupal stage is generally short, within a period of 2 - 5 days. Although the pupae are usually motionless, those in tree-holes have been observed to control their depth by extending and contracting the abdomen.

Remm (1976) described a few species of Culicoides from amber, and suggested that the genus was present about 100 million years ago, in the Upper Cretaceous period. This is particularly interesting because, as most extant Culicoides bite warm blooded animals, their presence in the Late Cretaceous infers the presence of warm blooded vertebrates during this period.

1.2. THE PHYLOGENETIC RELATIONSHIP OF CULICOIDES TO OTHER DIPTERA

Culicoides and other Ceratopogonidae are typical of the dipterous suborder Nematocera in that they possess long antennae, consisting of many similar segments, palps of several segments, and wings with several longitudinal veins, but without a discal cell.

Oldroyd (1977) has rejected the traditional suborders Nematocera,

Brachycera and Cyclorrhapha, on the grounds that they are unnatural. In their place he has regrouped the Diptera into three new suborders (Superstata, Madescata and Arescata), based on general ecology and habit, rather than on structural grounds. Under this scheme, Culicoides (and other Ceratopogonidae) fall into the suborder Madescata, by possession of aquatic larvae and mandibulate, bloodsucking females.

Hennig (1968) place the Ceratopogonidae together with the Simuliidae and Chironomidae in the Chironomoidea, stating that the relationships of these three families are unknown. This throws doubt on the traditional view that the Chironomidae and Ceratopogonidae are sister groups. Female Simuliidae are vertebrate blood feeders, whereas adult female Chironomidae have non-biting mouthparts (except for one primitive genus Archeoclus Brundin).

Downes (1970) has suggested that the Diptera originated as bloodfeeders. Although he discusses the hypothesis extensively in relation to feeding on vertebrates, he notes (p. 253) that the mouthparts did not necessarily originate in this context, but may have been used for feeding on insects. The Ceratopogonidae exhibit a great range of feeding habits, apparently exceeding that of all other families of biting flies taken together (Downes, 1970:243).

The significance of the genus Culicoides in the evolution of the Ceratopogonidae has been expressed in opposite ways. Downes (1977a) suggests that Culicoides, as a bloodsucking genus, is one of the most primitive in the family. Remm (1975), on the other hand, puts the insectivorous Ceratopogonids close to the origin of the group and Culicoides as one of the most advanced genera. One of the essential points of the controversy is whether feeding on insect blood or vertebrate blood is the most primitive condition. Both hypotheses accept that aphagia is an advanced condition. These alternative views do not substantially change the relationship of Culicoides within the family, only the supposed origin of the Ceratopogonidae. Relationships of Culicoides within the Ceratopogonidae are discussed in Section 4.1.1.

1.3. IMPORTANCE OF CERATOPOGONIDAE AS PESTS AND VECTORS OF PARASITIC ORGANISMS

Ceratopogonids are economically important in several respects. Some are vectors of pathogenic organisms to both man and domestic animals. Little is known of the parasites transmitted by these

midges to wild animals. Owing to their large numbers and persistent biting, many species are serious pests and may affect the use of extensive areas of land. The economically important species are those which feed on the blood of vertebrates and belong to four genera: Culicoides, Leptoconops, Forcipomyia (Lasiohelea) and Austroconops. The vector status of the three bloodsucking genera other than Culicoides is poorly known, but nevertheless, many cause considerable annoyance to man and his animals.

Before discussing details of parasite transmission, it is necessary to outline some epidemiological terms, as they are important in determining the significance of recorded vector capability of Ceratopogonids.

Parasite transmission by a vector may be of two types, cyclical and mechanical. In cyclical transmission, the insect is an obligatory vector of the pathogenic agent, thus becoming an indispensable link in the cycle of the disease, or a reservoir of the infective agent for an extended period of time. The parasite undergoes cyclical modifications in the body of the insect and may or may not increase in number. In mechanical transmission, the insect plays a more passive rôle and is mainly an accidental vector of the parasite. However, this apparently casual transmission may play an important rôle in the dissemination of the disease.

In epidemiological studies it is important to distinguish between the recovery of a parasite from a midge, and the positive incrimination of it in the transmission of the parasite. Parasites can be found in any bloodsucking Ceratopogonid that feeds on an infected host. It does not follow that the parasite will complete its development (or remain alive), or that the species of midge is a suitable host for the parasite. Many arboviruses have been isolated from midges, but relatively few of these have been shown conclusively to be the natural vectors of the viruses.

1.3.1. Culicoides

This is the most important genus of the Ceratopogonidae with respect to disease transmission. In discussing the effect of these midges on man and his animals, the subject may be divided into nuisance and dermatoses caused by bites, and transmission of parasites (which may be bacteria, protozoa or nematodes).

Nuisance and Dermatoses.

The large numbers of biting Culicoides midges have had considerable impact on the economic development of some regions. Linley & Davies (1971) give a comprehensive account of the relationship between 'sandflies' and tourism in Florida, the Bahamas and Caribbean area, where severe human annoyance by Culicoides has been a major factor affecting the growth of the tourist industry. Some $8 million is spent annually on Culicoides control in Florida alone (Linley, personal communication). An interesting case is discussed by Wirth & Arnaud (1969) where C. belkini Wirth & Arnaud has recently become a pest in French Polynesia. The species appears to be endemic, previously existing in low numbers, but recent land development has provided suitable breeding habitats by the creation of areas of shallow brackish water.

Agricultural work is often rendered difficult or impossible because of the attention of biting midges. Szabó (1965) describes an area of Hungary where agricultural workers were unable to open their eyes for several days because the lids were swollen after bites from Culicoides. Draught horses were rendered uncontrollable due to the persistent biting of C. nubeculosus (Meigen). Hill (1947) suggests that the backward state of croft farming in Western Scotland may, in part, be attributed to the activities of C. impunctatus Goetghebuer. The annoyance of cattle by C. brevitarsis Kieffer (approximately 5000 bites per hour) in Australia was described by Standfast & Dyce (1968). They suggested that this may result in losses (presumably in weight gain), due to restlessness.A number of papers have been published discussing dermatoses due to Culicoides bites, the most thorough of which is a series by Sherlock (1963 and 1964). He reported a public health problem in Brazil, of an extensive dermatitis as a reaction to the bites of C. paraensis (Goeldi). An increase in biting Culicoides in and around the city of Salvador, was correlated with decreasing standards of sanitation (refuse accumulation and water pollution), and the end of a control campaign against Aedes aegypti breeding sites. Arean & Fox (1955) describe severe reactions to bites of C. furens (Poey), and Arnaud (1956) reports eczema in some Japanese caused by the bite of C. erairi Kono & Takahashi, and sores that took some three to four months to heal, following bites of C. obsoletus (Meigen).

In Australia, a condition of horses known as Queensland Itch is described by Riek (1954). He determined the cause to be a reaction to the bites of C. brevitarsis (Taylor) (= robertsi Lee & Reye), and not to the activity of Onchocerca, as had previously been thought. The dermatitis of horses that occurs in India, the Philippines and the United States, is thought to be identical to that outlined by Riek.

Transmission of Viruses

The main veterinary importance of Culicoides in tropical and subtropical regions is as vectors of viruses. Approximately 26 arboviruses have been isolated from Culicoides and many of them are closely associated with ungulates, particularly cattle, goats and sheep (Boorman, In press).

DuToit (1944) was the first to incriminate Culicoides in the transmission of Bluetongue disease, by recovering the virus from wild caught midges in South Africa. The disease is endemic in East and Southern Africa, where the principal vector is C. imicola Kieffer (= C. pallidipennis Carter, Ingram & Macfie). Recent work on the rôle of Culicoides in the epidemiology of the disease in Kenya is discussed by Walker & Davies (1971) and Walker & Boreham (1976). Bluetongue is present in many parts of the Middle East (Boorman, 1974; Braverman et al., 1971).

The isolation of Bluetongue virus from sheep in Texas (Price & Hardy, 1954) initiated much research in the U.S.A.. The colonisation of the vector, C. variipennis Coquillett by Jones (1960) was a landmark in subsequent laboratory transmission studies (Foster et al., 1963).

DuToit (1944) experimentally transmitted the virus of African Horse Sickness by intravenous injection of an emulsion of wild caught Culicoides into a horse. This disease is endemic in eastern and southern Africa and the Sudan, but rare in West Africa. Kettle (1965) outlines the radiation of this disease into Iran, Afghanistan, W. Pakistan, India, Turkey and Cyprus, putting some thirteen million horses at risk (Huq,1961). Wind borne Culicoides have been suggested as the route by which this disease entered Spain, via Morocco, in recent years (Boorman, 1979).

Culicoides have long been suspected as vectors of Bovine Ephemeral Fever in Australia (Standfast et al., 1972), but this

has not been confirmed (Doherty et al., 1972, 1973). However, Davies & Walker (1974) isolated B.E.F. virus from wild caught Culicoides (principally C. kingi Austen) during an outbreak in Kenya, and found the C.schultzi (Enderlein) group (including kingi Austen) the most common feeder on cattle at disease sites (Walker & Boreham, 1976).

A North American disease of Deer - Epizootic Haemorrhagic Disease - has proved to be transmitted by Culicoides (Boorman & Gibbs, 1973). Although confined to North America, this disease is of potential economic importance with the increase in farming red deer in Scotland. It is particularly significant that the disease can be transmitted by British species of Culicoides (Boorman & Gibbs, 1973).

Karstad et al. (1957) reported the recovery of Eastern Equine Encephalitis from Culicoides in Georgia (U.S.A.) and cite a private communication from R. Levi-Castillo, stating that Venezuelan Equine Encephalitis had been isolated from Culicoides in Ecuador, during an outbreak affecting both man and animals. Although not directly incriminated, C. arubae Fox & Hoffman and Leptoconops kertezi Kieffer were suspected as possible vectors, permitting the movement of Venezuelan Equine Encephalitis from South and Central America into Texas (Jones et al., 1972).

Many other arboviruses have been isolated from Culicoides, e.g., Button-Willow and Akabane (Hartley et al., 1975), or have been transmitted in the laboratory, e.g., Main Drain (Mellor et al., 1974), but the epidemiological significance of these results awaits further research.

The transmission of Fowlpox virus by Culicoides was first suggested by Shiraki (1913) and later by Tokunaga (1937), on the slender evidence that Culicoides were thought to breed in poultry pens. Subsequent study revealed that the principal poultry biting Culicoides - C. arakawae (Arakawa) - actually bred in paddy fields.

In a laboratory study by Seledtov et al., (1969), Culicoides were shown to be possible vectors of West Nile Virus, a dengue-like virus of birds.

Transmission of Bacteria

The rôle of the genus in transmission of bacterial infections of domestic animals has also been investigated, but with inconclusive results. Turner et al. (1963) attempted to incriminate

Culicoides in transmission of Infective Synovitis of Poultry. They demonstrated limited transmission when freshly engorged C. variipennis (which had fed on diseased birds) were macerated and innoculated into healthy chickens. Transmission did not occur after 24 hours between feeding and maceration. They concluded that neither direct mechanical nor cyclical transmission by Culicoides occurs.

Nielsen (1971), working on Culicoides biting cows, suggested them as being possible vectors of the bacteria responsible for Summer Mastitis. Sorenson (1974) isolated the primary agent of Summer Mastitis (Corynebacterium pyogenes) from Culicoides feeding on the teat of a healthy cow (note, only one individual was tested). He went on to suggest that the most likely rôle of Culicoides in the epidemiology of the disease is that they are responsible for teat injuries in cows. It may be that these injuries encourage the attention of Hydrotaea irritans Fallén, a known vector of Summer Mastitis bacteria.

Transmission of Protozoa

In general, the Protozoa tansmitted by Culicoides are parasites of birds. The notable exception is Hepatocystis kochi (Laveran), a malaria-like parasite of monkeys, transmitted by C. adersi Ingram & Macfie (Garnham et al., 1961; detailed review in Garnham, 1966). The bird parasites are members of the genera Parahaemoproteus Bennett, Garnham& Fallis and Akiba Bennett, Garnham & Fallis. Prior to Garnham (1966), species of these genera transmitted by Culicoides were recorded as Haemoproteus Kruse and Leucocytozoon Danilewsky.

Transmission of Protozoa by Culicoides was first shown by Fallis & Wood (1957), following the development of Parahaemoproteus nettionis (Johnston & Cleland) in an unidentified species of Culicoides. They, and other Canadian workers, produced a series of papers on the transmission of Parahaemoproteus in both wild and domestic ducks (Fallis & Bennett, 1960, 1961a, 1961b). A list of Culicoides and the species of Parahaemoproteus they transmit is given in Greinier & Bennett (1977), who suggest that each Protozoan parasite may be transmitted by more than one species of Culicoides, and that one species of Culicoides may transmit more than one species of Haemosporidian.

Unpublished work by Bennett indicates that avian trypanosomes

will develop in, and can be transmitted by Culicoides (Kettle, 1965).

Akiba et al. (1959) and Akiba (1960) showed Culicoides arakawae (Arakawa) to be the vector of Akiba caulleryi Mathis & Leger (Fam.: Leucocytozoidae). This parasite is responsible for Leucocytozoonsis, economically a very important disease of poultry in the Far East. Frequently the parasite is so pathogenic that a large percentage of whole flocks die. The disease is also known as Bangkok Haemorrhagic Disease, and is capable of killing a bird overnight (Garnham, 1966). The development of the parasite in both host and vector is outlined by Garnham (1966) and Morii & Kitaoka (1968a, 1968b). In a laboratory study, Morii et al. (1965) incriminated C. circumscriptus Kieffer and C. schultzi (Enderlein) as subsidiary vectors in Japan.

Garnham (1966) suggested that the discovery of Culicoides as vectors of Leucocytozoidae may help to explain the mystery of the transmission of this family in regions where the usual vector, Simulium spp, is absent. He cites an example of Kisumu in Kenya where Culicoides probably replace Simulium as the natural vectors of local species of Leucocytozoidae.

In the early part of this century, Culicoides were suspected vectors of Kala Azar (Leishmaniasis) by Fletcher (1924) and Christophers et al. (1925). However, the authors were unable to find live Leishmania in the gut of C. maculatus Shiraki three days after a blood meal, and concluded that Culicoides are unlikely to be important vectors.

Transmission of Nematodes

This aspect of parasite transmission by Culicoides has received considerable attention as it directly affects man, by the transmission of Filariasis due to Dipetalonema Diesing and Mansonella Faust.

Sasa (1976) gives a very good review of the complex story following the discovery of the Culicoides vectors of Dipetalomena perstans (Manson) and D. streptocerca (Macfie & Carson) by Sharp (1928) and later Duke (1954, 1956 and 1958); Hopkins (1952); Hopkins & Nicholas (1952) and Chardrome & Peel (1951). Neither of these parasites are pathogenic in man. D. perstans (Manson) is distributed throughout much of Africa, West Indies and tropical South America, while D. streptocerca (Macfie & Carson) is restricted to

the tropical rain forest zone of Africa. C. milnei Austen and C. grahami Austen are vectors of D. perstans in Africa and C. grahami and C. inornatipennis Carter, Ingram & Macfie are vectors of D. streptocerca.

Mansonella ozzardi (Manson) is a new world filaria recorded from some West Indian islands, Central America and South America, where it is distributed amongst peoples (principally Amerindians) living in the interiors of these areas. The vector in the West Indies was shown to be C. furens Poey by Buckley (1934) and C. phlebotomus (Williston) by Natham (1978), but in Brazil, the vector is reported to be the Simulium amazonicum complex.

Gibson & Ascoli (1952) investigated the potential rôle of Culicoides as vectors of Onchocerca volvulus (Leuckart) in Guatemala and found although C. paraensis Goeldi and C. stigmalis Wirth ingest the microfiliae, none complete their development, thus the longevity of the host fly is reduced.

Steward (1933) showed that Culicoides transmit Onchocerca cervicalis (Railliet & Henry),which is associated with fistulous withers of horses. He failed to transmit this filaria with Simulium or Haematopota, but did so with C. obsoletus (Meigen) and C. nubeculosus (Meigen). Steward traced the development of the worm in C. nubeculosus and found that it took 25 days before reaching the infective stage for the equine host.

In Australia and Malaysia, heavy infection rates of Onchocerca gibsoni (Cleland & Johnston) have been found in cattle, and these may lead to carcasses being condemned. In his investigation to find the vector, Buckley (1938) collected 20 species of Culicoides and Lasiohelea off cattle. Of these, C. punoens de Meijere and C. schultzi (Enderlein) (= C. oxystoma Kieffer), were the most abundant, but the natural infection rate was low (0.3%). Even after feeding on infected cattle, the infective rate did not rise above 1%. However, as Culicoides can be collected from cattle at the rate of 500 per hour, a very low infection rate in the vector would be adequate to maintain a high parasite rate in the host (Kettle, 1965).

Spratt, Dyce & Standfast (1978) observed the development of larval Onchocerca sweetae Spratt & Moorhouse in the thorax of Culicoides (recorded as species 'M'), collected whilst feeding on water buffalo. They concluded that the Culicoides sp. 'M' was the natural intermediate host of the buffalo parasite in Australia.

1.3.2. Leptoconops

Leptoconops is a widely distributed genus throughout the tropical and subtropical regions of the world. Wirth & Atchley (1973) review their biology, give a general outline of taxonomy and provide an extensive bibliography. The vicious daytime biting of this genus has been described many times, and is known to have seriously impaired both agricultural work and the development of tourism (Linley & Davies, 1971).

Kimura (1959) reported dermatitis in Japanese patients, followino the bites of L. nipponensis Tokunaga, and Howell (1970) describes a condition called Leptoconops-mange of sheep in S. Africa.

1.3.3. Forcipomyia (Lasiohelea)

This is the only subgenus of Forcipomyia which has been found biting man, the remainder are flower feeders, with a few species attacking other insects. Lasiohelea have been reported as very troublesome biters over much of the Old World (particularly Asia). F. (L.) taiwana Shiraki is a serious pest in China (Chang & Wang, 1958; Chung et al., 1964) and Taiwan (Shiraki, 1913). F. (L.) stimulans (de Meijere) is a pest in S.E. Asia, especially Indonesia. F. (L.) stylifer (Lutz) is the only recorded New World pest species. Ortiz (1952) has outlined its pest status in Venezuela. Chan & Saunders (1965) record F. (L.) anabaenae Chan & Saunders causing a mild fever after severe attacks in Singapore. The fever was thought to be a reaction to injected protein, since the collector later lost sensitivity to the bites, and the fever subsided after two days. In Queensland, F. (L.) townsvillensis (Taylor) is a serious pest of man and domestic animals and was suggested as a possible vector of Bovine Ephemeral Fever by Lee et al. (1962). They also suggested that 'worm nodule' of cattle (Onchocerciasis), is widespread in Australia, and is likely to be Culicoides transmitted (as it is in Malaya). Moorhouse (1978) found the parasites responsible for this condition, Onchocerca gibsoni Cleland & Johnston, O. gutturosa Neuman and O. linealis Stiles, in the guts of F. (L.) townsvillensis feeding on an infected cow from Queensland.

The only incrimination of Lasiohelea in the transmission of human disease is that of Wu & Wu (1957), who isolated the virus Japanese B encephalitis, from F. (L.) taiwana Shiraki.

1.3.4. Austroconops

This is a monotypic genus, about which remarkably little is known, even its taxonomic position is uncertain (Boorman & Lane, 1979). It has been recorded as biting man in Western Australia (Wirth & Lee, 1959).

1.4. SPECIES CONCEPTS AS APPLIED TO DIPTERA

The 'Species Concept' has been the subject of much discussion by biologists. It is not the purpose of this outline to further this controversy, but to present some of the attitudes in the study of the Diptera. To this end, two aspects of the species problem are discussed. Firstly, the theoretical basis of two conflicting concepts (Typological and Biological species concepts), and secondly, how these theoretical ideas are applied.

The range of species concepts used by recent and current workers on the order Diptera is probably as great as that found anywhere in the animal kingdom. They range from the strict adherents to the typological species concept, to those like Dobzhansky, who used examples of Drosophila to develop his interpretation of the biological species concept.

1.4.1. Theoretical Aspects

Two basic concepts have been, or still are, used by Dipterists, the typological and biological concepts.

For the adherents of the typological species concept (= morphological concept of some authors), the degree of morphological dissimilarity is the primary criterion for the species status. This concept was elaborated by Linnaeus and his followers (Cain, 1954) and holds that there are a limited number of universals, or types, of organsim. Individual variation is seen as the result of imperfect expression of the essential qualities of each species. The reasoning used by proponents of this concept has been paraphrased by Mayr (1970) as "Natural populations considered by general consent to be species, are morphologically distinct. Morphological distinctness is thus the decisive criterion of species rank. Consequently, any natural population that is morphologically distinct must be recognised as separate species". For the essentialist

therefore, morphology is of paramount importance in delimiting the 'essence' of their species.

This concept does not currently have many strict adherents, but was the main hypothesis of past taxonomists. For example, Robineau-Desvoidy worked on Tachinidae during the 1860's and described many species, of which some 250 were later found to be synonyms of one species, Phryxe vulgaris (Fallén). This polyphagous parasite is very variable in details of colouring, size and wing venation, and each of the 250 'species' represented minute morphological variations in these characters. As recently as the 1930's, the typological species concept was advocated by Townsend (1935) who stated his beliefs most comprehensively: "Two species, the progeny from whose crossing is sexually infertile, belong to separate natural genera. All those species which can produce fertile crosses belong to the same natural genus" (p. 38) and "such differences almost uniformly mark species groups or physiological genera, whose members differ among themselves practically only in colouration, size and minutae of structure" (p.59). Ironically, although Townsend adhered to an outdated and unpopular concept, he was ahead of his time in proclaiming that the study of chromosomes would define, and resolve, the identity of most taxa.

Oldroyd (1966) challenged the view that the typological species concept was no longer widely used in taxonomy. He suggested that many taxonomists were compelled to retain this hypothesis through the necessity of designating types when describing a species, to comply with the rules of the International Commission of Zoological Nomenclature. Crowson (1970) supported this view and wrote: "The current convention that a single specimen, the holotype, is the only satisfactory basic criterion for a species would be difficult to justify logically on any theory but that of Special Creation". It is of interest that this theory of special creation was fundamental to the early proponents of the typological species concept, since it stated that a fixed number of discrete species were formed, and that it only remained for taxonomists to discover the intrinsic or essential quality of each of these!

Most Dipterists now follow the biological species concept developed by Dobzhansky (1937) and Mayr (1942), defined by the latter in 1969 as: "Species are groups of interbreeding natural populations

that are reproductively isolated from other such groups". This definition stresses that species consist of populations, and have genetic continuity. The species concept is termed biological not because it deals with biological taxa, but because the definition is biological. It expresses concepts that separate biological classification from that of inanimate objects and associated restrictions, particularly essentialism (Hull, 1965). Species may be defined by their relational properties rather than because of any intrinsic property. Very few species are known to fit the definition and Scudder (1974) suggests that the data required for its use are virtually impossible to collect. In the Diptera however, there are a few studies to which the concept has been rigorously applied, e.g., in Drosophila (Dobzhansky, 1951) and in Mosquitoes (see White, 1979).

Although the concept outlined by Mayr (1969) is widely accepted, many authors have added their own qualifications or extensions, e.g., Paterson (1978) who worked on Muscidae, accepts in principle the biological definition of a species, but changes the emphasis on isolating mechanisms to give the following definition: "members of a species share a common specific mate-recognition system".

Scudder (1974) rejects the single definition and maintains that, while that of Mayr and Dobzhansky solve the problems of the typological species concept, they do not solve everything. He discusses two shortcomings in particular, (a) the inapplicability in some circumstances, e.g., ring species, species showing introgressive hybridisation, and asexually reproducing species; (b) the general operational inapplicability. This second aspect has been noted by many biologists (Dipterists included), and has been dealt with in a formal manner by Sokal & Crovello (1970). The main point of their argument is that most species recognised at present, and also those now being described, are still determined by reference to morphology and this alone. Rarely are data available for a decision on the biological nature of a species. Cain (1954) was among many who suggested that a variety of species should be recognised, depending on the nature of the evidence used to recognise them.

1.4.2. Working Concepts

As noted above, Cain (1954) and Scudder (1974) suggest that a

number of different species concepts should be used according to the particular biological problems of the animals being studied. It is possible that the diversity of species concepts used in the study of Diptera reflects the application of varying practical techniques, rather than a marked difference in theoretical models.

Crowson (1970) surveyed the problem from a different angle, in that he does not discuss different concepts, but the range of criteria (in all, five) which are commonly used. In this sense he obliquely suggests that those concepts held by biologists are a function of the technique available for the study of any group of animals. There are many alternative lists of types of species, but the criteria of Crowson (1970) will be used here, to demonstrate the range of working species concepts employed by Dipterists.

Museum Criteria

These are the most commonly applied criteria, and are based mainly on morphology. Crowson suggests that among members of one species, there is normally a limited and continuous variation in characters of structure and pigmentation, whereas a discontinuity in one or both these respects will normally show itself when members of two different species are compared. These criteria were also given by Mayr (1969, p. 21). Species recognised by this method have often been termed morpho-species - they are only morphological in that the data used are morphological, not that the morphological differences are used as the sole criteria. There is an important difference between using this similarity as a primary criterion for species rank and using it, hopefully with other evidence, as an indication of the specific status of a population. Thus, these criteria commonly accept the theoretical model of one definition (biological species concept), but use evidence which has been associated with another (typological species concept). The biological species concept does not imply any morphological distinction between species, but nevertheless, for the overwhelming majority of species, genotypic isolation may be inferred from their phenotypic discreteness (Simpson, 1961).

Although relatively few species have been studied in sufficient detail to fulfil the requirements of the biological concept, the idea is commonly used to equate morphological distinction with reproductive isolation. It is arguable whether species recognised by the criteria

given above necessarily demonstrate the discontinuity which Crowson and Mayr suggest is typical. Many morphologically overlapping species are recognised, for example, the pulicaris group studied here.

Unlike some, these criteria do not offer any reliable means of deciding whether or not two specimens belong to the same species by comparison alone. For a satisfactory comparison, several specimens are needed to overcome such problems as geographic variation. It would appear that most taxonomists evaluate variation both within and between species either intuitively or, more reliably, by statistical means. Therefore they hold some concept of the expected ranges of morphological diversity associated with these types of variation, and are able to assign individual specimens to taxa.

One of the main advantages of species recognised by the museum criteria is that of easy comparisons. It is possible to compare several forms of one genus in a week by the museum method, which would take years of research using ecological, physiological or genetical characters. The museum criteria therefore allows a provisional classification to be constructed for groups of flies, for which it would otherwise be difficult to obtain the necessary biological information to satisfy the biological species concept.

Ecological Criteria

Members of different species usually show variation in habit and behaviour. This aspect of biology is often the first to be discovered or experimentally tested in the field, e.g., there have been many studies of the monospecificity of swarms of male flies, particularly Culicidae and Chironomidae (Downes, 1958, 1969).

Those species of Diptera which have an intimate association with plants have often been the subject of ecological studies, and species are frequently defined in terms of the host-plant association. This was the principle employed by Barnes' study of the Cecidomyidae, in which numerous host transfer experiments were carried out. These experiments attempted to establish specific status of populations of midges found on a restricted range of host plants, by showing that they were unable to develop on plants harbouring other potential species of flies. Hering worked during the 1940's and 1950's on the leaf-mining Agromyzidae, and although he never explicitly discussed his species concepts, it is most likely that he first

recognised species by their host association and subsequently sought structural characters to substantiate them.

Differences in larval habitat have been found in the sibling species of Anopheles gambiae complex, in which the adults are either difficult or impossible to identify morphologically (Davidson et al., 1967).

Physiological Criteria

Within a species there is normally the same kind of limited and continuous variation in physiological and biochemical characters, as there is in structural ones.

Physiological differences may be related to ecological requirements as shown by Davidson et al. (1967), in larvae of Anopheles gambiae complex. First stage larvae of five species were immersed in a saline solution, resulting in the death of three species and the survival of two. These same authors were able to separate the morphologically indistinct A. gambiae and arabiensis of the gambiae complex, by chromatic treatment of pterine pigments.

Genetic Criteria

Because the biological species concept is phrased in genetic terms, these criteria are held to be the most reliable. Three basic techniques have been employed in establishing the genetic basis of many species studied in the Diptera: cross sterility test; cytology; isoenzyme studies.

Sexual crosses between members of one species are normally fully fertile, whereas interspecific crosses usually yield infertile offspring, or none at all. This test has been used extensively to establish the affinities of mosquitoes and to identify wild specimens by crossing them, or their progeny, with reference stocks. An interesting and often overlooked aspect of cross-sterility tests is emphasised by White (1979); whereas hybrid sterility signifies a post-mating barrier (sensu Mayr, 1970) between species, hybrid fertility is an equivocal condition. Allopatric species which do not have any premating barriers may produce fertile offspring when tested. Hence, this test is most useful for distinguishing between sympatric species. Such a practical observation complies with the theoretical objections raised by Dobzhansky (1970) to his own species definition, whereby species share a 'common gene pool'.

Therefore the definition is most useful in delimiting sympatric or parapatric species, but difficulties may be expected if it is applied to populations living in geographically separated localities.

Drosophila pseudobscura Frolova and persimilis Dobzhansky & Epling were recognised solely as a result of breeding experiments, after geneticists failed to obtain fertile offspring when the two species were crossed. Subsequently, further study showed some small morphological differences between species (Dobzhansky, 1951 for review).

Fortunately for dipterists, many species of Diptera possess 'giant chromosomes' and consequently have received considerable attention by cytogeneticists. Numerous cryptic species have been discovered and maps have been prepared of the banding pattern of many more. Studies of this type have been carried out in the Simuliidae, Chironomidae, Sciaridae, Culicidae and Drosophilidae. The problems of sibling species have benefitted enormously from chromosomal studies, and these will be discussed in more detail in the following section.

The use of isoenzymes in characterising species has also become very popular in recent years. The technique uses the population frequencies of polymorphic enzymes to recognise demes or species, and to give an estimate of the genetic distance between different species. The Culicidae, with over 100 papers published, and the Simuliidae, in particular have benefitted from the application of such sophisticated techniques. The above outline is by no means a complete directory of working concepts, for it is possible to produce numerous lists of these specialised 'concepts'. Scudder (1974) lists at least twelve. Each proposed working concept fulfils different interests or requirements, e.g., the superspecies of Hennig (1966), and Amadon (1966) is of interest to phylogenetic systematics, to describe a collection of fully differentiated sister species, which retain an allopatric distribution. However, such a concept is of limited use to the systematist with little phylogenetic evidence available. The Palaeospecies, beloved of palaeontologists, is not commonly applied in the study of the Diptera which have an incomplete fossil record and occur in relatively few geological strata.

However, in the Diptera at least, although there are many practical definitions, there is little disagreement over the theoretical basis, most accepting the principles of Mayr (1969) and Dobzhansky (1970

One practical concept that uses many of the criteria outlined above and has attracted the attention of many dipterists, is the 'sibling species'. This is the subject of the next section.

1.5. THE TAXONOMIC PROBLEM OF SPECIES COMPLEXES IN BITING FLIES

Groups of species whose constituants are morphologically indistinct are frequently encountered by the taxonomist. Mayr (1969) outlines their occurrence in the animal kingdom, noting that they seem especially common in the Diptera, the biting flies in particular. This may be due to the attention such insects have received rather than some intrinsic quality, although species complexes appear to be more common in those insects with well developed olfactory senses.

Such species groups do not constitute a type of species set apart from others, but are merely near the invisible end of a broad spectrum of diminishing morphological differences between species. Although too few of these complexes have been adequately analysed, to facilitate generalisations, for the taxonomist the problems are similar: a number of taxa (variously ranked as biological races, 'subspecies', or varieties) are recognised, which are difficult to separate on morphology alone. Normally, the taxa are morphologically overlapping, so that individual specimens are difficult to identify with precision or certainty.

The term 'sibling species' was suggested by Mayr (1942) to describe such similar species and was defined as "morphologically similar or identical natural populations that are reproductively isolated". Dobzhansky (1972) paraphrased this definition as "pairs or groups of species that are morphologically indistinguishable, or distinguishable with difficulty, are called sibling species". Although this definition omits the important biological aspects of reproductive isolation (a cornerstone of Mayr's biological species concept), it is useful to the practising taxonomist to aid recognition of a possible pair of sibling species. Many definitions of sibling species exist, the practical difference between them lying in the emphasis that each places on the degree of morphological distinction between the components.

Downes (1973), in discussing the impact of sibling species on taxonomy, suggested that the number of existing species of Simuliidae

is not merely slightly greater than the number now recognised, but more like five times as great. Furthermore, he also suggests that this phenomenon is not restricted to the blackflies, for studies have shown similar results in other families of Diptera.

The problem of sibling species has encouraged a wide range of techniques to test the genetic composition of the taxa concerned, and also to demonstrate the reproductive isolation of populations. Even though powerful and complex biochemical and cytological methods have been used to investigate and resolve the problem, a recent review of species complexes in insect vectors of disease (W.H.O., 1977) recommended that morphology should remain the basis of species discrimination. Presumably, this is not a return to the typological species concept, but a suggestion that wherever possible, morphological characters should be sought to facilitate easy recognition of species in the field. Referring to the numerous and difficult sibling species in Anopheles, White (1977) comments that they "seldom lack diagnostic morphological characteristics, although their distinctive specific features are often very small, compound or non-absolute". Rothfels et al. (1978), working on chromosomal differences in North American Simuliidae, have suggested that a priori it is to be expected, and empirically it has been found, that morphological and biological differences between siblings can be found, once certifiably pure material of all stages is obtained. They go on to suggest that, from their experience it should be possible to extend morphological studies by one of the following ways:

(i) Certifiably pure lines of larvae, pupae and adults, resulting from cytological studies.
(ii) Adult and larval chromosomes.
(iii) Electrophoretic studies to characterise allozymes (=alloenzymes), providing that sibling differences in larvae persist to adults.

Because of their superficial morphological similarity, sibling species are often difficult to recognise. They are often discovered through various differences in habits, ecology or physiology. Among various attributes that distinguish siblings, Mayr (1969) lists biometric differences as being amongst the most important. Although qualitative structural differences may be absent between species, their distinctiveness may be substantiated by biometric studies. White (1977) upholds this view for the future of sibling

species studies in mosquitoes, by suggesting that the field should not be monopolised by geneticists and other experimentalists, but that computers should be used to process morphometric data.

The difficulty in using morphological data for the study of some species complexes has meant that considerable effort has been spent in obtaining genetic evidence to show the reproductive isolation of different populations. Although this information is often more difficult to collect, it is more useful in the application of the biological species concept. These sibling species are not different from other species in the biological sense, but for theoretical and applied reasons they have attracted much attention.

Genetical data collected from the field populations has led to a much greater understanding of the structure of species and mechanisms of speciation under natural conditions. For example, the orthodox theories of speciation (Mayr, 1970) do not concede that sympatric speciation occurs naturally. However, many sibling species are sympatric (more often than not) and although there is no direct proof, there is a growing body of evidence from their structure to show that sympatric speciation does in fact occur in the wild (Downes, 1973).

The techniques used in the study of sibling species complexes to establish their status within the biological concept fall into two categories.

(i) Techniques which investigate the genetic composition of the species to demonstrate that interbreeding does not occur in nature, e.g., cytogenetics and allozyme studies.

(ii) Techniques which demonstrate the presence of an isolating mechanism by direct observation or experiment, e.g., cross-mating and hybrid sterility tests.

1.5.1. Cytological Studies

Cytotaxonomic studies have played an important part in understanding many difficult taxonomic problems. It is now generally accepted that chromosome studies can distinguish sibling populations, provide evidence of reproductive isolation, and assist in tracing lines of phylogenetic descent (Rothfels *et al.* 1978). In these ways, cytotaxonomy has been of paramount significance in the taxonomic, ecological, and epidemiological analysis of groups such as the Simuliidae, Culicidae (and Drosophilidae) in all parts of the

world (W.H.O., 1977, p.9).

The chromosomes of many Diptera are of the polytene type, and exhibit a succession of transverse bands (discs) that reputedly reflects the genetic structure. These light and dark bands, together with 'puffs' and constrictions are often characteristic of a population. Polytene chromosomes have been studied mainly in larval tissue, usually the salivary glands or malpighian tubules, but good preparations have also been made from adult females of some species, by examining the ovarian nurse cells. All of these tissues are sites of high protein synthesis.

Often part of the chromosome becomes inverted so that the gene sequence of a section of the chromosome is reversed. Inversions are usually recognised by a reversal in the order of the banding patterns and also by local failure of pairing. The complex banding pattern allows an inversion to be recognised with great accuracy. The large number of bands and the variable length of an inversion makes it very unlikely that any one inversion will be repeated by chance alone. Hence, problems of evolutionary convergence (so frequently found in morphological structures) is very rare. In practise, the polytene sequence for one member of a group is taken as a standard and subsequent patterns (described as rearrangements of the standard) are necessary to produce the new pattern.

Most individuals with a new inversion can interbreed successfully with those having standard chromosomes. The progeny from such crosses are heterozygous (for the inversion) and will show the characteristic inversion loop on the pair of polytene chromosomes. The various proportions of individuals homozygous for the standard, heterozygous for the inversion and homozygous for the inversion, is characteristic for a population and may be easily recognised by studying the polytene chromosomes from a large sample of individuals.

However, polytene chromosomes do not always show differences between species. Carson (1967) found groups of known Drosophila species which were well differentiated morphologically, but showed no obvious difference in chromosomal banding.

Although such chromosomes have been found in all mosquito genera examined, preparations of suitable clarity for taxonomic use have been found in relatively few. They have proved most useful in the taxonomy of Anopheles. In this genus several species complexes have been studied, the two most significant of which are the maculipennis and gambiae complexes. White (1975) reviews the

cytotaxonomic studies of the Anopheles vectors of malaria and gives a detailed account of the techniques used.

In the A. gambiae complex, six species have been recognised. They were originally discovered from crossing characteristics, and confirmed by subsequent work on chromosome banding. The chromosomal characteristics are particularly important for the field identification of all six sibling species (W.H.O., 1977) and apart from the lengthy process of cross-mating, cytotaxonomy is the only reliable method of distinguishing the members of this group.

In the A. maculipennis complex, there is a similar situation with most of the component species having different banding patterns. Certain fixed inversions separate the species. A. labranchiae Falleroni and A. atroparvus Van Thiel have virtually identical chromosome banding patterns (homosequential). The complex has been studied in Europe and North America by Kitzmiller et al. (1967).

In the Simuliidae, much pioneering work on cytotaxonomic methods was carried out by Rothfels (1956 et seq) and his coworkers. Much of this work has been done on North American fauna, where 22 of the 150 species were detected through chromosomal studies. Many species still remain unstudied. One of the first 'species' studied by Rothfels was 'Prosimulium hirtipes', now recognised to be a complex of twelve species (Downes, 1973).

In Africa, the Simulium damnosum complex has been studied extensively (because of its importance as a vector of Onchocerciasis). One of the first new species named on the basis of chromosomal differences was published along with a cytotaxonomic identification key (Vajime & Dunbar, 1975). It is now known that this complex is composed of at least 25 taxa, which can only be distinguished with any certainty on chromosomal evidence (reviewed in W.H.O., 1977).

1.5.2. Enzyme Studies

Studies of enzyme differences controlled by different alleles at a single locus (alloenzymes) provide valuable information about the genetic variation in natural populations (Ayala & Powell, 1972). Recent work has been undertaken on biting flies to identify these polymorphic enzymes and to use their population frequencies to distinguish demes or species, evaluate taxonomic affinities, and to map genetic linkage of enzyme loci.

The genetic differentiation between populations is usually calculated by averaging the differences between the populations for all the loci studied. By this method, the same calculated value may be obtained from two types of variation in enzyme polymorphism. Firstly, when few loci are found showing complete, or nearly complete differences and secondly, when many loci are studied, which exhibit only moderate differences between populations. Usually only allozymes showing complete or nearly complete differences are used in taxonomy for the specific identification of populations, or individual specimens.

As in many other studies of population genetics, considerable use is made of the Hardy Weinberg Equilibrium to test for non-random mating in a sample, and consequently to demonstrate a lack of interbreeding in two or more demes.

The allozymes may be detected by the use of electrophoresis. This is a technique of separating molecules by means of an electric field, according to characteristics of their charge, size and shape. For any one specimen, staining of the electrophoretogram for a particular enzyme (over 100 are testable) reveals several bands of activity. Each band represents a different enzyme, termed an isoenzyme. Only those isoenzymes coded for at a single locus are called allozymes.

Numerous technical problems render the application of electrophoresis a specialised line of taxonomic research. It may be necessary to maintain cultures of known allozyme type to act as standard markers for each run. This technique is still in its preliminary stages for taxonomic purposes in the study of sibling species. In the Simuliidae, this technique has been used on the Simulium damnosum complex in W. Africa. So far, 18 enzymes have been tested, but have failed to produce unequivocal identification of the species within the complex (W.H.O., 1977). However, one enzyme has distinguished species of the complex in South Ghana.

In the mosquitoes more extensive studies have been made on the Anopheles maculipennis, A. gambiae, Aedes mariae, Aedes scutellaris and Culex pipiens complexes. Over 100 papers have been published on the subject (White, 1979). Of the complexes listed above, the Aedes scutellaris complex has benefitted most from the use of electrophoretic studies to separate all of its seven species.

Most of the studies in mosquitoes have differentiated local

populations, but to date there are comparatively few studies which have claimed reliable diagnostic allozymes for the separation of sibling species (White, 1979).

1.5.3. Cross-Mating Tests

The biological species definition of Mayr (1970) hinges on the reproductive isolation of two populations. Therefore direct observation of individuals of each is a good test of the status, specific or otherwise, of the populations. Hybrid sterility or inviability in one or both sexes proves that the parents belong to genetically incompatible species.

Hybrid sterility constitutes a post-mating isolating mechanism (Mayr, 1970) and is therefore of considerable importance in maintaining the integrity of the sympatric species. The use of cross-breeding studies for sympatric species is therefore of use to the systematist in confirming the degree of isolation of two populations. Results of this test are not so easy to interpret when studying allopatric species. Geographic separation constitutes a pre-mating barrier to hybridisation and so under natural conditions, the opportunity for mating would not occur, thus eliminating the need for a post-mating barrier such a hybrid sterility. When cross-mating tests are carried out in the laboratory, on allopatric populations, the production of viable hybrids is not an absolute criterion for distinctness.

Sibling species may be identified by crossing unknown individuals with those of known identity (usually laboratory colonies) (White, 1979). The use of cross-mating studies has been of particular use in the study of anopheline sibling species. In the Anopheles maculipennis complex, cross-mating tests have been used to establish the validity of most of the recognised species in both Europe and North America, and to explore the divergence of the species from these two regions (Kitzmiller et al., 1967).

In the A. gambiae complex, the six component species were first recognised by crossing experiments and confirmed by subsequent chromosomal studies. From the 30 possible crosses between sibling species, only two were considered fertile, but in these cases , the reciprocal crosses were sterile (Davidson et al., 1967; Davidson & White, 1972; W.H.O., 1977). Laven (1967) has shown that some of

the incompatibility in sympatric populations of Culex pipiens has its basis not in chromosomal genes but in cytoplasmic factors, which is termed cytoplasmic incompatibility. To date, the colonisation of Simuliidae has not been at all successful and therefore the techniques of cross-breeding tests, unlike many others, has not been applied to this group.

1.6. SPECIES COMPLEXES IN THE CERATOPOGONIDAE

The taxonomic problems of the Culicoides pulicaris complex are not unique in the genus. On the contrary, it is typical of a number of difficult groups, e.g., salinarius, nubeculosus, variipennis and obsoletus groups.

Only the C. variipennis complex in North America has been studied in any detail (morphological and ecological), but no genetic data have been obtained, unlike the species complexes of mosquitoes and blackflies. Most complexes in the Ceratopogonidae have been investigated only at the morphological level, for example, in the Leptoconops kerteszi complex of pestilent biting midges in North America (Clastrier & Wirth, 1978), eleven species have been recognised, where only one was known before. The description of the ten new species was based solely on morphological characters.

Compared with the sophisticated level to which the study of species complexes has risen for Simuliidae, Culicidae and Drosophilidae, that of a group such as the C. pulicaris complex can be seen in perspective as one whose study is in its infancy. At present, complexes in the Ceratopogonidae require that morphology be fully exploited by the use of biometric methods. The considerable variation within the C. pulicaris complex has led to the proposal of several nominal taxa. However, in common with other species complexes, the boundaries of the taxa overlap to the extent where it is difficult, if not impossible, to determine where intra-specific variation ends and inter-specific variation begins. The difficulties have been summarised by Campbell & Pelham-Clinton (1960): "most of the species are very variable, and since they are also closely related, individuals frequently occur which are difficult to place without much experience. It is possible too that natural hybrids occur between some species..... female structural characters are little better (than males) for identifying individuals though very importantfor defining populations". This point of view was also

shared by Wirth & Blanton (1969) in their review of the North American species, in which they suggested that a series of specimens is to be preferred for population identification purposes and the most representative specimens should be selected and studied.

The morphological problems of adults in the pulicaris complex are paralleled in the larvae (Kettle & Lawson, 1952). Glukhova (1977) has found that larvae exhibit group differences which, as a rule, are quite clear-cut, and that within groups of species great similarity occurs. It is not always possible to differentiate individual species within a group, based on larvae alone.

Section 2. OBJECTIVES OF THIS STUDY

2.1. OBJECTIVES

In the context of general considerations outlined in the preceding section, the objectives of the present work may be summarised as follows:

- To test how far some techniques of multivariate morphometrics can discriminate between sibling species of *Culicoides*.

- To develop a system for the specific identification of adult female specimens belonging to the *Culicoides pulicaris* complex.

- To investigate biometric variation of some anatomical characters in *Culicoides* and its taxonomic implications.

- To devise a system for specific classification of morphologically overlapping species (sibling species).

Although some of the objectives are rather ambitious, the thesis also attempts to make a critical study of the extent to which these objectives can be realised, through the use of available methods. An attempt is also made to evaluate the reasons for success or failure of the methods.

2.2. RATIONALE AND LIMITATION OF THE APPROACH

It has become increasingly apparent in recent work on *Culicoides* taxonomy, that the traditional approach of using wing and mesonotal disc patterns, to differentiate the species, is not sufficient (Atchley, 1967). Furthermore, lack of simple qualitative differences to distinguish between species belonging to the more difficult complexes, requires a method that considers a number of characters simultaneously. Such an analysis is achieved by using methods of multivariate statistics. Before dealing with this approach, some of the additional or alternative methods should be considered.

The study of chromosomes is one of the methods more commonly applied to investigate species complexes. They are usually taken from the final stage larvae or, in some cases, from adult ovarian tissue. Although this technique has been most useful in the biting flies, there are two drawbacks to its use in the study of Culicoides:

1. It has been shown by Atchley & Jackson (1968) that polytene chromosomes are "virtually impossible" to prepare well enough for cytotaxonomic studies of Culicoides.
2. Cytological techniques make it desirable to establish colonies of the species concerned, which is difficult with these midges owing to their host preference, mating conditions and other factors.

The second problem also makes laboratory cross-mating tests of various taxa difficult. Colonies have been established in only a few species (C. variipennis (Coquillett), nubeculosus (Meigen) and furens (Poey), and require considerable laboratory and labour facilities. Jones (1960) states that the full time work of a trained technician was required to maintain their colony of C. variipennis. Unfortunately such facilities were not available for the present study. Attempts at colonisation of the ground mating C. melleus (Coquillett) and the swarming C. hollensis Melander & Brues, by Koch (personal communication) were hindered by the complex conditions required for mating. Less than 5% of all attempted matings were successful in laboratory cages.

Consequently, at least for the present, studies involving species complexes of Culicoides will have to be based on morphology. Thus, in the present study, the methods of multivariate morphometrics have been adopted to identify and, if possible, to emphasise any small, compound and relative characters in the Culicoides pulicaris complex. These multivariate methods are particularly suited to the problem of morphologically overlapping species, because many characters may be considered simultaneously, and the contribution of each character to a classification evaluated.

Once a framework describing the variation in the complex has been established, working hypotheses, concerning the specific status of the recognised morphological entities, may be advanced. Against these hypotheses, further observation and experimentation must be made to determine their validity. Whatever the practical problems

associated with non-morphological studies, the point remains that, even when the morphologically defined taxa have been recognised, their significance in terms of the biological species concept is unresolved.

Section 3. MORPHOLOGY

3.1. ASPECTS OF THE TERM 'CHARACTER'

Prior to a study of character variation, it is necessary to establish which definition of a character is to be used, as several have been advanced.

Mayr (1969) gives the following definition: "any attribute of a member of a taxon by which it differs or may differ from a member of a different taxon". As Mayr concedes, this definition makes a number of important assumptions. Principal amongst these is that characters only distinguish organisms and therefore leads to a classification that is a mechanism for identification rather than a predictive, and an information retrieval system. Mayr(1969) does recognise this problem in an informative discussion of biological classification as a scientific theory, although it does not prompt him to expand his definition.

Another troublesome assumption of this restrictive use of the term 'character' is that it emphasises an important dilemma of conventional taxonomic procedure, summarised by Sneath & Sokal (1973): "characters are restricted to differences between members of taxa, but the taxa cannot be recognised without the characters themselves being first known".

Cain & Harrison (1958) consider a character in a much broader context, defining it as "anything that can be considered as a variable independent [logically] of any other thing studied at the same time". Such a definition is prevalent in the works of numerical taxonomists and has obvious differences to that of Mayr and his adherents. One principal difference is that views like those of Mayr are a product of practical taxonomy involved in the day-to-day problems of producing a workable classification. Numerical taxonomists on the other hand, especially in the earlier days, seek a more theoretically based system to encompass many of the advances in biology. The objective of this preamble is not to further this discussion, but to recognise it and indicate the viewpoint taken in the present work. For the purposes of this study, the definition of Cain & Harrison has been used. However, as one of the overall objectives is to produce a system for identifying

unknown specimens, many of the characters used will be discriminatory.

According to the definition accepted here, the presence or absence of cibarial teeth is a 'character' with two states - presence or absence. For quantitative characters, e.g. wing length, the states are not discrete, but vary continuously and may represent as many states as the observational technique will allow. It should be noted that Blackwelder (1967) and others would refer to the 'states' defined above as 'characters'. They have no category equivalent to a character as it is used here. Jardine & Sibson (1971) distinguish between an attribute (= character used here) and its characters.

Sokal & Sneath (1963) introduced the term 'unit character', thus initiating a controversy between numerical and conventional taxonomists. After accepting a character in the terms outlined above, they went on to define a unit character as a fundamental concept in numerical taxonomy. Their definition - "a taxonomic character of two or more states which, within the study at hand, cannot be logically subdivided" was aimed at producing information content in terms of Information Theory's 'bits'. If ratios are considered as a measure of shape in their own right, rather than in terms of two components, then a number of characters used in this study conform to this definition. There has been considerable criticism of the unit character concept and by many who misinterpret it. For example, Griffiths (1972) states that it inherently implies logical atomism. By using the phrase 'in the study at hand' there is no suggestion that the nature of the logical division is absolute, i.e. the unit character is not an atomic fact _per se_. The definition acknowledges that any classification is a function of the methods employed in its construction e.g. biochemical, physiological, immunological, etc., and that the unit character of one study is not congruent with that of another. In this sense, Russell's (1918) complete analysis of complex facts has not been satisfied and the character does not have the status of an atomic fact. Therefore the definition used by Sokal & Sneath does not aim at a 'universal' (in Russell's sense), but a logical and practical concept.

To summarise, a character does not have to be discriminatory to be useful and its variants are termed character states.

3.2. THE ANATOMY OF CULICOIDES

Culicoides are small flies with a humped thorax (Figs 1 and 2). Their compact nature enables them to run through the host's hair before biting, in contrast to the 'land and bite' tactics of the long legged mosquitoes and sandflies.

The following section is a brief outline of adult structure in Culicoides, together with a discussion of specific characters used in this study. A number of new characters are discussed in Section 8 (p.142). More detailed descriptions are given by Linley (1976), Atchley (1967, 1970), Arnaud (1956), Tokunaga (1937). Jobling (1928) studied the structure of the head and mouthparts of C. pulicaris and Gad (1951) reviewed the head structure of a range of Ceratopogonidae including C. impunctatus. The anatomy of the male genitalia of Culicoides is discussed by Pomerantzev (1932) and the anatomy and histology of the alimentary tract by Megahead (1956). The electron microscope has been used to examine the fine structure of antennae and palps in North American Culicoides by Rowley & Cornford (1972) and Chu et al. (1975).

3.2.1. The Head

The head of adult Culicoides is hypognathous with prominent antennae and an elongate proboscis (Fig.3). The lateral region is composed of large, reniform compound eyes which, in the females of the pulicaris complex, may either touch dorsally, or be separated by a narrow projection of the frons. The broad vertex occupies most of the dorsal region of the head and may be arbitrarily delimited from the frons by the interocular suture. This suture indicates the presence of an internal phragma, presumably serving as a cross-strut to strengthen the front of the head. Another small suture is often present in the centre of the frons when the eyes do not touch. Contiguity of the eyes was coded by measuring the separation or contiguity of the eyes in units equal to one facet width (Fig.7). This measurement was then coded in 11 steps of 0.5 units from eyes separated by 2.0 facet widths, to touching for 3.0 facet widths. There is rarely any interommatidial hair and the ocelli are poorly developed.

The facial region is composed primarily of the fused frons

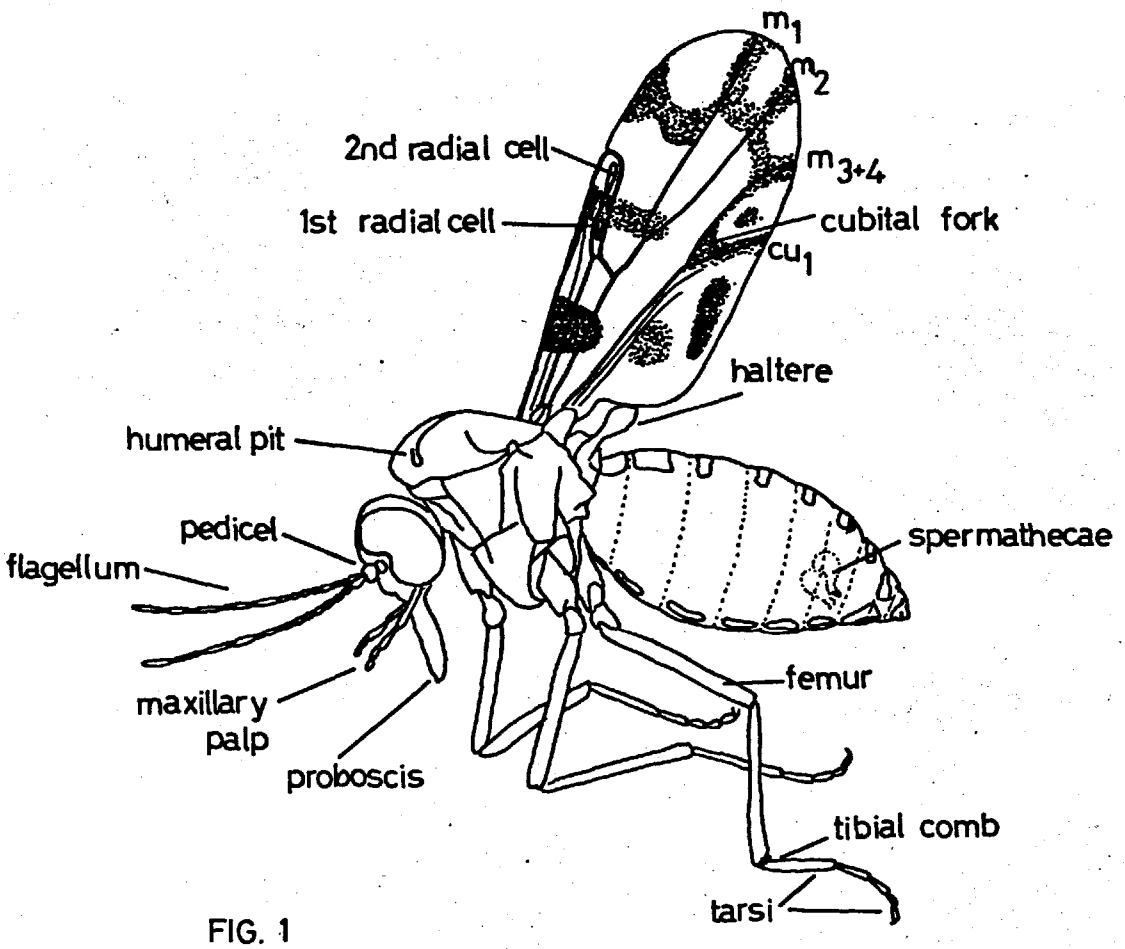
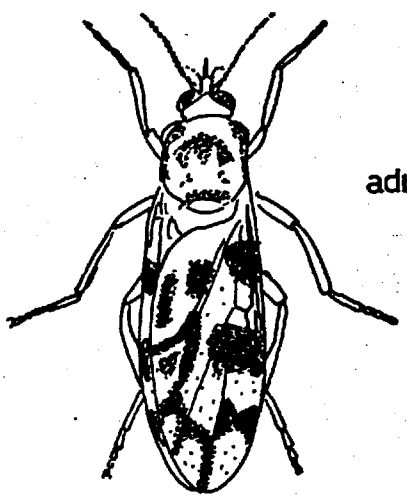


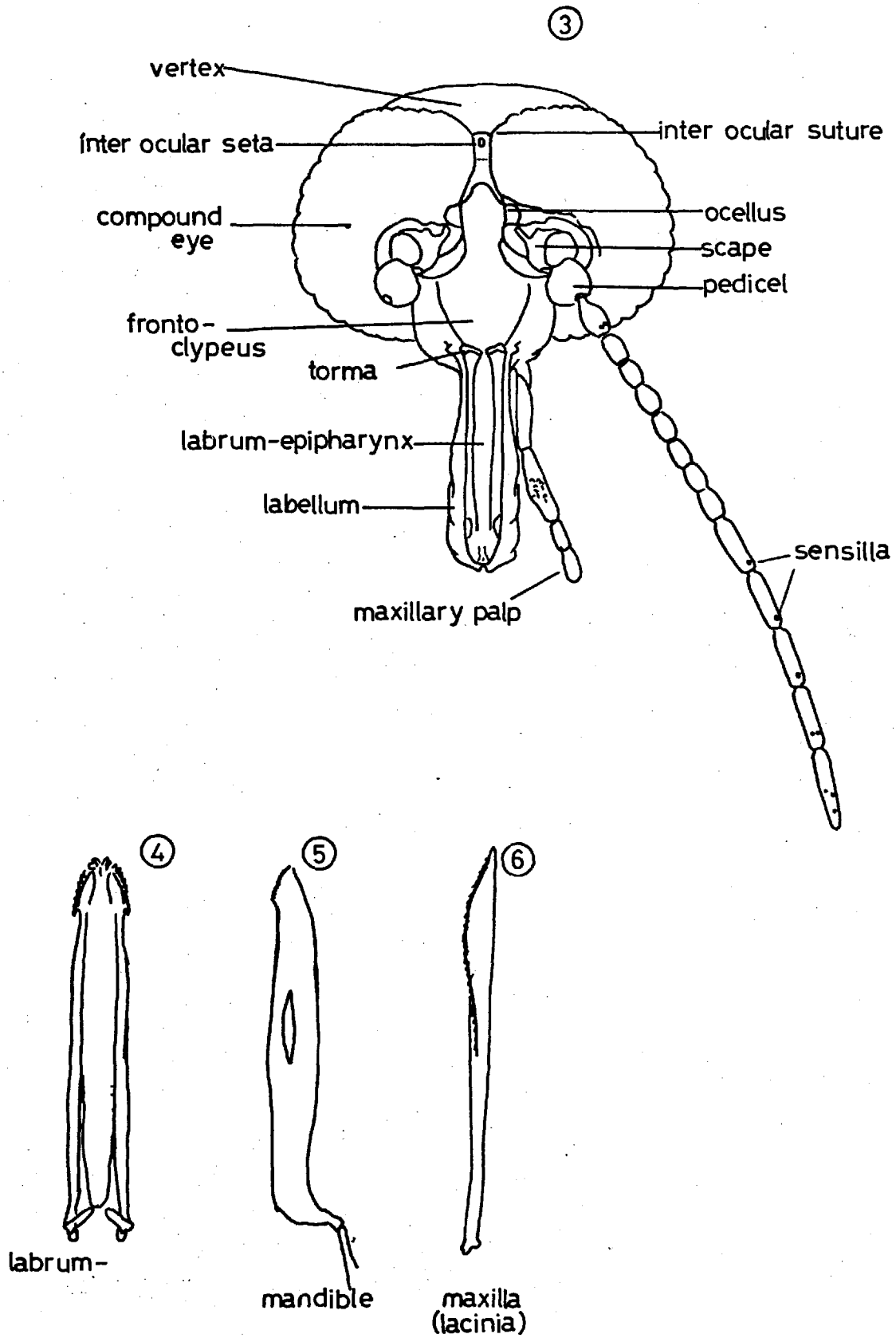
FIG. 1



adult female of C.pulicaris at rest

FIG. 2

FIGS 1-2 GENERAL MORPHOLOGY OF CULICOIDES



FIGS 3-6
HEAD AND MOUTHPARTS OF FEMALE CULICOIDES

and clypeus, the frontoclypeal suture having been lost in the consolidation of the head capsule. At the base of the frontoclypeus are two small sclerites, referred to as tormae, joining basally to the labrum. The tormae are intimately concerned with the articulation of the mouthparts and head capsule. Their shape is sexually dimorphic, reflecting a difference in feeding habits of the two sexes - only the females use their mouthparts to pierce the skin of larger animals and suck blood. The head length is the distance measured from the tormae to the interocular seta (Fig.8).

The mouthparts form an anterio-ventrally projecting structure called the proboscis. This is well developed in the female and somewhat atrophied in the male. Jobling (1976) suggests the term 'syntrophium' as a general term to include the mouthparts of all biting Diptera. He further proposes the establishment of two classes of mouthparts - achilophorous and chilophorous - based on their structure. Culicoides do not have the labium as the principal piercing component and therefore are of the achilophorous type.

The biting apparatus consists of the labrum, mandibles, maxillae, labium and hypopharynx. In much of the literature on Culicoides morphology, the term labrum-epipharynx is used. Since the epipharynx is an outgrowth of the inner face of the labrum, there does not appear to be any real justification for using this term. It is therefore simply referred to as the labrum.

The labrum of the C. pulicaris group is elongate and blunt, with six terminal teeth and a number of laterals (Fig. 4). There are small clear areas apically, presumably sensory in function. The distance from the tormae to the tip of the labrum is termed proboscis length (Fig. 8). This length, relative to that of the head, varies within the genus Culicoides and may be correlated with the degree of autogeny (Downes, 1970).

The mandibles are pointed and bladelike (Fig. 5), with a series of outward projecting lateral teeth. They are articulated basally with the mandibular condyles, and are crossed with an interlocking area midway along their length, giving a scissor-like action.

The maxilla consists of three parts: cardo, stipes and maxillary palp. The cardo and stipes are reduced to small sclerites, lying in the membranous areas below the foramen magnum. The

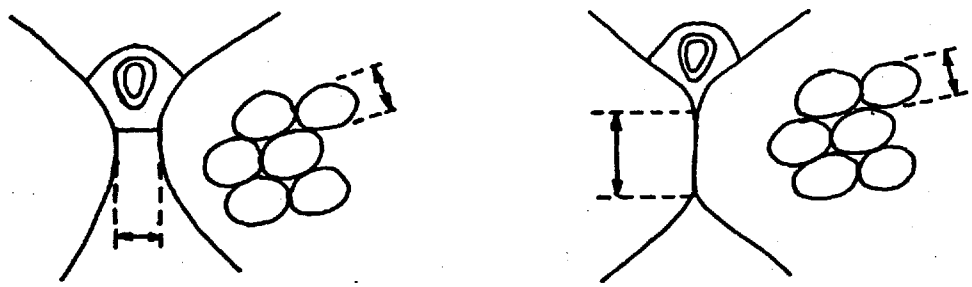


FIG. 7 CONTIGUITY OF EYES

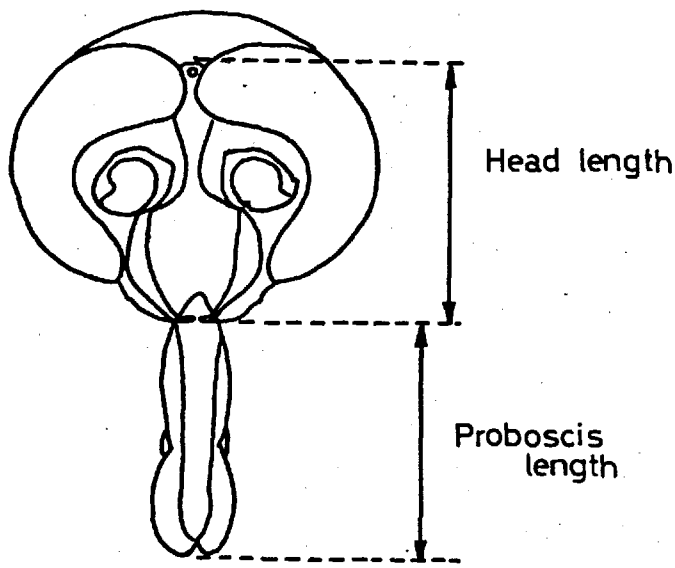


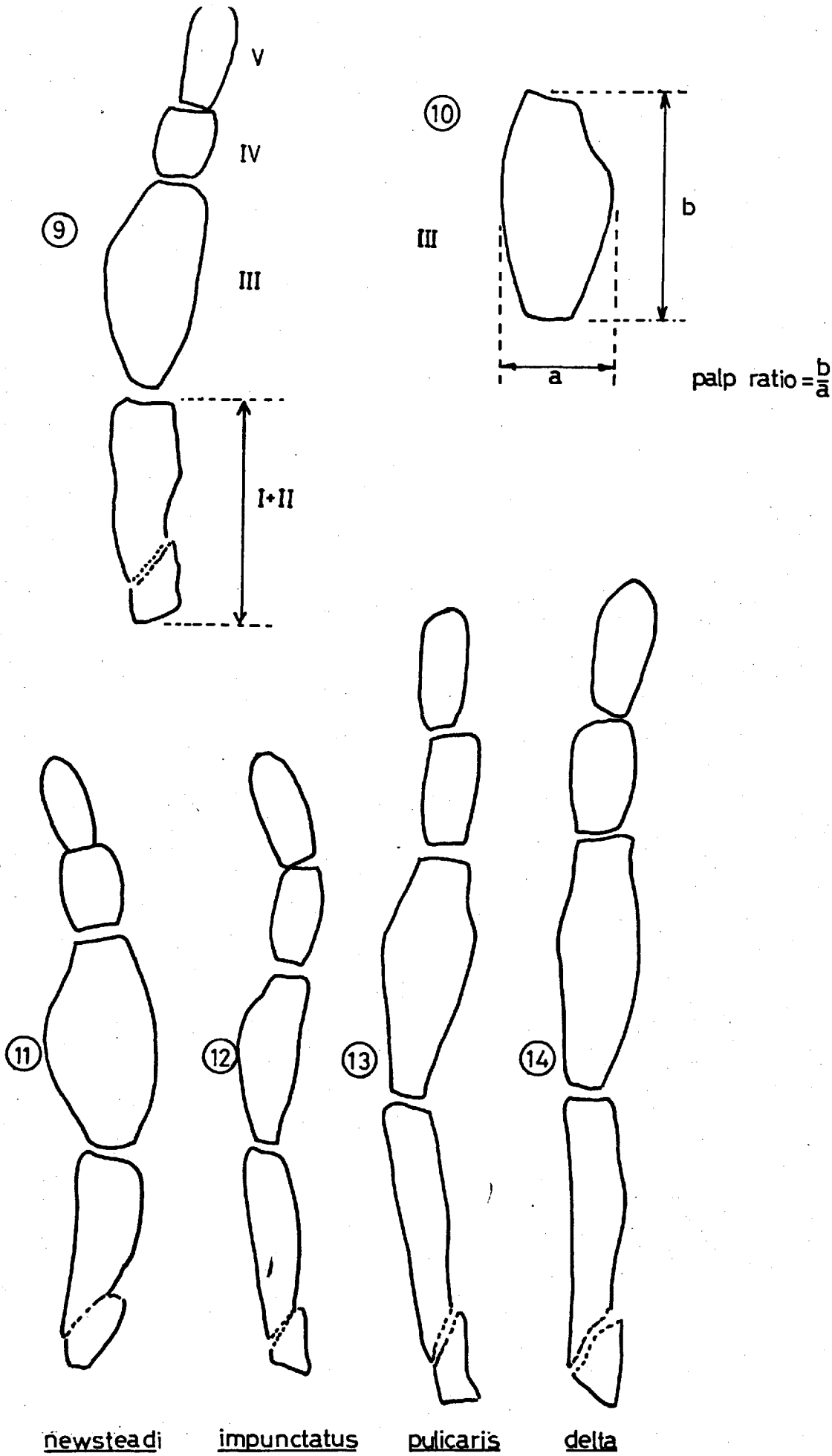
FIG.8 MEASUREMENT OF HEAD

'maxillary stylet' commonly referred to in the literature was thought to be the galea by Jobling (1928) and many subsequent authors. However, this has been proved erroneous by Matsuda (1965) following Gad (1951) and Imms (1944), who suggest that the galea has been lost and that it is the lacinia which is present. The lacinia is slender with a series of long discrete teeth on the apicolateral margin (Fig.6). The number of teeth was recorded for both the mandible and maxilla.

The maxillary palps are five-segmented, with a swollen third segment, bearing specialised sensory hairs. The first segment is small and weakly sclerotised, often varying between individuals. As the boundary between the first and second segment is so often diffuse, both segments are measured together (Fig.9). In the pulicaris group, the sensory hairs of the third segment are distributed throughout a number of small pits, a feature which distinguishes them from many other species groups. Using Tokunaga's (1937) classification of palpal sensory organs, the pulicaris arrangement is typical of his 'scattered type'. The sensory hairs are bulb-shaped sensilla, similar to those of the Simuliidae (Mercer & McIver, 1973) and the Culicidae (McIver, 1971). Electrophysiological and behavioural studies in the mosquito Aedes aegypti have shown that such bulb-shaped sensilla are sensitive to carbon dioxide (Kellog, 1970). Linley (1976) illustrates some scanning electron microscope photographs of these receptors in C. hollensis Mellander & Brues.

The use of the palpal ratio is an attempt to describe the shape of the third segment, which varies considerably, yet systematically, within the pulicaris group. It is derived by dividing the length of the segment by its width at the broadest point (Fig.10). In species such as C. newsteadi (Fig.11), with a short swollen segment, the ratio is low and contrasts sharply with the slender palp of C. delta (Fig.14). C. pulicaris (Fig.13) and E. impunctatus (Fig.12) are intermediate.

The significance of the variation in shape of the third palp segment and the palpal pit is difficult to elucidate. It is mostlikely related to host specificity, but the exact nature of this relationship is not clear. In some species of Culicoides, such as the marine C. circumscriptus Kieffer, which has not been found biting any bird or mammal, the third palp segment is greatly swollen.

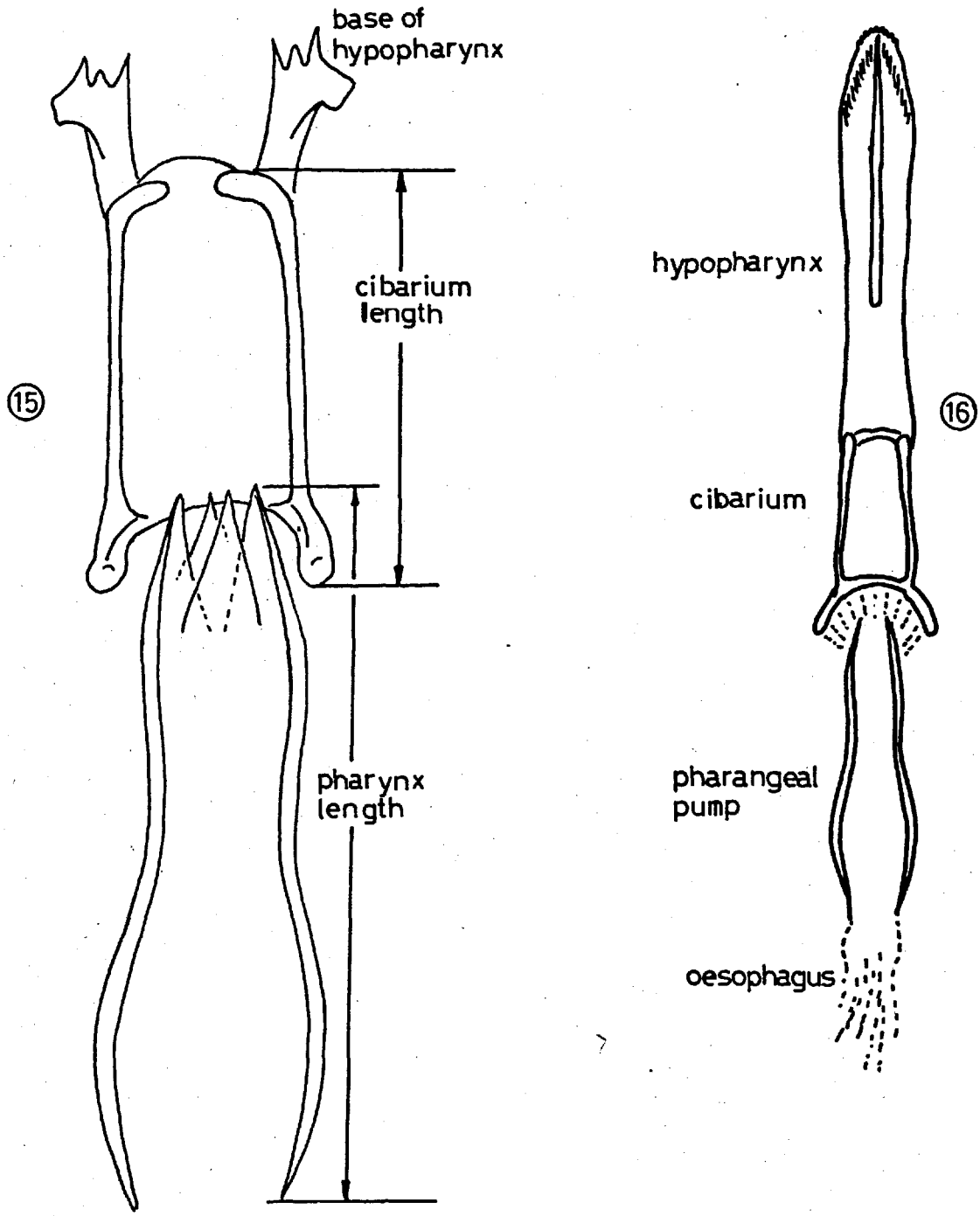


FIGS 9–14
MEASUREMENT AND VARIATION IN PALPS OF C. PULICARIS COMPLEX

However, Hair & Turner (1968) suggest that host preference is not clearly demarcated in Culicoides, many species attacking both birds and mammals. Kettle (1962) suggests that each species has a range of hosts on which it will feed but generally prefers one in perticular. In considering the apparently conflicting data, a broader hypothesis may be postulated in that the structure of the third palp segment may not be related purely to host location but to host location in relation to the habitat. For example, in Britain there are a number of species with a swollen palp segment: C. salinarius Kieffer, circumscriptus Kieffer, duddingstoni Kettle & Lawson, maritimus Kieffer and newsteadi Kieffer, representing three subgenera. All these species inhabit coastal marshes, which are typically flat and exposed. This contrasts with the woodlend, meadowland and some moorland frequented by other species. Within the pulicaris group, the only species inhabiting brackish environments is newsteadi, which, es noted above, possesses the most swollen palp segments in the complex. The above hypothesis is rather speculative, but would benefit from further research, using a larger sample of species. The mexillary palps and lacinia are not well developed in males - undoubtedly related to their non-biting habit.

The labium of Culicoides is represented by a postmentum, prementum and a two-segmented palp. As in other Nemetocera, the ligula has been lost and the palps are modified to form the two-segmented labellum dominating the labium (Fig.3). In females, the hypopharynx is bladelike, with apical teeth, and together with the labrum, is joined basally to the anterior opening of the digestive tract (Fig.16). The first section of this tract is a well sclerotised tubular structure, variously termed 'cibarium' by many authors,'pharynx' by Jobling (1928) and 'cibarial pump' by Gad (1951). It lies almost vertically in the head with large muscles attached posteriorly, and was measured as shown in Fig.15.

Posterior to the cibarium is anothercomponent of the buccopharyngeal region, which also has a confused nomenclature. Jobling (1928) refers to it as the'oesophageal pump', whilst Ged (1951) calls it the 'pharyngeal pump'. Subsequent authors have followed either one or the other of these interpretations. The cause of the controversy lies in establishing homologies for the muscles attached to these parts. Snodgrass (1944) has shown this



FIGS 15-16
ANTERIOR SECTION OF DIGESTIVE TRACT

area to be a modification of the pharynx and this interpretation is followed in this study. Whatever the nomenclatural differences, most authors are agreed that the cibarium and pharynx function as pumps during feeding. The pharynx was measured as shown in Fig.15.

The antennae are composed of 15 units in both sexes which, although are not strictly segments, are often referred to as such. In North America, the term'flagellomere' is frequently employed, but the term 'segment' is used throughout this study. The antennae are of considerable taxonomic importance. The first segment, called the scape, is flattened and ring shaped, whereas the second segment, the pedicel (pedicle of Jobling) is considerably enlarged to house the Johnston's organ. This is well developed in the male and together with the plumose antennae, serve as a receptor to detect the sound of the female wingbeat. This organ is present to some extent in most Diptera, indicating that the pedicel of the more generalised Dipteran form, demonstrated by the Nematocera, is homologous with the same segment of the more specialised Diptera. Further discussion on the function of this structure is given in Section 7 (p.116).

The remaining 13 segments of the antennae comprise the flagellum. The first segment of the flagellum is nearly oval with an elongated neck which fits into the funnel-shaped depression of the pedicel. The proximal segments are slightly elongated, usually decreasing in size to the tenth segment. Thereafter, there is an abrupt change in size, so that segments xi - xv are elongated and cylindrical.

There are five main types of sensilla on the antennae, outlined by Cornet (1974). The most taxonomically significant of these are termed sensilla campaniforma, often abbreviated to 'sensory pits' or 'sensilla' in much of the Culicoides literature. These structures are circular pits with a series of setula around their perimeter and a thin-walled peg (sensillum coeloconica) in the centre. Their fine structure has been described by Chu et al. (1975). These sensilla are found in a number of Ceratopogonid genera, as well as other insects, where the term Picket-fence receptors is currently used (Callahan, 1975). The number and distribution of the sensilla on differing antennal segments is of considerable importance in Culicoides taxonomy. However, considerable

variation has been found in their distribution (Kremer & Delécolle,1974), casting doubt on their reliability for separating closely related species.

Jamnback (1965) has suggested that the number of sensilla on the antennae of the female is correlated with host preference. Of the few species for which he provides data, those with sensilla on a total of eight to thirteen segments were primarily ornithophilic, whilst mammilophilic species had sensilla present on only four to six segments. It is noteworthy that no real evidence, other than circumstanstial, has been put forward concerning the function of these receptors, although their function has always been presumed olfactory. Chu et al. (1975) have shown that the central peg of this compound sensillum has a lumen with a series of small pores 100Å in diameter, communicating with the exterior, and suggests that they may function as hygroreceptors. Slifer (1970) however, has suggested that thin walled pegs may be olfactory in function.

The antennal ratio (in the female only) is the summed length of segments xi - xv, divided by the summed length of segments iii - x.

3.2.2. The Thorax

The thorax is convex dorsally, extending slightly over the head. The presence of two pits on the humeral corners has been given much significance as a generic character, but their function is not known, or rarely discussed. Linley (1976) suggests they may be sensory in function.

The pits are located in a slight depression, immediately behind the anterior margin of the dorsum and appear as shining patches. Their depth and shape varies between species, in the pulicaris complex they are elongate with the main axis running dorso-ventrally. The pits are asymmetrical in cross-section, with the most abruptly sloping edge anteriorly. Overall, they are inclined in a posterio-lateral direction. Within the pit are small pores, usually most dense at the dorsal margin. Two hypotheses to ascertain the function of the pits are suggested here. Firstly, they are secretory and produce a pheromone. Due to the shape of the pit, air moving over the thorax during flight would cause turbulence and draw the secretion into the airflow

and away from the insect. The second hypothesis suggests that the pits are scars, resulting from the detachment of the tracheae in the pupal horns during eclosion. The respiratory horns in the pupa arise from the prothorax, and correspond closely with the humeri of the adult thorax. It is difficult to determine whether the small pores function as spiracles in the adult (Lane, 1979).

Of these two hypotheses, the latter seems more feasible, as it requires less supposition (i.e., Culicoides communicate by pheromones in flight and also by the nature of airflow around the thorax). In addition, the latter would also explain the presence of humeral pits on other Nematocera with aquatic pupae.

The thorax is usually patterned, and often used in the taxonomy of Culicoides. These patterns were not used in the present study for the following reasons:

1. In their study of the British pulicaris group, Downes & Kettle (1952) noted that the thoracic markings were very variable, and when used in conjunction with wing pattern, conflicting identifications resulted. This mixture problem occurred in material from a number of localities.

2. Pattern is only reliable when seen in freshly collected material. Such material was not available for this study because specimens were studied from a wide geographical range.

3. Pattern is only readily visible in dry mounted material, but the present study demanded slide mounting of all specimens for accurate measurement of numerous other characters.

4. As wing patterns were also used, the problems of attempting to encode both wing and thoracic pattern objectively would present considerable difficulties in accuracy.

The legs are slender with five tarsal segments, the last bearing a small pair of equal claws. Grooming organs, consisting of combs and spinose hairs, are found on the inner apices of the fore tibiae, and as a comb on the hind tibiae (Linley & Cheng, 1974). The length and number of spines on the hind tibial comb is often used in Culicoides taxonomy, but unfortunately has not been found useful in the pulicaris complex.

When living, the wings are folded over the back whilst at rest, and usually cover the whole abdomen (Fig. 2). They are covered with dense microtrichia and a number of macrotrichia, especially towards the distal half of the wing. Shape is usually

a sexually dimorphic feature, being relatively more slender in the male. Whether this has any functional reason, perhaps associated with swarming in the male, is not clear. The majority of the well sclerotised wing veins are concentrated in the anterior section of the wing, where two radial cells are also present. Wirth (1952, p.103) gives a very useful table comparing the wing vein nomenclature used by different workers on the Ceratopogonidae. The system used at present is that of Comstock-Needham, modified by Tillyard.

The wings of Culicoides are frequently patterned, details of the dark areas having been used extensively in the taxonomy of the genus. The underlying mechanism of the pattern has not hitherto been investigated. Atchley (1970) states that the patterns of spots are "formed by varying densities of microtrichia". By 'densities', he presumably means spacing of the microtrichia. This hypothesis was tested by using the scanning electron microscope, after first drawing the wing pattern using the light microscope, and noting important landmarks such as venation. Fig.17 shows a low power photograph of the wing, revealing the lack of any overt pattern when the surface detail is considered. If the patterns were entirely due to surface features, such as density of microtrichia, a pattern would be evident in the photograph. Taking the analysis further, the distances between the microtrichia were then measured at a magnification of 2000 times. The microtrichia are approximately 6 μm long and are quite distinct at this resolution. As the microtrichia are present in rows, the distance to the nearest neighbour, both within a row and between rows, was measured for a 'pale area' and a 'dark area'. The positions of the costal spots were used since they could be easily located using structural landmarks of wing venation. Measurement of both inter- and intra-row distances would detect a pattern based on either overall microtrichia density, or row density. The results of 't' tests comparing spacing for pale and dark areas are shown in Table 1.

TABLE 1

	Mean distances in μm	Value of 't'	Degrees of freedom	Probability
intra-row distances	light=22.67 dark =21.36	0.159	16	P=> 0.1<0.2
inter-row distances	light=20.10 dark =20.44	0.375	23	P=>0.2<0.3

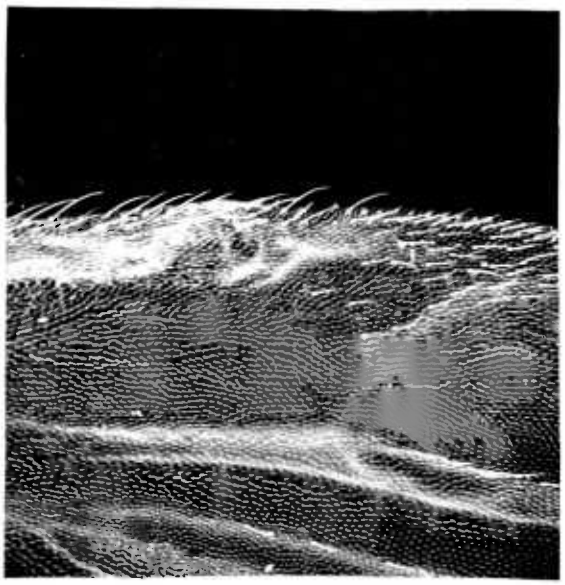


FIG. 17. Anterior margin of Culicoides wing, as seen by the scanning electron microscope. The boundary between light and dark pattern runs down the centre of the photograph.



FIG. 18. Wing of Culicoides punctatus, viewed with dark-field illumination.

These results clearly show that the null hypothesis cannot be rejected, and that there is no significant difference between the spacing of microtrichia in pale and dark areas of the wing. If the wing is viewed under dark field illumination (Fig.18), it is clear that the wing pattern is due to pigmentation of the microtrichia. The coding of wing patterns is discussed in Section 8.

3.2.3. The Abdomen

The abdomen of the female is broad, tapering posteriorly, and commonly expands and contracts with engorgement of blood.

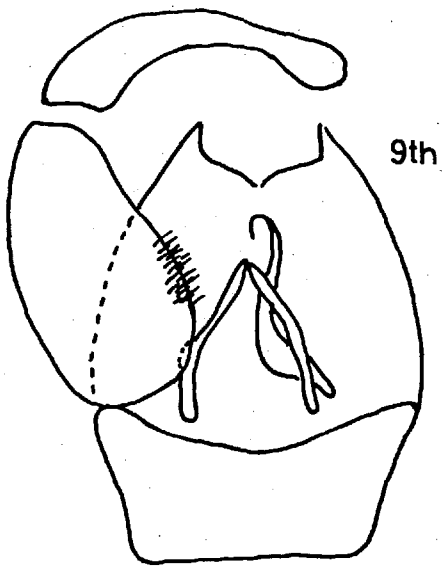
There are two well developed spermathecae in the pulicaris group, with the occasional presence of a diminutive third. A sclerotised ring surrounds the common duct leading from the fusion of the two small spermathecal ducts, to the exterior. Unfortunately the internal reproductive organs of the female were of no taxonomic use in the pulicaris complex for this study.

As in most other Diptera, the male genitalia are frequently used in the Culicoides taxonomy. The genitalia consist of a ninth sternite and tergite, paired appendages (basi- and dististyle), a single aedeagus and paired parameres. In the pulicaris group, as in the nubeculosus and salinarius groups, the genitalia are of little use in discrimination of different taxa. In fact, they provide a reliable means of identification for only two species in the pulicaris group. C. grisescens and C. fagineus may be separated from other species by the presence of a convex ninth tergite (Fig. 19).

3.2.4. Size

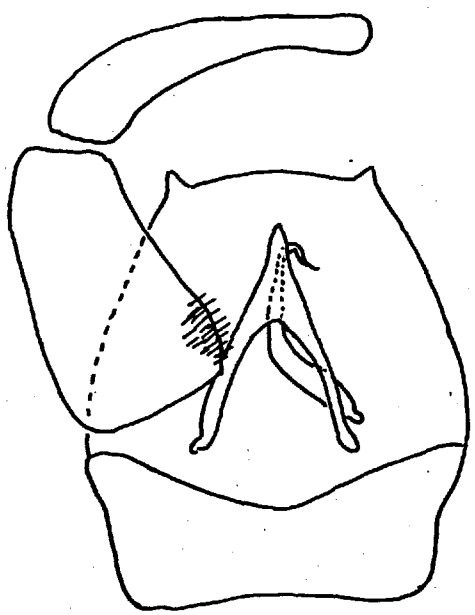
Absolute size has been used as a discriminatory character in a number of keys to the pulicaris and other species groups of Culicoides (Campbell & Pelham-Clinton, 1960; Wirth & Blanton, 1969). In several species groups, the size range of the constituent species are significantly different, but this is not always the case, especially if the specimens compared are from widely separated localities.

A number of studies into factors influencing size in the Ceratopogonidae have been made, principally by Linley, on salt marsh species. Linley(1969) showed adult size was dependent on



9th tergite (concave)

C. pulicaris



9th tergite (convex)

C. grisescens

FIG.19 MALE GENITALIA

larval diet in C. furens (Poey) and on larval development in Leptoconops bequaerti (Kieffer) (Linley, 1968a). Linley(1968b) has also shown that L. bequaerti is polymorphic for wing length, and suggests that this is under genetic control, associated with the degree of autogeny (larger insects were anautogenous). The seasonal variation in size of C. melleus (Coquillett) is discussed by Linley & Hinds (1976), Linley et al. (1970), and is related to ambient temperature. Hensleigh & Atchley (1977) made similar studies, augmented by work on an established laboratory colony of C. variipennis Coquillett.

All these studies revealed that absolute size is negatively correlated with temperature, a common phenomenon in insects. In addition to the factors mentioned above, adult size in other Diptera appears to be under the control of larval competition in Psychoda (Lloyd & Golightly, 1939) and also in Calliphoridae (Lane,1975). The possible factors influencing seasonal variation in size of adults in the pulicaris complex are discussed in Section 7.

Section 4. TAXONOMIC REVIEW OF THE CULICOIDES PULICARIS COMPLEX

4.1. DEFINITION OF HIGHER TAXA

4.1.1. Ceratopogonidae

For many years the biting midges were placed in the genus Ceratopogon, in the family Chironomidae. The ceratopogonids were, at most, regarded as a subfamily of the Chironomidae until 1901, when Grassi elevated them to family rank, followed by Malloch in 1915. They exhibit a considerable range of structure and biology and yet form a compact group, showing many differences with the non-biting Chironomidae. Although earlier workers had suggested the division from the Chironomidae, it was not until Edwards (1926) that a coherent definition was given for the Ceratopogonidae:

Head: rounded behind. Mouthparts complete; mandibles well developed in both sexes and toothed; blade of maxilla present. Second segment of palp with sensory organ. Antennal flagellum (with rare exceptions) with 13 segments in both sexes, the last three or four in the male lengthened.

Thorax: rounded; pronotum with small anterior division, placed low down and hidden between the head and scutum. Scutellum usually with bristles. Postnotum gently rounded. Sternopleurite not very prominent, reaching only a little below the tip of the front coxa.

Abdomen: Spermathecae strongly chitinised. Hypopygium with distinct cerci, parameres and aedeagus. Legs short and stout, the hind pair the longest. Hind tibia with double comb at tip. Pulvilli never present. Wings almost invariably superimposed over the back when at rest. Vein R_{2+3} absent. Media nearly always forked (except Leptoconops and Brachypogon). Alula scarcely indicated, this area of wing sometimes fringed. Squama small, never with a fringe of hairs.

In this definition, the two most significant features are the complete mouthparts and forked media.

In recent years, Wirth, Ratanaworabhan & Blanton (1974) have reviewed the classification of the family. They recognise four subfamilies: Leptoconopinae, Forcipomyiinae, Dasyheleinae and Ceratopogoninae. Separate family status for the Leptoconopinae (containing only Leptoconops) has been proposed by many modern workers, based on the extreme reduction of the larval head capsule.

The significance of larval and adult characters supporting this division has been reviewed by Downes (1977b).

The Forcipomyiinae contains two genera - Forcipomyia and Atrichopogon - and its classification, based on immature stages, is probably more advanced than any of the other three subfamilies.

The Dasyheleinae contains only one genus - Dasyhelea - which is intermediate in structure between the Forcipomyiinae and the Ceratopogoninae.

The Ceratopogoninae contain the remaining 59 genera of the family, grouped into seven tribes: Culicoidini, Ceratopogonini, Stenoxenini, Palpomyiini, Stilobezziini, Heteromyiini and Sphaeromiini.

The genera placed in the Culicoidini (including Culicoides) appear to be as primitive and non-specialised as any in the family and, together with the Ceratopogonini, may give more clues to the ancestoral lineage than other sections of the family (Wirth et al., 1974). The six tribes other than the Culicoidini are principally insectivorous in the adult stage and consequently exhibit a variety of modifications of the legs and mouthparts associated with this habit.

At present, the Culicoidini contains four genera: Culicoides Latreille, Neoculicoides Boorman & Lane, Paradasyhelea Macfie and Austroconops Wirth & Lee. Recently, the taxonomic validity of the tribe has been independently challenged by Remm (1975), and Boorman & Lane (1979). Remm (1975), using morphological, ecological, palaeontological and zoogeographic evidence, has proposed that the tribes Culicoidini and Ceratopogonini be united.

4.1.2. Culicoides and the subgenus Culicoides (s.s.)

The following definition of the genus follows Wirth (1952).

Genus Culicoides Latreille, 1809.

Culicoides Latreille, 1809 (General Crustaceorum et Insectorum, IV:251-252).

Type species Culicoides punctata (Meigen) by monotypy.

Body moderately slender, somewhat hairy. Eyes usually bare. Male antennae plumose, last three segments long; female antennae with segments iii - x rounded or oval, segments xi - xv more cylindrical and longer. Mesonotum usually dull, often with pruinose pattern; with short hair and often with longer bristles; humeral pits always large and distinct. Legs slender; femora

without spines; hind basitarsus at least twice as long as second segment, fourth segment shorter than fifth but rarely cordiform; claws small and equal in both sexes; empodium very short. Wings with dense microtrichia ; macrotrichia present, usually abundant, often confined to wing tip; wings often with spotted pattern; costa extending beyond middle of wing; anterior radial cells two, more or less equal; crossvein r-m slightly oblique; median fork distinctly petiolate, branches parallel; base of M_2 often interrupted; intercalary fork present as a fold close and parallel to M_1, a second fold within median fork from near base of M_2 to near tip of M_1; mediocubital fork widely open, M_{3+4} somewhat arched at base; anal vein straight. Male genitalia with ninth sternite short, emarginate on posterior margin; ninth tergite usually with apicolateral processes; basistyle usually with distinct ventral and dorsal roots at base; dististyle usually slender and curved; aedeagus usually a conical sclerite with distinct anterior lateral arms; parameres usually separate, bent or irregular sclerites, with slender distal points bent ventrad above tip of aedeagus.

Although the genus Culicoides has attracted considerable taxonomic attention, the general classification of the 1000 species is not very satisfactory, as Wirth et al. (1974) have summarised: "We still have a long way to go in the classification of Culicoides, for although many subgenera have been proposed, most of these are valid in only one particular geographic region and must be revised or supplemented to bring the other species of the world into the system".

The subgenus Culicoides (s.s.) was defined by Fox (1955) as: Female large, 1.5 - 2.5 mm. long; eyes contiguous or very close together. Wings with the second radial cell in a light spot and markings prominent, often with three broad transverse stripes, the middle one enclosing the basal portion of the cubital fork in a dark area; spermathecae double. Male hypopygium with ninth tergite rounded, often projecting at the midline, apicolateral processes usually reduced or absent, sometimes long but unsclerotised. Aedeagus various, if triangular without a basal marginal band. Inner margin of the basistyle with prominent setae. Ventral root short, not longer than the dorsal root. This subgenus has been recorded from the Nearctic, Neotropical (Fox, 1955),

Palaearctic (Gutsevich, 1973) and oriental regions.

As hinted above, not all Culicoides taxonomists have accepted the use of subgenera, turning to the alternative informal concept of species-groups. The studies of Khamala & Kettle (1971) on East African Culicoides and Kremer (1965) and Campbell & Pelham-Clinton (1960) on the Palaearctic fauna, are noteworthy examples of this approach.

In the Palaeartic region the subgenus Culicoides contains the following species: C. pulicaris (Linnaeus), punctatus (Meigen), newsteadi Austen (= halophilus Kieffer), impunctatus Goetghebuer, delta Edwards, grisescens Edwards, lupicaris Downes & Kettle, fagineus Edwards, nipponensis Tokunaga, variifrons Glukhova & Ivanov, flavipulicaris Dzhafarov, (see Gutsevich, 1973).

In his work on the bloodsucking midges of Transcaucasia (i.e., the Soviet republics of Armenia, Azerbiadjan and Gruzia) Dzhafarov (1964) described three new species - subgrisescens, achkamalicus and flavipulicaris, which he placed in the subgenus Culicoides. Gutsevich (1973) subsequently placed subgrisescens Dzhafarov in the subgenus Oecacta Poey, and suggested that the position of achkamalicus Dzhafarov was 'doubtful', but probably closely related to C. saevanicus Dzhafarov, also in the subgenus Oecacta Poey.

4.1.3. Definition of the Culicoides pulicaris group and the C. pulicaris complex

The Culicoides pulicaris group is widely distributed throughout the Holarctic Region and constitutes the basis of the subgenus Culicoides as defined by Fox (1955). One of the members of this group, C. punctatus (Meigen), is the type species of the genus Culicoides.

Although some of the species are extremely abundant, and include several pest species, their classification has remained confused and difficult. Generally, single specimens are difficult to identify with confidence, and therefore a series of specimens have to be used.

A number of definitions of the pulicaris group have been proposed. That of Kremer (1965) and Campbell & Pelham-Clinton (1960) is followed here:

Wing in both sexes with second radial cell ending in a pale area. Female with eyes touching or separated; no sensilla on antennal segments iv - x; third segment of maxillary palps with sensory hairs distributed over many small shallow excavations; two spermathecae. Male genitalia with small apicolateral processes; ventral processes short or absent; parameres separate with few small hairs at tip.

On this definition, the group is represented by nine nominal species in the Palaearctic Region: C. delta Edwards, fagineus Edwards, grisescens Edwards, newsteadi Austen (= halophilus Kieffer), impunctatus Goetghebuer, lupicaris Downes & Kettle, pulicaris (Linnaeus), punctatus (Meigen), and flavipulicaris Dzhafarov.

Unfortunately, material of C. flavipulicaris Dzhafarov was not available for the present study. The remaining eight taxa are the subject of this study and are referred to as the pulicaris complex.

Gutsevich (1973) included two other Palaearctic species in the subgenus Culicoides: nipponensis Tokunaga, which has sensilla on the proximal segments of the antenna, and variifrons Glukhova & Ivan, which has the sensory hairs in a well defined pit on the third palp segment. Neither of these species comes within the definition of the pulicaris group mentioned above.

Wirth & Blanton (1969) reviewed the pulicaris group in North America, recognising 15 species, nine of which were described as new, based mainly on numerical characters. In their definition of the group (1969), they included the character of 'presence of sensilla on antennal segments iv - x ,to encompass a new species from Utah.

The Culicoides pulicaris group has in the past been divided into a number of sections, mainly by North American workers, none of which have proved useful. For example, Wirth & Blanton (1969) split the group into two sections - the Cockerelli section and the Pulicaris section. The former included those species without a dark patch in the middle of wing cell M_4; dark apices of wing veins (except R_s), and a convex margin to the ninth tergite in the male. However, as the authors suggest, the division does not work well - for impunctatus came within the Cockerelli section on wing pattern, but within the Pulicaris section on male genitalia. Working on Palaearctic species, Kremer (1965) managed to circumvent this problem to some degree by dividing the whole species group into

three sections. Essentially this split the Cockerelli section of Wirth & Blanton into two - an Impunctatus section and a Grisescens section, thus eliminating the ambiguous position of impunctatus and delta described above, but produced a section to include grisescens and fagineus. This contained two species united by their dissimilarity to the rest of the pulicaris group, rather than their similarity to one another. Further comments on the validity of these groups, as revealed by multivariate morphometrics, will be discussed in Section 9, p.274.

4.2. FORMAL DESCRIPTION AND NOMENCLATURE

4.2.1. Culicoides pulicaris (Linnaeus)

Culex pulicaris Linnaeus, 1758 (Syst. Nat., 1(10th Edn), p. 603).
Culicoides setosinervis Kieffer, 1913 (Bull.Soc.Hist.nat.Metz, 28:8).
Culicoides pulatus Kieffer, 1915 (Arch,Hydrobiol., 2:474).
Culicoides stephensi Carter, 1916 (Ann.trop.Med.Parasit., 10:135).
Culicoides cinerellus Kieffer, 1919 (Annls hist-nat.Mus.natn. Hung., 17:40).
Culicoides quinquepunctatus Goetghebuer, 1921 (Mem.Mus.r.Hist. nat.Belg., B:177).
Culicoides flavipluma Kieffer, 1924 (Bull.Soc.Hist.nat.Metz, 30:19).
Culicoides pulicaris var A Edwards, 1926 (Trans.ent.Soc.Lond., 74:406).
Culicoides pulicaris ssp. kasachstanicus Schakirzjanova, 1963 (Izv.Akad.Nauk kazah.SSR, 1963:63).

Edwards (1926) treated C. pulicaris in a broad sense, distinguishing four varieties (A - D) based on wing pattern alone. These varieties have subsequently been raised to species rank (pulicaris, punctatus, newsteadi (as halophilus) and delta). Edwards' first variety was arbitarily chosen as representing the typical pulicaris, a move followed by subsequent taxonomists. In 1939, Edwards wrote "The precise identity of the midge to which Linnaeus actually applied the name Culex pulicaris is to some extent a matter of conjecture, but it was certainly one of the larger Culicoides, and the name is now in general use for a common species which has somewhat milky wings, clothed with hair on the greater part of their surface, and with rather sharply defined dark markings, which include three blackish spots on the

front margin and a small dark spot in the middle of the cubital fork".

In Britain, the taxon recognised by Edwards as the typical pulicaris, is the most widely distributed and abundant. It is also the most variable (Campbell & Pelham-Clinton, 1960).

This species is very variable in size, although generally large. Wing length = 1.57 (1.39 - 1.73, s.d.= 0.16)mm. The wings are clear whitish with dark spots. There are three costal spots equally spaced along the anterior margin. The first of these extends a little beyond the medial vein; the second covers the distal part of the first radial cell and the proximal part of the second radial cell, and extends down to join a dark area over the medial fork. The shape of the distal spot is an important characteristic in distinguishing this species from C. punctatus. In pulicaris it is like an hour-glass in shape and broadest on the wing fold in cell R_5. (In punctatus, the hour-glass shaped spot is broadest above the fold). The two branches of the medial fork (M_1 and M_2) are dark to the extreme tips. There is a distinct spot in the cubital cell, and dark areas along both its bordering veins. A triangular spot extends down from vein An_2 and a cresent shaped mark curves up towards the cubital fork. The wing pattern varies considerably and is therefore not always as distinct as this description may imply. Variation in the wing pattern of pulicaris and punctatus is much the same. An example of the variation in pulicaris wing pattern at one site in Cyprus is given in Fig.20 . The pattern ranges in a continuous manner from a reduced form (often call var. setosinervis Kieffer) through an intermediate state to the normal pattern.

The pigmentation of the mesonotum is very variable. It may be unmarked or have three small and usually separate spots.

The male genitalia are structurally indistinguishable from punctatus, delta and lupicaris and are only separable from newsteadi and impunctatus on size.

The eyes of the female are usually touching, with a superior transverse suture above the juncture of the eyes. The presence of the suture may be used to separate members of the pulicaris group from the closely related obsoletus group, although it is not always reliable. The third palp segment of the female is moderately swollen. Palp ratio = 2.71 (1.90 - 3.88, s.d. =0.38).

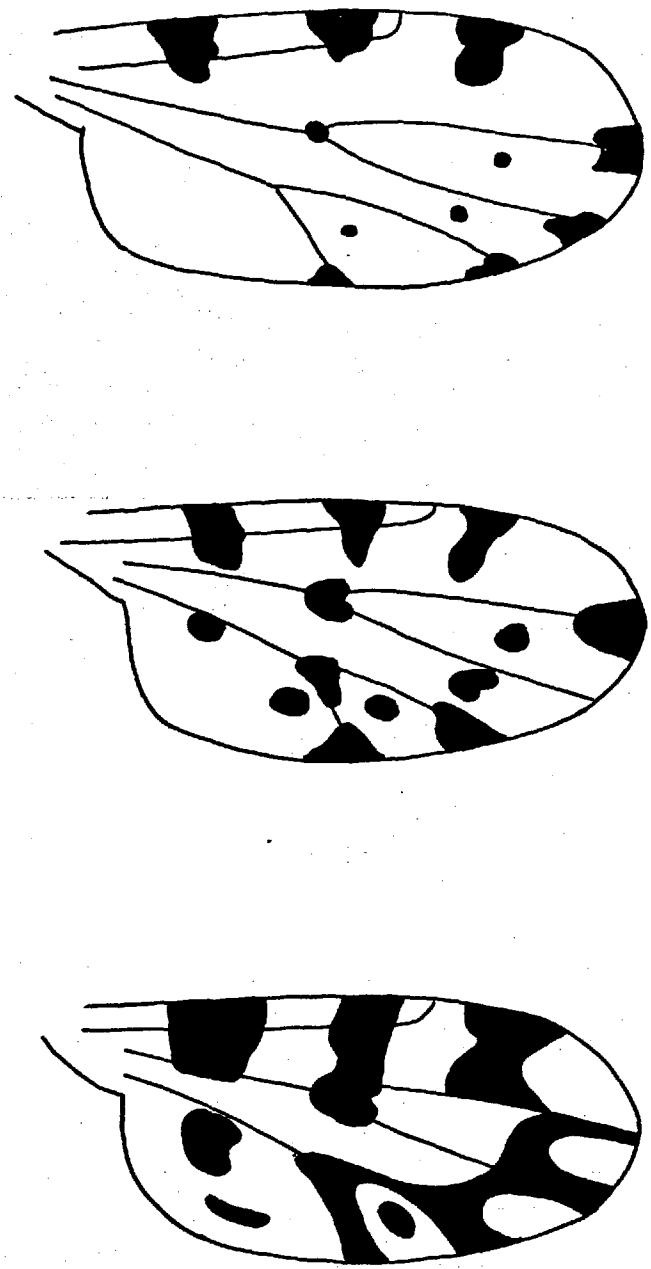


FIG. 20
VARIATION IN WING PATTERN OF C. PULICARIS
FROM CYPRUS

Antennal ratio = 1.10 (0.96 - 1.18, s.d.= 0.06).
Hind tibial comb with 5 to 7 spines.

Recorded Distribution

Owing to the grouping of several species under the name C. pulicaris, during the early part of this century, only recent records are quoted here. However, this does not completely alleviate the confusion, since some taxonomists (Arnaud, 1956; Dzhafarov, 1964; Gutsevich, 1973) still treat pulicaris in a broad sense, to include punctatus and lupicaris.

France: Kremer (1965); Spain: Havelka (1979); Callot, Kremer, Rioux & Descours (1967); Corsica: Kremer, Leberre & Beaucournu (1971); Switzerland: Kremer & Callot (1961); Germany: Havelka (1976); Denmark: Nielsen (1964); Poland: Kremer, Doby & Skierska (1965), Bilinsky (1964); Czechoslovakia: Orszagh(1977), Knoz (1977); Hungary: Remm (1973), Zilahi-Sebess (1936); Morocco: Kremer, Hommel & Bailly-Choumara (1971); Cyprus: Boorman (1974); Israel: Braverman, Boorman & Kremer (1976); Iraq: Khalaf (specimens uncertain, 1961); USSR:- Karelia: Glukhova (1962); Ukraine: Schevtshenko (1967); Crimea: Remm & Zhogolev (1968); Turkmenia: Muradov (1965); Tuva ASSR: Violovich (1965); Ussuri: Amasova (1957); Kirgiz ASSR: Kornubayev (1965);Azerbiadzhan: Dzhafarov (1964); Koni ASSR: Belokur (1960); West Siberia: Mirzayeva (1963); Primorje Territory: Ivanov & Glukhova (1967); Korea: Arnaud (1956); Japan: Arnaud(1956).

This species was recorded by Tokunaga (1937) from Japan, but the wing illustrated lacked a spot in the cubital cell. This is most unlike pulicaris and more like delta or grisescens. However, both these species lack pale rings at the base of the tibia which Tokunaga (1937) describes and the palps are not as swollen as Tokunaga's figure illustrates. Such inflated palps are found in C. newsteadi but the wing pattern is different from this species. Attempts to locate the material on which Tokunaga based his description revealed that the collection has been lost and therefore the exact identity of this interesting material must remain in doubt.

Arnaud (1956) is the only recent author to have given a full description of this species from outside Europe. His description and illustration of Japanese material exemplifies the problem of using wing patterns to distinguish between this species and punctatus. Arnaud's illustration of the wing shows the distal

costal spot to be of the punctatus type, but agrees with pulicaris in the lack of white spots at the tips of veins M_1 and M_2. Furthermore, the mesonotum is more similar to that of the punctatus type than that of the pulicaris type. Arnaud (1956) also cites the 'pulicaris' of Tokunaga discussed above and C. pulicaris ocellaris (a synonym of punctatus) as synonyms of C. pulicaris. It is therefore reasonable to suppose that Arnaud has taken a broad interpretation of pulicaris as a species, and that his comments on distribution and biology should be considered with this in mind.

As noted above, Gutsevich (1973) and Dzhafarov (1964) regard pulicaris in a wide sense to include punctatus. Campbell & Pelham-Clinton (1960) suggest that the figured wing of Gutsevich (1952) is typical of punctatus and therefore there is little evidence to show that pulicaris extends beyond Europe and the Mediterranean regions. Specimens of typical pulicaris have been received from Gutsevich for inclusion in this study, and show that the range of pulicaris does indeed extend well into the Soviet Union.

4.2.2. Culicoides punctatus (Meigen)

Ceratopogon punctatus Meigen, 1804 (Klass.Beschr.Eur.Zweifl., 1:29).
Culicoides pulicaris var. ocellaris Kieffer, 1921 (Annls Soc. scient.Brux., 40:276).
Culicoides pulicaris var. B Edwards, 1926 (Trans.ent.Soc.Lond., 74:406).

This species was designated the type species of Culicoides by Latreille (1809, p.251).

Culicoides punctatus is very closely related to C. pulicaris and is often considered no more than a variety. The principal differences lie in the details of the wing pattern. In punctatus, the hour-glass shaped costal spot is broadest above the longitudenal fold in cell R_5 (in pulicaris, this spot is broadest on, or below, the fold). Other distinguishing characters of punctatus are the presence of white areas at the tips of veins M_1 and M_2, and occasionally M_{3+4}; the costal spot over the radial cells does not usually extend down to meet the medial fork; no distinct triangular spot based on vein An_2, but a spot is present which is usually ill-defined and usually fused with dark marks in the anal cell. Like C. pulicaris, the wing pattern varies a great deal. Downes & Kettle

(1952) point out that the type of mesonotal marking is characteristic: these markings are nearly always well developed and distinguishable from those of pulicaris by their very angular outline. The male genitalia are similar to those of C. pulicaris.
Wing length = 1.55 (1.41 - 1.79, s.d. =0.118)mm.
Palps in the female are also similar to C. pulicaris.
Palp ratio = 2.78 (2.00 - 3.90, s.d. =0.33). Antennal ratio = 1.11 (0.86 - 1.21, s.d. =0.07).
Recorded Distribution

The distribution of this species may extend further than is given here, as some workers have recorded this species under the name C. pulicaris. Where figures of wings are given in the original records, these havebeen used to establish the identity of the taxa concerned.

Spain: Havelka (1979); France: Kremer (1965), Callot, Kremer, Rault & Bach (1966); Switzerland: Callot, Kremer & Deduit (1962); Germany: Havelka (1976); Italy: Callot, Kremer & Coluzzi (1965); Hungary: Zilahi-Sebess (1933); Czechoslovakia: Orszagh (1977), Knoz (1977); Denmark: Nielsen (1964); Rumania: Albu & Georgescu (1971); Morroco: Callot, Kremer & Bailly-Choumara (1968); Ukraine: Gutsevich (1964); Cyprus: Boorman (1974); Israel: Braverman, Boorman & Kremer (1976); Crimea: Remm & Zhogolev (1968); Mongolia: Tokunaga (1940); Manchuria: Takahashi (1941); Iraq: Khalaf (1961); Japan: Arnaud (1956).

Wing pattern varies enormously and samples from Sweden and Norway (for details, see newsteadi) are difficult to assign to either punctatus or newsteadi. The main difference is the presence of an extra band of dark pigment at the base of cell M_1, and a higher palp ratio in newsteadi.

4.2.3. Culicoides impunctatus Goetghebuer

Culicoides impunctatus Goetghebuer, 1920 (Mem.Mus.r.Hist.nat. Belg., 8:55).
Culicoides arcuatus Edwards, 1926 (not Winnertz), (Trans.ent.Soc. Lond., 74:136).

This is the smallest member of the pulicaris group. Its size, together with the unmarked thorax and absence of a dark spot in the

cubital cell, distinguish most specimens of the species from others of the group.

Wing length = 1.26 (1.09 - 1.43, s.d. =0.106)mm.

The second and apical spots are confluent with other dark areas to form two continuous dark bands, running from the anterior to the posterior part of the wing. Thorax usually without any markings, rarely, there are two small 'comma-shaped' patches centrally. Third palp segment in female not markedly swollen. Palpal ratio = 2.83 (2.04 - 4.11, s.d. =0.38). The third palp segment with few sensory depressions. Antennal ratio = 1.01 (1.00 - 1.09, s.d. =0.034). Eyes of females touching for a variable distance, usually about one facet's length. As in other members of the group, there are two functional spermathecae plus one vestigial one. Occasionally, the vestigial spermatheca is enlarged to become approximately the same size as those that function. Hind tibial comb with five or six spinea.

The male genitalia are similar to those of pulicaris, although the inner swelling of the coxite is less pronounced. The style is generally not longer than 110 μm, whereas in pulicaris, punctatus, delta and lupicaris it is not less than 120 μm. The somewhat dubious use of size as a diagnostic character is discussed in Section 7.2.

Recorded Distribution

This species occurs in montane and submontane zones. In Britain, it is more common in Scotland and northern England than in the south.

Spain: Havelka (1979); France: Kremer (1965); Belgium: Goetghebuer (1920); Germany: Kremer (1965); Denmark: Nielsen (1964); Sweden and Norway: Kremer, Geiss, Delecolle & Hommel (1975); Czechoslovakia: Orszagh (1977); Poland: Kremer, Doby & Skierska (1965); Rumania: Albu & Georgescu (1971); USSR:- Karelia, Estonia, Caucasus, Lenningrad district, Novgorod district, Tumen ASSR, Donskaya district, Koni ASSR, Kol'skiy Plouostrov Peninsular, Primorskiy Maritime Province: Dzhafarov (1964); Japan: Tokunaga (1941).

4.2.4. Culicoides newsteadi Austen

Culicoides newsteadi Austen, 1921 (Bull.ent.Res., 12:113).
Culicoides halophilus Kieffer, 1924 (Archs Inst.Pasteur Alger., 2:404).
Culicoides biclavatus Kieffer, 1924 (Bull.Soc.Hist.nat.Metz, 30:141).
Culicoides pulicaris var. C Edwards, 1926 (Trans.ent.Soc.Lond., 74:406).

Culicoides var. edwardsi Goetghebuer, 1933 (Fliegen palaearkt. Reg., 13a:46).

Culicoides pulicaris var. edwardsianus Goetghebuer, 1933 (Bull.Soc. ent.Belg., 73:367).

This is a small species with a wing length of 1.38 (1.09 - 1.73, s.d. =0.183)mm. The wing pattern is very similar to C. punctatus but generally less distinct. The principal difference is the presence of an extra dark spot at the base of the medial cell. As in punctatus, there is a dark area over the fork of M_1 and M_2, which extends below M_2, then curves upward through the medial cell and joins the darkened area along M_1, thus forming an extra dark area in the cell. There is some variation in the development of this spot, illustrated by Kremer (1965). C. newsteadi and C. punctatus are often difficult to distinguish on structural characters as well as wing characters. Samples examined from Sweden and Norway (both from well within the Arctic Circle) have structural characters e.g., palp ratio, intermediate between typical newsteadi and punctatus. For a further discussion on these specimens, refer to Section 7.4 (geographical variation) and Section 9.6. (morphometrics).

Eyes in the female touching by a short distance. Third palp segment of female swollen, with the lowest palpal ratio in the complex = 2.26 (1.66 - 2.60, s.d. =0.26). Antennal ratio = 1.06 (1.00 - 1.18, s.d. = 0.05). Tibial comb with five spines. Male genitalia as in pulicaris. Callot, Kremer & Braverman(1969) describe a gynandromorph of this species.

Recorded Distribution

France: Kremer (1965); Germany: Thienemann (1954); Denmark: Nielsen (1964); Czechoslovakia: Orszagh (1977); Morocco: Callot, Kremer & Bailly-Choumara (1968); Cyprus: Boorman (1974); Poland: Bilinski(1968); USSR:- Gruzia ASSR, Armenia ASSR, Crimea, Latvia, Tadzikistan, Chechino-Inushskaya ASSR: Dzhafarov (1964); Iraq: Mesghali (1963).

The immature stages of this species have always been found in brackish coastal habitats, although the adults have been collected well inland on several occasions e.g., Orszagh (1977), Mesghali (1963) and Dzhafarov (1964). This apparent anomaly may be due to two causes. Firstly, owing to the difficulty in separating punctatus from newsteadi, these records may be based on mis-identifications. Secondly, the saline larval habitat may be

available inland. Rieb & Kremer (1977) found the halophilic species C. circumscriptus Kieffer and C. salinarius Kieffer in mud along the river Ill (Alsace), where the water was contaminated from potash mines. In the neighbouring region of Lorraine, where there has always been salt (sodium chloride) in the water, they found all the European species of salt marsh Culicoides.

The synonymy of halophilus and newsteadi was first proposed by Edwards (1939), who suggested that the names would be best applied to the northern and southern forms of a single species. This has subsequently been supported by several authors, notably Kremer (1965), Bailly-Choumara & Kremer (1970) and Boorman (1974). It was not until Kremer et al. (1975) that the names were formally synonymised.

4.2.5. Culicoides delta Edwards

Culicoides delta Edwards, 1939 (in Edwards, Oldroyd & Smart, British Bloodsucking Flies, B.M.(N.H.) p.48, 145).

Culicoides pulicaris var. D Edwards, 1926 (not Goetghebuer) (Trans. ent.Soc.Lod., 74:407).

This taxon was first recognised by Edwards in 1926 when he divided C. pulicaris into four varieties. Its distinguishing features are the lack of spot in cubital cell; pigment along veins bordering cubital cell and filling base of fork; dark area in anal cell along posterior edge of wing; no dark pigment along complete length of M_1 and M_2 . The wing pattern is very similar to that of impunctatus, differing in the larger size of delta. Wing length = 1.82 (1.67 - 2.09, s.d. =0.137)mm.

The male genitalia show no obvious difference to pulicaris, lupicaris and punctatus. In the female , the third palp segment is relatively slender. Palp ratio = 3.23 (2.55 - 4.16, s.d. = 0.47). Antennal ratio = 1.09 (1.02 - 1.20, s.d. =0.05).

The species was described from specimens collected by F.W. Edwards in Scotland, but the type description does not indicate how many specimens were included in the type series. The localities of the type series were also vague, therefore the following details of holotype and paratypes, based on material in the British Museum collection, is given to help clarify this point.

Holotype♂, SCOTLAND: Arran, Brodick, 22-25.v.1919 (F.W.Edwards).

Paratypes. SCOTLAND: Arran, 3♀, Brodick, 22-25.v.1919 (F.W.Edwards); 1♀, Catacol, 29-30.v.1919 (F.W.Edwards); 1♀, Corriegills, 2-4.vi.1919 (F.W.Edwards); 1♀, Sannox, 26-28.v.1919 F.W.Edwards). Perthshire, 2♀, Loch Rannoch, vi.1931; 1♂, Loch Kinardochy, vi.1931.

The specimens above are believed to constitute the entire series of type material.

Recorded Distribution

Germany: Havelka (1976); Czechoslovakia: Orszagh (1977), Knoz (1977); Poland: Skierska (1972); Rumania: Georgescu (1972).

Kremer & Callot (1961) recorded this species from Grimentz in Switzerland, but Kremer (1965) notes that these specimens were misidentified and should be assigned to lupicaris.

Campbell & Pelham-Clinton (1960) suggested that during the early part of the year, when delta, pulicaris and punctatus were flying, the following features were sufficient to separate delta from the other members of the complex: absence of spot in cubital cell, extensive dark mesonotal markings, and a yellowish tinge to the wings. This apparent distinction between the species becomes more unreliable as the year progresses. Campbell & Pelham-Clinton also note that during the summer months, the incidence of delta specimens with spots in the cubital cell increases. They also describe the wide range of variation in the mesonotal markings of C. pulicaris, a further factor which hinders the clear separation of delta from pulicaris.

The distinction between C. delta and C. lupicaris is particularly difficult, and has led some authors to synonymise the two under the name delta. Campbell & Pelham-Clinton (1960) recognise this problem and suggest that "any specimens taken after the end of June, which key to delta, should probably be referred to as lupicaris". A more detailed discussion of these two species is given under the species description of C. lupicaris.

4.2.6. Culicoides fagineus Edwards

Culicoides fagineus Edwards, 1939 (in Edwards, Oldroyd & Smart, British Bloodsucking Flies. B.M.(N.H.): 147-148).

A generally small species. Wing length = 1.47 (1.34 - 1.67, s.d. =0.129)mm. Wings yellow by reflected light. Pattern shows little variation, probably the least of any species within the

complex, but since only a few specimens are known, this is difficult to be certain of. Wing markings well defined. Vein M_2 is free from pigmentation on its basal half, and by this character fagineus may be separated from impunctatus, in which the whole of M_2 is darkened. Cubital cell without a spot (a character found unreliable by Campbell & Pelham-Clinton, 1960). Pattern on thorax as in pulicaris and delta but less well defined. Legs brownish, with pale areas at tips of femur and pale tarsi. Hind tibia with pale tip and base. Eyes of female in contact. Hind tibial comb usually with six spines. The females of fagineus may be distinguished from all other species in the complex by the presence of vertical teeth in the cibarium. Male genitalia with distinctive ninth tergite; apicolateral processes very short, distal margin between processes convex. Rest of genitalia as in pulicaris.

Recorded Distribution

France: Callot, Kremer, Rault & Bach (1966); Italy: Coluzzi & Kremer (1964); Czechoslovakia: Orszagh (1977), Knoz (1977); Algeria: Clastrier (1957); Morocco: Kremer, Hommel &Bailly-Choumara (1971); Israel: Braverman, Boorman & Kremer (1976); USSR:- Crimea: Remm & Zhogolev (1968); Ussuri: Amasova (1957); Azerbiadshan, Gruzinia ASSR, Tuvinskaya A.O., Lenningrad District, Primorski Province: Dzhafarov (1964).

The larval habitat and cibarial teeth in the female render the position of this species doubtful as a member of the pulicaris complex.

4.2.7. Culicoides grisescens Edwards

Culicoides grisescens Edwards, 1939 (in Edwards, Oldroyd & Smart, British Bloodsucking Flies. B.M.(N.H.): 146-147).

Culicoides impunctatus Tokunaga, 1941 (not Goetghebuer) (Insecta matsum., 15:97).

Culicoides impunctatus var. minor Tokunaga, 1941 (Insecta matsum., 15:97).

This species is one of the largest in the pulicaris complex. Wing length = 1.75 (1.37 - 2.01, s.d. =0.184)mm. The wing markings are usually diffuse, not showing the contrast between the dark and pale areas typical of other species in the complex. Cubital cell without a spot. The wing markings are similar to those of C. impunctatus

and C. delta in general appearance. The species may be separated from impunctatus by size alone and from delta by the presence of a large basal costal spot, confluent with the dark area on the posterior area of the wing, (thus forming a continuous dark band from anterior to posterior of wing). There are three pronounced costal spots, the apical spot often confluent with pigment on vein M_1. The thorax is more or less unmarked, similar to impunctatus in this respect. Occasionally, small wedge-shaped patches are found behind the humeral pits. Third palp segment of female usually slender. Palpal ratio = 3.99 (3.28 - 4.52, s.d. =0.40).
Hind tibial comb with five or six spines.
The male genitalia are the most distinctive of the group, with the ninth tergite having a strongly convex posterior margin without a median notch.

Recorded Distribution

Most commonly found in montane and submontane zones.

France: Kremer (1965); Switzerland: Kremer & Callot (1961); Czechoslovakia: Orszagh (1977); Poland: Kremer, Doby & Skierska (1965); Bilinski (1968); Denmark: Nielsen (1964); Rumania: Albu & Georgescu (1971); USSR:- Krasnoyarsk Region: Glukhova & Berzina (1963); Tuva ASSR: Violovich (1965); Murmansk Province: Solovey & Likhoded (1966); Estonia, Karelia, Lenningrad District, Vladimir District, Ryazan District, Carpathians, Kazikistan, Komi ASSR, Baykal, N. Caucasus, Azerbiadshan, Gruziya, Armenia: Dzhafarov (1964); Iran: Mesghali (1963); Manchuria: Tokunaga (1941).

Kremer, Doby & Skierska (1965) note that specimens they collected in northern Poland were intermediate between grisescens and pulicaris. The specimens collected by Solovey and Likhoded (1966) are claimed to be the most northerly collection of any species of Culicoides (69°24'N.).

4.2.8. Culicoides lupicaris Downes & Kettle

Culicoides lupicaris Downes & Kettle, 1952 (Proc.R.ent.Soc.Lond.(B), 21:76-77).

A large species of the pulicaris complex, closely allied to pulicaris and delta. The name is an anagram of pulicaris.
Wing length = 2.17 (1.82 - 2.41)mm. (data from description).
Wing pattern is similar to pulicaris. Second costal spot not

prolonged under the second radial cell and third costal spot hour-glass shaped, widest on fold of wing above vein M_1. Dark areas extending along veins M_1 and M_2 to tip; well marked spot in cubital cell. In all these features and general colour of wing, intensity and definition of markings very like pulicaris. It differs from pulicaris by the pigmentation in the anal region, which extends along the hind margin, and does not rise up from the margin as it approaches the cubital cell. The extent of the wing markings varies, but not to the extent of pulicaris or punctatus (Downes & Kettle,1952).

The mesonotum is heavily marked with large tridentate marks, by which it differs from pulicaris, punctatus and newsteadi. Male genitalia show no differences from other members of the group, except the distinctive grisescens and faqineus.

Recorded Distribution

France: Kremer (1965); Corsica: Kremer, Leberre & Beaucournu-Saguez (1971); Switzerland: Kremer & Callot (1961); Germany: Kremer (1965); Czechoslovakia: Orszagh (1977); Poland: Skierska (1972); Rumania: Georgescu (1972); Estonia: Remm (1956); Ukraine: Gutsevich (1964).

Downes & Kettle (1952) suggest that the only difference between lupicaris and delta is the presence of a spot in the cubital cell of lupicaris. The thoracic pigmentation and anal area of the wing are the same. Campbell & Pelham-Clinton (1960) redefined the species, using mesonotal markings, described as: triangular medio-lateral marks and short median vittae, and by this means, claimed the ability to distinguish Scottish specimens of lupicaris and delta. They also initially regarded lupicaris as a form of delta, but concluded that it should have specific status, as Kettle & Lawson (1952, p. 443) found the larvae to be distinct. However, in their redefinition of lupicaris, Campbell & Pelham-Clinton do not state whether differences in the adults of newly defined species correlate with differences in their larvae.

When Campbell & Pelham-Clinton regrouped their Scottish specimens (based on newly defined species), they were then able to show a statistical difference in wing length between lupicaris and delta. Working with Russian specimens, Gutsevich (1973) found that the ranges of variation in wing length overlapped considerably between the two species (although he gives no statistical parameters). He supported his synonymy of lupicaris with delta by describing

the variation in wing and thoracic pattern. He found that in samples from the Carpathians, the presence of a spot in the cubital cell (one of the principal distinguishing features) varied a great deal. Some specimens were found in which the spot was present on one wing, but absent on its complement. He also noted that the variation in thoracic markings was too great to be considered as reliable scientific characters. Dzhafarov (1964) regarded C. lupicaris as a subspecies of C. pulicaris together with C. punctatus.

4.3. GEOGRAPHICAL DISTRIBUTION

The species complex is widely distributed in the Palaearctic Region, some species having been recorded many times. For brevity, only the limits of distribution are discussed.

Gutsevich (1973) recorded pulicaris (including punctatus), grisescens, delta, newsteadi (as halophilus), fagineus and impunctatus from various parts of the Soviet Union. In the Far East, Tokunaga (1941) worked on a collection of midges from Manchuria and recorded two species; C. punctatus (as ocellaris Kieffer) and grisescens (as impunctatus and impunctatus var. minor Tokunaga).

At the southern end of the range, Braverman et al. (1976) recorded fagineus, pulicaris, punctatus and newsteadi from Israel. Callot, Kremer & Bailly-Choumara (1968) recorded pulicaris and newsteadi from the Atlantic coast of Morocco.

Members of the pulicaris group have been recorded from the Afrotropical (= Ethiopian) region on two occasions. Macfie(1937) recorded a single male of 'pulicaris' from Ethiopia. Further examination of this specimen in the BM(NH) collection showed the genitalia and wing to resemble the African species C. brucei Austen and C. magnus Khamala & Kettle, and not pulicaris. Although these two species can only be reliably separated on the distribution of sensilla on the female antennae, the genitalia of this specimen most resemble brucei in the indentation of the ninth tergite, and broad aedeagus.

The second Afrotropical record, by Clastrier (1959), concerns a series of five female 'punctatus' from the Ivory Coast. Examination of these specimens showed the wings to be very clearly marked with distinct pale spots at the tips of veins M_1, M_2 and M_{3+4} in some specimens (the characters which led Clastrier to

identify them as punctatus). However, the distribution of sensilla on the antenna and features of the wing pattern identify them as brucei Austen.

C. brucei belongs to the C. magnus species group as defined by Khamala & Kettle (1971). This species group is the nearest taxon to the subgenus in the Afrotropical region, but as species groups are the only infrageneric rank used for this region, the recognition of the subgenus Culicoides in Africa must await further taxonomic studies.

As noted in Section 4.1.3., the pulicaris group has a Holarctic distribution, extending into the Nearctic region, where it is represented by 15 species. The link between Palaearctic and Nearctic members of the pulicaris group is to be found in Greenland, where a North American species, C. sordidellus Zetterstedt, occurs. This species is closely related to the Palaearctic C. grisescens.

There are many comparable Holarctic distributions in the Nematocera (e.g. in the Chironomidae: Fittkau & Reiss, 1978) with certain species having circumpolar distributions (e.g., Lindeberg, 1971, on Tanytarsus sp.). It would not be unreasonable to expect that some members of the pulicaris group were also circumpolar, but this does not appear to be the case. Wirth & Blanton (1969) compared the adults of the North American and European species carefully with rather inconclusive results, but found that none of the American species conclusively matched those from Europe. This unexpected observation presents many interesting problems, but is beyond the scope of this study.

4.4. IMMATURE STAGES

The larvae of some species of the pulicaris group have been studied by Kettle &Lawson (1952). They found a similar problem with the identification of the larvae as others have found with the adults. They wrote "Larvae are very similar to each other so that their range of variation overlap and it is often impossible to determine a specimen with certainty". This has been found in other species groups of Culicoides (Glukhova, 1977).

Although they only had small samples for some species, Kettle & Lawson (1952) were able to find some characters in the thoracic pattern of the larvae to distinguish impunctatus and lupicaris from other species. This is of particular interest with respect

to lupicaris as it is almost impossible to distinguish it from pulicaris in the adult stage. Some evidence of differences between other species was found in head length and breeding site. C. newsteadi (as halophilus) was only found in brackish water, whist the other species occurred in marsh and swampy ground.

A few differences have been found between the species in the pupal stage by Kettle & Lawson (1952); body size; pigmentation of abdominal segments; number of papillae on the respiratory horn; shape of tubercles on the abdomen. As with larvae and adults, it is unfortunate that the distinctions are not clear. Glukhova (1977) found that as a general rule, pupae of Culicoides are more uniform in structure and have less stable characters than larvae.

4.5. BIOSYSTEMATIC DATA ON THE C. PULICARIS COMPLEX

Of the eight nominal taxa studied here, relatively little detail is known of the biology for several of them, e.g., the host ranges have not been investigated for any of the species. Biological characteristics have been used to support some taxonomic decisions, for example, specific status of punctatus, but have generally proved inconclusive, especially when compared over the whole distribution of the species concerned.

The following is a brief review of the biological attributes which have been used.

C. newsteadi (as halophilus) is distinct among the complex in that its larvae have always been found in saline habitats. Reasons for the apparent deviations from this distinct habitat, implied by inland collection of adults, is discussed in Section 4.2.3.

Although the larvae of C. fagineus have not been described, the species has been bred on two occasions, once from a rot hole in Beech (Edwards, 1939), and once from a rot hole in Elm (Coluzzi & Kremer, 1964). The species appears to be the only one of this complex which does not breed in soil, amongst growing plants. No other members of the complex have been bred from rot holes. The host preference of fagineus is not known. It has never been collected biting man or domestic animals. This negative evidence, together with the presence of cibarial teeth (the only species of the complex to have these) suggests that C. fagineus may feed on a very different host to the rest of the complex,

possibly birds or reptiles.

The breeding sites and adult behaviour of C. impunctatus have been well studied (Kettle, 1950 and references cited in Section 4.6.), and emphasise the distinct nature of this species. Kettle (1950) found that the seasonal distribution of impunctatus at Loch Lomond was bimodal, suggesting that this revealed the existance of two biological races. Critical examinations of pattern and wing length in females and the genitalia of the male by Kettle failed to reveal any morphological difference between the two adult populations. However, a statistical difference in sex ratio and vertical and horizontal distribution was found. Much contradictory evidence has been gathered from different sites as to whether impunctatus is bivoltine or univoltine (see Section 4.6. for references). Onyiah (1971) found that impunctatus was unimodal in three sites during one year, but bimodal in the same sites in the following year. He attributes this to the development of immature stages under different environmental conditions.

Biological data supporting the recognition of pulicaris and punctatus as separate species has received much attention in the literature on British Culicoides. Edwards (1939) regarded punctatus as a variety of pulicaris, on morphology alone, and it was not until 1952 that Downes & Kettle brought biological considerations into the discussion. They describe a number of observations to substantiate their elevation of punctatus to specific rank. The evidence may be summarised as follows:

1. Although the larvae of both species normally occur on the same ground, the microhabitats selected by each species appears to be different. They sampled one breeding site using a 3 inch diameter core sampler and proved that neither species is uniformly distributed over the breeding site. Most core samples contained larvae of both pulicaris and punctatus, but in differing proportions, ranging from 7-100% pulicaris with a mean of 52% pulicaris over all samples. Whether this is biologically significant will require further study, as both species have such a wide range of larval habitats.

2. Three swarms of males were observed in habitats where both species occur. The first two swarms consisted of 10♂ pulicaris and 16♂ pulicaris respectively, and the third was composed of 3♂ punctatus.

Campbell & Pelham-Clinton (1960) attempted to use the biting habits of pulicaris and punctatus as biological characters to distinguish samples of these two species. In their experience, British pulicaris and punctatus rarely bite man, and used this as evidence that non-anthropophilic Japanese species, C. sawamotoi Kono & Takahashi, is not synonymous with either pulicaris or punctatus. However, there are a number of British specimens of both species in BM(NH), collected biting man. Edwards (1939) records pulicaris as commonly biting man in Britain, and Kremer (1965) records both species as troublesome in France. In this example, it would seem that the use of biological characteristics such as host selection has not proved very reliable.

Campbell & Pelham-Clinton (1960, p.277) found that whenever collections of pulicaris and punctatus were taken separately from the backs and bellies of cattle, collections from the back consisted almost entirely of pulicaris and those from the belly almost entirely of punctatus. Nielsen(1971), however, collected Culicoides biting cattle in Denmark, and found punctatus biting legs, belly and rump. This contradictory observation is interesting in that Nielsen found no pulicaris in his collections.

Close examination of the two species does not reveal any structural difference associated with moving through varying length and thickness of hair on the rump and belly of a host cow.

A difference in frequency of males of the two species captured in suction traps was noted by Campbell & Pelham-Clinton (1960, p.278). Males of punctatus were frequently trapped, but males of pulicaris appeared in the traps only singly or in small numbers at a time, although females of both species were numerous on the site.

In conclusion, there are reasonable data on the larval habitats of species such as newsteadi and impunctatus to show that they are quite distinct biologically. The data supporting the specific status of punctatus is not so clear, and is often contradictory.

The paucity of biological information for some of the species in the complex does not allow many biological characteristics to be assessed in relation to the taxonomy of the complex, but rather reveals how much further work needs to be done.

4.6. APPLIED IMPORTANCE OF THE PULICARIS GROUP

The C. pulicaris group contains a number of species which are

of considerable annoyance to man. In northern Britain and much of northern Europe, C. impunctatus is the most important biting species of Culicoides. In Scotland, although this species may constitute only 60% of the Culicoides population on the wing, it is often the only one which attacks man (Kettle, 1952).

The nuisance of this midge in Scotland prompted the Department for Health of Scotland to set up a Sub-Committee to begin investigations into the nature of the midge problem. This interest generated much field work and research studies based at the universities of Glasgow and Edinburgh during the 1940's. Studies were undertaken on flight activity (Kettle, 1950, 1951a, 1957 and 1960; Parker, 1949), egg and larval biology (Parker, 1950; Kettle, 1951b, and 1956; Hill, 1947; Reuben, 1959); and influence of weather conditions on the activity of adults (Kettle, 1957; Reuben, 1963).

Recent studies of seasonal incidence and flight activity of C. impunctatus in southern England have been undertaken by Boorman & Goddard (1970) and Service(1969a, 1969b, 1971). Blood digestion in relation to oviposition was studied by Sevice (1968).

Kettle (1952) reports that after concerted efforts at control by barrier spraying of vegetation with D.D.T., the populations of C. impunctatus biting man were unaffected. This was not the result of resistance to insecticide, but to unknown features of the behaviour pattern (Kettle, 1949).

C. pulicaris and C. punctatus have commonly been reported as a pest of man (Edwards, 1939; Service, 1969b;Tokunaga, 1941), but more often are found biting farm animals (Campbell et al., 1960; Nielsen, 1971).

In North America, the group includes a number of troublesome pest species. C. tristriatulus Hoffman is extremely annoying to man in southern Alaska (Jenkins, 1948; Wirth & Blanton, 1969; Sailer et al., 1956) whilst C. yukonensis Hoffman and C. canadensis Wirth & Blanton are frequently troublesome in the interior parts of Alaska and western Canada.

The rôle of the species group in disease transmission remains poorly known. Boorman (1974), working on Culicoides from Cyprus, urged that further work should be undertaken to determine the involvement of members of the pulicaris group in the transmission there of bluetonge virus of sheep and cattle. In the U.S.S.R., Gutsevich & Vigovskii (1960) succeeded in isolating one of the chorio-meningitis group of viruses from wild Culicoides pulicaris (presumably sensu lato), but the virus was of low pathogenicity

in laboratory animals.

Nielsen (1971) suggested that Culicoides are the possible vectors of Summer Mastitis in cows, and that large numbers of C. punctatus found feeding on the belly of the cow made this species a more likely candidate than the other species of Culicoides that feed on different parts of the cow's body.

Section 5. ENTOMOLOGICAL METHODS

5.1. GENERAL METHODS AND TECHNIQUES

The following account outlines general procedures and materials employed in this study. Any alternative methods are detailed in individual sections.

5.1.1. Specimens

Samples of specimens used in the study were of two sorts, according to the purpose for which they were used. Firstly, large homogenous samples for the study of variation of characters. This was the basis of the allometry study, and the details of collection and subsampling are given in Section 7.5.

The second category, for which the majority of specimens were used, was for the taxonomic aspect, and represented as many geographical localities and morphological forms as possible. A few specimens from the large homogenous samples were also included. Material was either from the collection of the British Museum (Natural History), or donated or lent by many specialists (names and institutions given in the acknowledgements). Since neither principal component nor canonical variate analysis can be carried out on a data matrix with missing values, imperfect specimens were rejected. This resulted in rejection of a large number of specimens, but ensured a higher standard of accuracy.

5.1.2. Preparation

All specimens were mounted in Berlese mounting medium on microscope slides (for details, see Lewis, 1973, p.173). Dried specimens first had their wings removed and immersed in glacial acetic acid. The remainder of the specimens were treated with cold potassium hydroxide until soft, and then put into glacial acetic acid with their respective wings. After five to ten minutes, the specimens were removed and transferred to Berlese medium. Fresh specimens, and those which had been preserved in alcohol, were mounted directly into Berlese medium. Some material received from other museums was mounted in Canada Balsam. There was no noticeable

difference in shrinkage between the two mountants, only in optical properties, which did not affect measurement.

Before mounting, specimens were dissected under a low power stereoscopic microscope, as is the usual technique for Culicoides. The head, wings, thorax and abdomen were separated and, where possible, mounted under separate 5mm. diameter coverslips, in a thin layer of medium. This minimised measurement errors resulting from parallex.

5.1.3. Making and Recording Measurements

Measurements were made with a 'Wild' micrometer eyepiece. When calibrated with a stage micrometer, this yielded measurements with an accuracy of up to 2µm (at x400). The measuring eyepiece was fitted to a Wild M11 microscope. To further minimise error, measurements were made at appropriate magnifications to ensure that the structure filled most of the field of view. Segments of the antennae and palps were measured at x400, wings at x40 and the remaining characters at x100. The majority of characters are defined in Section 3, describing the morphology of Culicoides, and wing pattern characters are described in Section 8.

All lengths refer to midline measurements. For each segment of the antennae, palps and legs, length included the basal process inserted into the apex of the previous segment (Fig. 21). Total length of antennae and palps was calculated as the sum of all segments, measured individually. This is more accurate than making a single, overall measurement, because it eliminates error incurred by 'telescoping' of segments. Segment width was taken at the widest point.

Wings were examined under dark field illumination to emphasise the pattern. Detailed drawings of wings (and other structures) were made with a Wild drawing tube.

Measurements and observations were made in a standard sequence and noted on specially prepared printed forms. All measurements were recorded as microscope eyepiece graticule divisions. A series of short programs were written for conversion to millimetre values, summarisation of the data, and the calculation of complex ratios. Unless stated otherwise, calculations were carried out on CDC Varian computer in the Biometrics & Computing Section of the BM(NH).

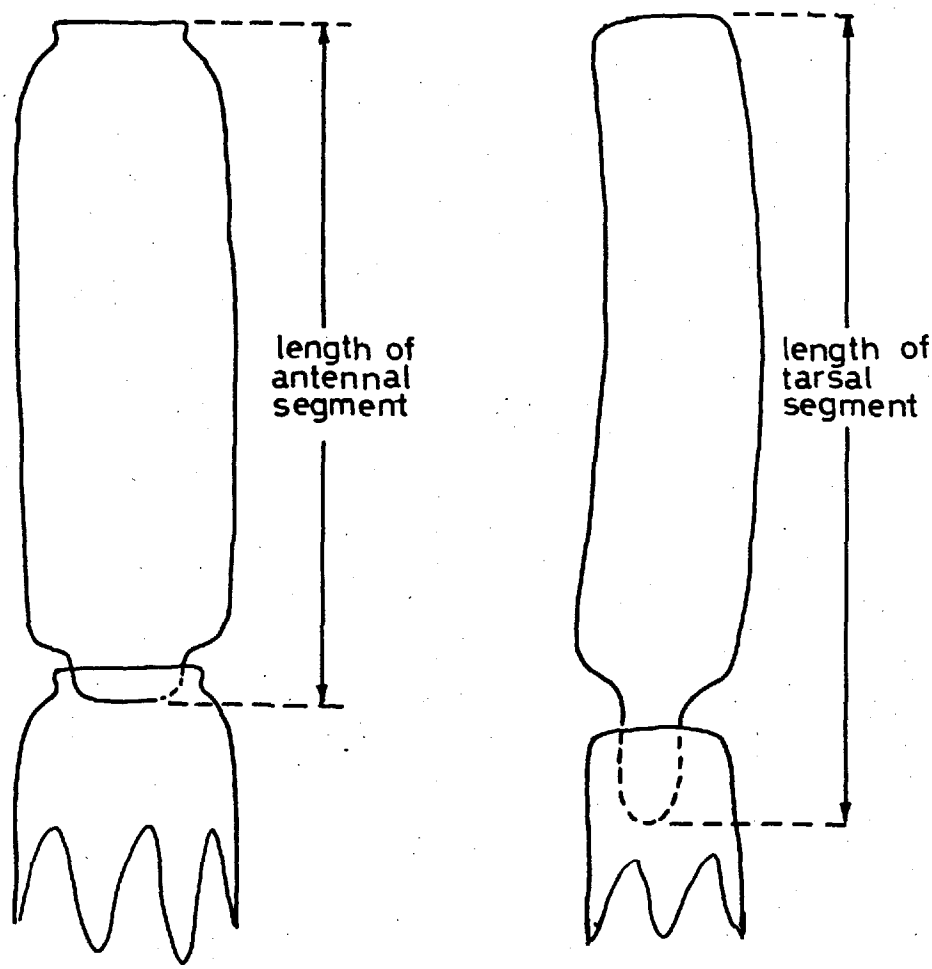


FIG. 21
MEASUREMENT OF SEGMENTS

5.2. RESTRICTION OF STUDY TO FEMALES

In deciding whether one or both sexes were to be used in the study, the following points were considered:

(i) Only females are attracted to hosts and light traps. They are also pests and vectors of parasites (or possible vectors), and consequently have frequently been collected and studied. Males, in contrast, are seldom collected and therefore only small numbers of specimens are available for study. Applied biologists and field workers require methods for accurate identification of females, in preference to males.

(ii) Because several members of the C. pulicaris complex are sympatric, the correct association of males and females presents considerable problems.

(iii) In morphological terms, females show greater interspecific variation than males. For this, and biological reasons (e.g., host selection), it seems likely that the adult female is the sex and phase of life cycle most susceptable, and responsive to, natural selection.

(iv) Principal component and canonical variate analyses do not allow any 'missing values' in the basic data matrix. Therefore, if males are unavailable for some samples, the females alone would be insufficient to characterise a sample, and it would have to be rejected.

(v) Male characters have so far proved useful for distinguishing only two species of the complex: C. orisescens and C. fagineus differ from all other species in the pulicaris complex by the presence of a convex ninth tergite.

After consideration of these points, it was decided to base this study on females alone.

Section 6. STATISTICAL METHODS

6.1. INTRODUCTION

Numerical methods employed here are typical of those commonly used in multivariate morphometric studies. They fall into three categories: clustering, ordination, and discrimination. Although the latter is essentially a special case of ordination, it is more convenient to discuss it separately. A number of other numerical methods, such as the allometry function, are employed, but these specialised methods will be discussed in the relevant sections that follow.

6.2. CHARACTER TYPES

All that is required at this point is a discussion of the type of characters used, as this has tended to govern the choice of statistical methods. The characters are described in the outline of Culicoides morphology (Section 3) and wing patterns (Section 8). A complete list of characters is given in Table 21.

Sneath & Sokal (1973) divided characters into three main groups, based on the type of coding and measures of similarity needed to employ them numerically:

(a) Two-state characters - presence or absence of a trait, also called binary or dichotomous characters.

(b) Quantitative multistate characters. These may simply be measurements or ratios varying in a continuous manner, or are characters represented by an ordered sequence of states, or are meristic characters (e.g., number of antennal segments). All of these characters may be expressed by a single numerical value.

(c) Qualitative multistate characters. The several coded states of these characters cannot be arranged in any logical order e.g., colour, or alternative forms of cuticular sculpturing.

All but one of the characters used in this study are of the type 'b'. The single exception is a two-state character - presence or absence of cibarial teeth.

Of the 71 quantitative multistate characters used in the multivariate analyses, 48 are simple measurements or ratios;

10 are meristic characters; and the remaining 13 describe wing pattern. For each of the wing pattern characters, it was possible to arrange the character states in a sequence of steps, of more or less equal magnitude. These characters could then be used for calculating measurements of association, in the same way as continuous and meristic characters.

6.3. PRIMARY DATA MATRICES

When a number of variates are measured on a set of OTU's (individuals, species, etc.), they may be tabulated to form a primary data matrix. If p variates are recorded on each of n OTU's, an n x p matrix is produced.

For the taxonomic part of this study (Sections 9,10), two primary data matrices were employed. The first was of the order 84 x 72 and used in Section 9 to determine (a) whether the recognised species were homogenous, or (b) whether a large number of characters were required for a reliable classification. The second matrix, of the order 145 x 10, was the basis of the discriminant analysis of Section 10. These two matrices are given in full in the appendix.

The data for the study of character variation (Section 7, p. 99) and wing pattern (Section 8 , p. 148) are described and summarised in the appropriate sections.

6.4. STANDARDISATION

The objective of standardisation is to compare characters expressed in different units of measurement. Because the characters used in this study are measured in a variety of units and scales, e.g., millimetres, microns, or pattern units, standardisation is essential. It usually involves subtracting from each observation the mean for the character, and dividing by the standard deviation. Each element of the standardised data matrix is therefore expressed in standard deviation units about a zero mean and is hence independent of the unit of measurement. Put another way, the standardisation of the character states makes all character means equal to zero and character variances equal to unity.

For most forms of factor analysis (of which principal components is an example), a variance-covariance matrix is computed, based

on standardised data. However, the correlation matrix is identical with a variance-covariance matrix based on data with unit variance and zero mean. This is because the correlation is the covariance divided by the square root of the product of two variances. When these variances have been standardised to 1, the denominator is 1, and therefore the correlation equals the covariance. Although the correlation matrix is the same as the covariance matrix, based on standardised data, the principal components (if R-mode) will differ according to whether the raw data or the standardised data were used to compute them. The alternative methods for calculating principal components are summarised in Fig. 22. Methods 1 and 3 provide identical results, which differ from the results of method 2. Throughout this study, method 1, i.e., correlation matrices, was used for the calculation of principal components.

6.5. ASSOCIATION MATRIX

The first major computational objective in a multivariate study is an association matrix (also called a similarity matrix). There are a number of ways in which the affinity of two OTU's may be measured, their use being governed by the form of the data (type of characters) and the statistical method used to display the relationships.

In this study, Gower's similarity coefficient and a correlation coefficient were used. The first coefficient was devised by Gower (1971) for application to all three types of character, and has proved very popular in numerical taxonomic studies. For two-state characters, the similarity is 1 for matches and zero for mismatches. For quantitative characters, the similarity between OTU's i and j is:-

$$S_{ij} = \sum_{k=1}^{p} \left[1 - \left(\left|x_{ik} - x_{jk}\right| / R_k\right)\right]$$

where X_{ik} and X_{jk} are the ranked or continuous measurements and R_k is the range of character k over all the taxa. The coefficient is 1 when both character states are identical and zero when the character states represent the extremes of variation for that character. Gower's coefficient also involved weighting coefficients

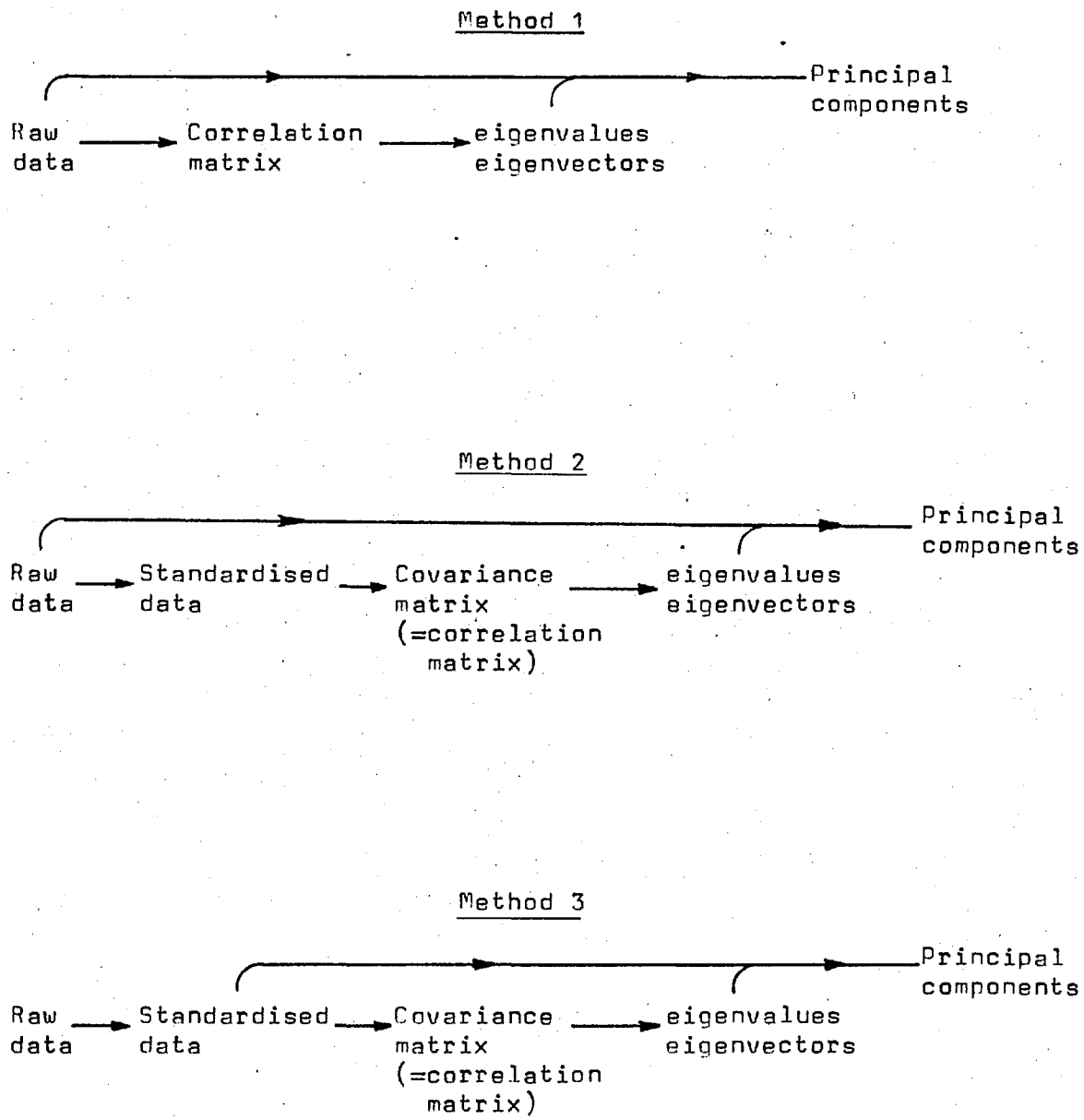


FIG. 22 Three Alternative Methods for Calculating Principal Components

which have been omitted because all characters are equally weighted.

A matrix of coefficients based on this formula formed the basis of the cluster analysis and, after conversion to a distance (see p. 95), principal coordinate analysis.

The second measure of similarity used between OTU's was the Pearson product-moment correlation coefficient. The similarity between two OTU's j and i over p characters is:-

$$r_{ij} = \frac{\sum_{k=1}^{p} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^{p} (x_{ik} - \bar{x}_i)^2 \sum_{k=1}^{p} (x_{jk} - \bar{x}_j)^2}}$$

Values of the correlation coefficient range from -1 (complete absence of co-variation) to +1 (complete co-variation). A matrix of correlation coefficients was used in principal component analysis.

There is an element in the n x n association matrix for every pair of OTU's compared, and the principal diagonal represents an individual compared with itself. In most commonly used data (other than immunological data), the resemblance of a to b is the same as b to a, so the upper triangle of the matrix is the same as the lower triangle, i.e., the association matrix is symmetric. As only one half is generally used in computations, and following the generally accepted procedure, the lower triangle is given in the present work.

The association matrix is generally quite large (each triangle has n(n-1) / 2 off-diagonal elements) and therefore it is necessary to summarise the information on relationships of OTU's implied by it so that the results can be easily comprehended and communicated. Since the results must be simplified, there will almost always be a loss of relevant information and some consequent distortion of the final presentation. The problem therefore, is to select a method which gives the proper balance between preservation of useful information (good fit) and simplicity (Rohlf, 1970). The four different methods described below are discussed in relation to these points.

6.6. CLUSTER ANALYSIS

This diverse subject has become very popular in recent years. Its main objectives are to group a number of objects and to display the relationships within and between the groups. Cluster analysis is usually presented in the form of branching diagrams called dendrograms. Because of its apparently easy interpretation, it is a frequently, and arguably overused, method. One of the basic assumptions is that there is some 'structure' or 'order' in the data and, with most clustering methods, that this structure is hierarchical. On first inspection this would seem most attractive to the taxonomist, but it does have disadvantages when used in circumstances where discrete groupings cannot be assumed. The taxonomic implications will be discussed further in the general discussion (Section 11).

The many types of cluster analysis are reviewed by Sneath & Sokal (1973), and others, who also give an extensive bibliography. Some of the underlying theories are formalised and discussed by Wolfe (1970) in terms of the analysis of multivariate mixtures. Unfortunately, there is no generally accepted classification, or even terminology, for clustering techniques. However, Williams (1971), in an excellent review, describes cluster anlaysis as a "strategy for classification" and gives an outline of the philosophy behind the various methods available, without recourse to complex symbolism. Some of the advances in methods which adapt to the type of variation in a cluster are given by Rohlf (1970). Different methods are suitable for data with known structure, but relatively little is known about the type of data structure for which each method is most suitable (Rohlf, 1963).

The most commonly used clustering techniques in biological classification are termed sequential, agglomerative, hierarchical non-overlapping (SAHN) techniques by Sneath & Sokal. They are typified by some important features which are of interest when specialised data are to be analysed. They begin with a number of discrete entities, which are then sequentially lumped into successively fewer sets, until all the entities are in one large set. An individual can belong only to one set at a time. At each stage of the lumping process, the admission of an individual to an established set, or the joining of two groups, will be decided by explicit rules. The nature of the decision rules will depend on

the exact technique used.

The clustering method used in this study was single linkage cluster analysis, and is one of the simplest available. In this method, an OTU which is a candidate for admission to an extant cluster, has a similarity to that cluster equal to its similarity with the closest member of the cluster. Therefore, connections between clusters and OTU's are established by single links between pairs of OTU's. This also means that two clusters may be joined by a single link between two of their members, even though some of the members of the cluster may not be very similar.

Basically, the algorithm searches the association matrix for those OTU's with the highest resemblence and joins them together. The level of admission to a cluster, or formation of a new one, is then lowered by a sequence of steps until all the OTU's are in a single set. One of the characteristic properties of single linkage methods is the production of a straggly dendrogram ("chaining").

Jardine & Sibson (1968) have argued on theoretical grounds that single linkage clustering is the best clustering strategy, because it conforms to certain postulates or axioms. Others point out that such axioms are too restrictive and thus exclude methods which nevertheless give a better description of 'natural' groupings. Farris (1969) for example, has shown that the unweighted pair group method of average linkage cluster analysis maximises the so-called cophenetic correlation coefficient, i.e., the correlation between lower triangular elements S_{ij} and the corresponding ultra-metric, derived from the dendrogram or linkage table. To date, there have been too few attempts at empirical evaluation of clustering techniques in animal taxonomy (or indeed, using artificial data sets) to confidently assess these opposing attitudes.

One generally accepted drawback of cluster analysis without character weighting is that, while it may depict relationships between individuals and small groups satisfactorily, it fails with relationships between progressively larger groups. Because the present data do not form compact and well separated groups, cluster analysis has not been used extensively in this study.

6.7. ORDINATION METHODS

The main objective of ordination is to describe a multivariate

sample in as small a number of dimensions as possible.

Two techniques were used, principal component analysis and principal coordinate analysis. Both are mathematically similar (Gower, 1966) and give approximately the same result, although they employ different methods to achieve this.

Ordination (mainly principal component analysis) is commonly applied to taxonomic problems, and in the present work, for the following purposes:

- Examination of correlations between different characters.
- Elimination of variables which contribute relatively little extra information to a classification.
- Examination of taxonomic groupings of individuals.
- Identification of individuals (or 'species') of doubtful or unknown identity.

No assumption need be made about the distribution of the variates in the hypothetical population, except where significant tests are of interest. Ordination gives a better representation of data where there is little tendency for the OTU's to occur in clusters, and therefore generally is more appropriate to the present problem of cluster analysis.

6.7.1. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis is described in detail by Seal (1964), Cooley & Lohnes (1962), Kendall (1975) and particularly clearly by Davis (1973), and Davies (1971).

The method may be described in geometrical terms by reference to Fig.23 . The method treats individuals as points in hyperspace, their position defined by the numerical values of all their measured variables. If a number of these points are plotted in a two dimensional space, they will form a cloud. The process seeks to find a set of new axes (Z_n) such that the first lies in the direction of the greatest variance of the cloud. The second axis will lie in the direction of the next largest variation and orthogonal to the first, i.e., uncorrelated and at right angles. These new axes, Z_1 and Z_2 are termed the principal component axes, and the coordinates on them of the points, are linear combinations of the original variables.

When computing principal components, the eigenvectors and

93

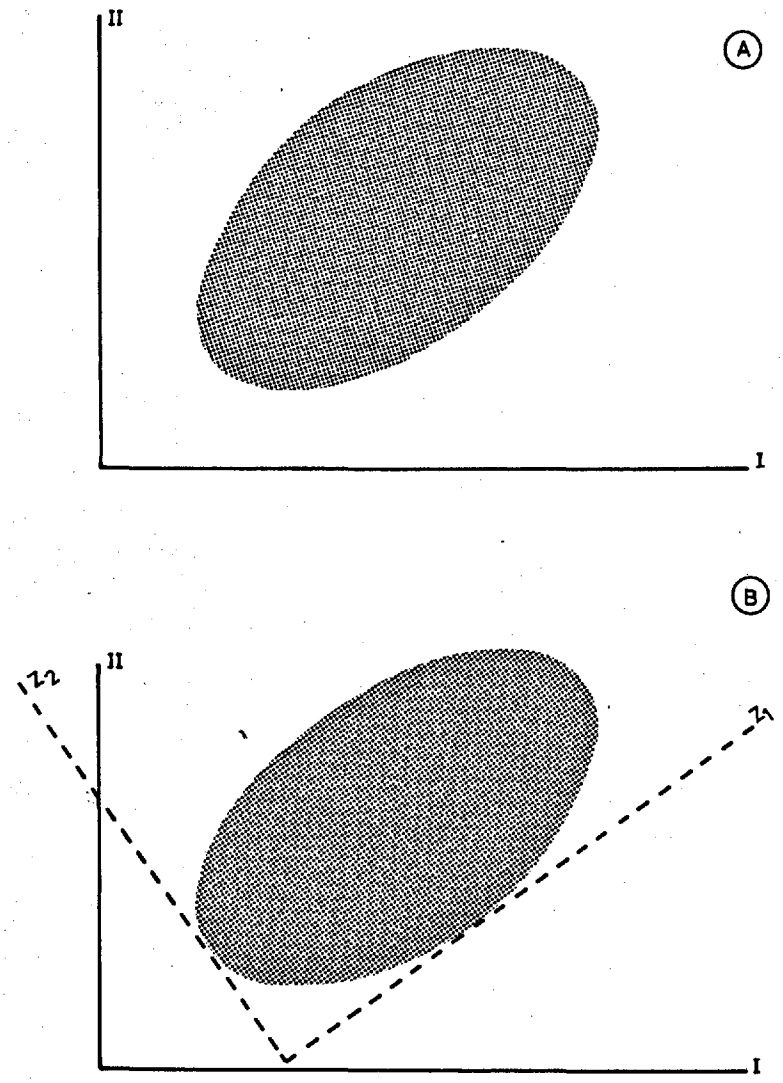


FIG. 23
TWO-DIMENSIONAL EXAMPLE TO SHOW GEOMETRICAL BASIS OF PRINCIPAL COMPONENT ANALYSIS

eigenvalues of a variance-covariance matrix (or correlation matrix in the present study, where standardisation was required) need to be found. The number of non-zero eigenvalues and associated vectors is the same as, or less than, the number of original variables.

Because the correlation matrix is always symmetric, eigenvalues are real and the principal components will be orthogonal, i.e., at right angles to each other.

Each of the eigenvectors was normalised so that the sum of squares of its elements was equal to 1.

An eigenvalue is equal to the variance along its corresponding axis, and therefore they may be ordered such that the first eigenvalue desribes the direction of the greatest variation. As the sum of the eigenvalues equals the trace of the matrix (a measure of the total variance within the matrix), it is possible to calculate the proportion of the total variance associated with each principal axis. As noted above, the positions or scores of each point on the principal axes are linear combinations of the original variables. The scores (i.e., principal components) are calculated by multiplying the value of the variable by the elements (termed loadings) of the corresponding eigenvector. These loadings are the coefficients of the linear equation which the eigenvector defines. The coordinate points of an OTU in the new space, defined by the principal axes, is found by calculating the scores for each axis.

The loadings are of considerable practical use in multivariate morphometrics, to determine the relative contribution each variate makes to the position of the individual in the new coordinate system. This aspect has been particularly useful in Section 9, to determine the relative importance of characters and which, if any, may be considered redundant.

The relative sizes of the eigenvalues gives an insight into the distribution of the OTU's in hyperspace. If the eigenvalues are similar in magnitude, it implies that the OTU's are distributed in a spherical fashion. If the first two eigenvalues are large and the remainder small, then the OTU's are concentrated in a plane. A large first eigenvalue followed by smaller and equal eigenvalues suggests a 'cylindrical' distribution. The nature of the OTU distribution can be most informative, helping to reveal some of the

factors likely to be influencing the analysis. It has been often found that a cylindrical distribution is frequently associated with size as the principal difference between OTU's. Therefore, not only does principal component analysis summarise the relationships between OTU's,but when used judiciously, it also allows considerable insight into the factors which shape these relationships.

6.7.2. PRINCIPAL COORDINATE ANALYSIS

Principal coordinate analysis was developed by Gower in 1966. One of its useful applications is that, unlike principal component analysis, it justifies the use of multistate or binary data. It is also said to place OTU's with less distortion than principal component analysis, when there are a number of missing entries in the matrix(Rohlf, 1972). Furthermore, it will be more efficient computationally than an R-mode principal component analysis, when there are more characters than OTU's. This is of particular importance when using a small computer.

The method of analysis begins with the construction of a similarity matrix, in this case using Gower's general similarity coefficient.The elements of the matrix are then converted to distances; $d_{ij} = 2(1 - S_{ij})$. The distance matrix is then transformed by subtracting from each element, e_{ij}, the mean of its row and column and then adding the mean of all elements in the matrix,i.e., $t_{ij} = e_{ij} - \bar{e}_i - \bar{e}_j + \bar{e}$.

The eigenvalues and eigenvectors of the transformed matrix are extracted and each eigenvector is normalised so that its sum of squares equals the corresponding eigenvalue. The resulting normalised eigenvectors give the OTU's on their principal axes (i.e., the method is a so-called Q-technique).

For multivariate morphometrics, an important feature of Q-techniques such as principal coordinates, is that the loadings, which are calculated directly in principal components (an R-technique), are not available. This is a disadvantage when the contribution of different characters to each of the successive eigenvectors is required. In the present study, principal coordinate analysis was used where eigenvectors were not required for detailed inspection.

6. 8. DISCRIMINATION

Of the discrimination techniques available, canonical variate analysis was chosen, because it enables individuals subsequently to be identified with relative ease. It enables an unknown individual to be assigned to a group with a known confidence, based on many characters taken simultaneously.

The calculation of canonical variates has some resemblance to that of components, and here again, transformed axes are produced. The main difference is that canonical variate analysis requires pre-established groups. Using group-means, it maximises the ratio of the variance between groups to the variance within groups. The weighting of variables is thus directed to those providing the best discrimination between taxa. The characters weighted by principal component and canonicalvariate analysis will therefore not necessarily be the same. The first axis produced is in the direction of the greatest variability between group-means. The second axis is inclined in the direction of the next greatest variability (but not necessarily orthogonal to the first), and so on for subsequent axes.

The geometrical basis of the analysis is shown in Fig. 24. Only two dimensions are figured, as was for the outline of principal component analysis. Referring to Fig. 24;

(i) The two-dimensional test space is represented by the axes X_1, X_2, of Fig. 24a. A principal component analysis is carried out 'within groups', effectively working on all points. The method assumes that the individuals within all groups have identical covariance matrices, i.e., the ellipses are of the same size and orientated in the same direction. At this stage, the variance of the groups and any correlations are ignored. A new set of axes Z_1, Z_2, are produced (Fig. 24b).

(ii) The ellipses of the individual samples are contracted to spheres, by setting the variance of all groups to 1 (Fig. 24c), to produce new axes W_1, W_2. W is related to Z by $W_1 = \frac{1}{\sqrt{\lambda_1}} Z_1$; $W_2 = \frac{1}{\sqrt{\lambda_2}} Z_2$ This scaling is non-orthogonal.

(iii) A second principal component analysis is then put on this W space, using only the group-means. The use of

97

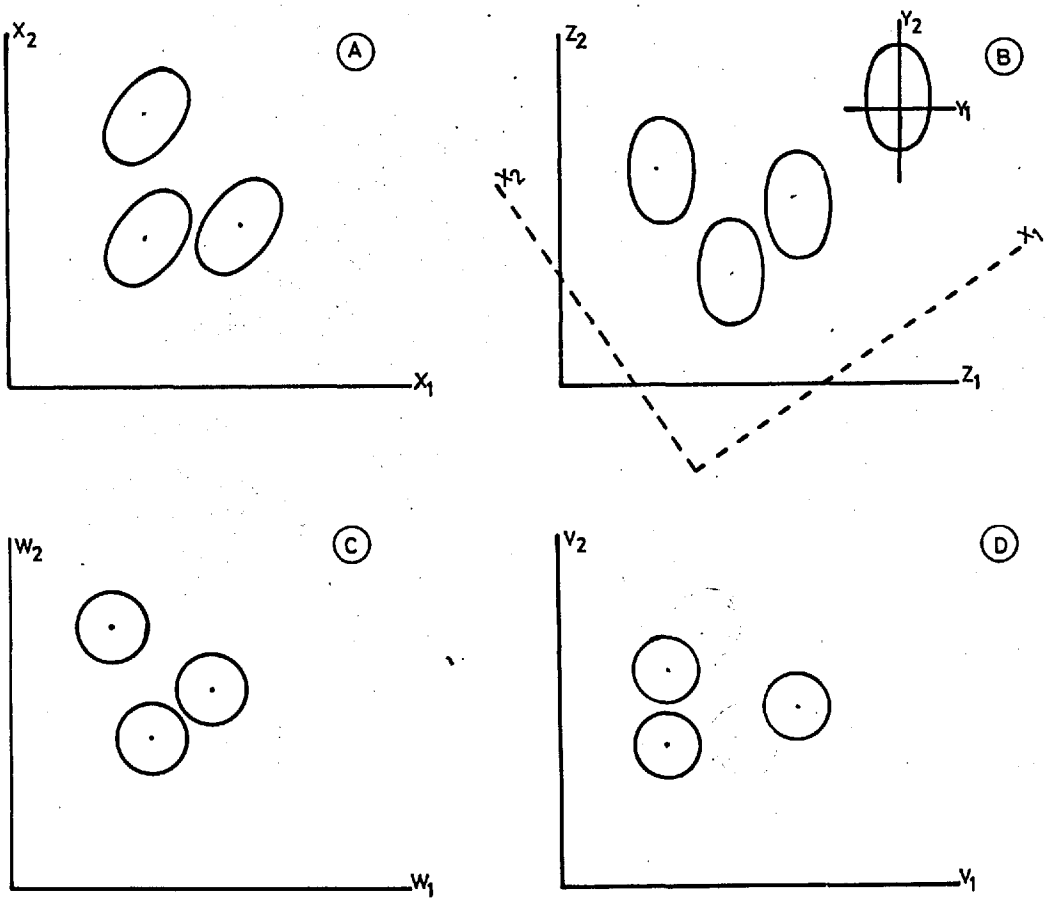


FIG. 24
GEOMETRICAL INTERPRETATION OF CANONICAL VARIATE ANALYSIS

scaling in step(ii) means that the distance between group-means is Euclidean and that there is no gross distortion due to uneven group sizes. This second principal component analysis produces another set of axes V_1, V_2, which are orthogonal to each other. These new axes are linear combinations of the variates given in terms of the W axes, which need simply to be converted into terms of the original X axes, so that the loadings on different variables can be investigated.

Blackith & Reyment (1971) discuss the often contradictory importance of statistical and biological significance of the results from canonical analysis. They advocate investigation of all axes, suggesting that some of the smaller axes may have biological significance. However, caution must be applied when using these lower axes for they may only reflect random effects in the data, and therefore be misleading. The confidence region around the mean of each sample may be calculated to allow an objective evaluation of the discrimination. In the two dimensions defined by the first and second canonical variates, the confidence region is a circle, and can be approximated to chi-squared distribution (χ^2), with 2 degrees of freedom. Thus, a 95% confidence limit for each group would be a circle of radius $r = \sqrt{\chi^2}_{(2;0.05)} = \sqrt{5.99} = 2.44$. To describe the 95% confidence region in three dimensions, a sphere of radius $\sqrt{7.81}$ (χ^2 with three degrees of freedom) is required.

In the section on discrimination (Section 10) these confidence limits are discussed in more detail.

Section 7. BIOMETRIC STUDY OF TAXONOMIC CHARACTERS AND THEIR VARIABILITY

7.1. REVIEW OF PREVIOUS WORK

Quantitative characters are important components of Culicoides taxonomy. This is readily demonstrated by the large proportion of taxonomic descriptions given to enumeration of these characters. Numerous measurements and ratios are calculated to define species. Furthermore, these ratios and measurements are an integral part of most taxonomic keys to species.

Ratios are frequently used to describe shape, e.g., palp ratio describes the degree of inflation of the third palp segment. They are also used for economy of description, e.g., the costal ratio 0.60 is briefer and more accurate than a phrase such as 'costa reaching beyond middle of the wing'. Ratios also describe a character in a way that enables it to be more accurately analysed for variation, or taxonomic reliability. The use of ratios in biology was condemned by Atchley et al. (1976) on statistical grounds for being misleading. This point of view has been strongly contested by Hills (1978) and Dodson (1978), who criticised this work on computer generated pseudo-random vectors as either not relevant to practical biology, or to be misleading itself of the unreliable properties of ratios. Both authors suggest that the benefits of ratios to biologists far outweigh the deliterious effects outlined by Atchley et al..

Although many authors give tables of quantitative characters, (Campbell & Pelham-Clinton, 1960; Wirth & Blanton, 1969) there have been relatively few rigorous studies of their variation.

Multivariate studies have been applied to the economically important C. variipennis (Coquillett) in North America by Wirth & Jones (1957) and McGuire & Wirth (1958). More recently, in a series of papers, Atchley (1970, 1971, 1973) has studied the small subgenus Selfia Khalaf in North America. The taxonomic separation of two species of Leptoconops Skuse was given by Atchley (1974).

Mcguire & Wirth used the technique of discriminant function analysis to distinguish five subspecies of C. variipennis, based on female morphology. Numerous frequency diagrams were produced

to evaluate characters of use in discriminate analysis. Four were chosen: palpal ratio, wing length, number of sensilla on antennal segments iv - vii, and number of teeth on the apex of the mandible. Calculations were initially made on samples containing specimens from scattered localities, but this gave poor results. Consequently, the analysis was rerun using samples restricted, so far as was possible, to one locality, and the localities were regrouped to represent five morehomogeneous subspecies. Throughout the study, it appears they assumed the infra-specific taxa were subspecies in the strict sense (i.e. geographically separated populations), although they gave no definition of their subspecies concept. The second analysis proved to be successful, allowing four of the five subspecies to be fairly accurately identified. The separation of the last two taxa was based on details of the male genitalia. The two authors finally produced a map of North America onto which were placed concentric contours around the postulated sites of 'typical' or 'genetically pure' populations.

The work of Mcguire & Wirth and Wirth & Jones was later criticised by Atchley (1967) and Hensleigh & Atchley (1977), who found that many of the divisions recognised by Wirth and his collaborators reflected environmental influences, rather than inheritable traits. It appears most likely that the shortcomings of Wirth & Jones' work lies in the use of the term subspecies. This is usually defined as 'an aggregate of local populations of species inhabiting a geographic subdivision of the range of species, and differing taxonomically from the other populations of the species' (Mayr, 1963). The subsequent findings of Atchley (1967) and others, of two or more subspecies coexisting in the same habitat, makes the validity of variipennis 'subspecies' rather dubious.

Hensleigh & Atchley (1977) undertook a detailed analysis of adult C. variipennis, reared from larvae in different conditions (using the laboratory colony of Jones) and suggested that the morphological divisions of the variipennis complex were "perhaps invalid" and that "true biological subgroupings can be described only as the result of more refined genetic analysis".

Much of Atchley's numerical analyses of Culicoides have been of the seven species in the subgenus Selfia Khalaf. He used the techniques of principal component analysis and discriminant functions (canonical variates) to investigate the geographical variation and

discrimination between three species. Morphologically intermediate specimens shown by the principal component analysis were interpreted as evidence for hybridisation. Numerous combinations of characters, presumably selected intuitively, were tried but it was found that using adult females alone, five of the seven species were considered inseparable. Despite the inseparability, a linear equation was given to describe the variation, but this resulted in a mis-identification of 25 - 30% of specimens. As is typical of multi-variate studies, some species had to be omitted from the analysis, because samples were too small.

In a later paper, Atchley (1971) advanced a number of hypotheses, couched in terms of Levins' theories of strategies and adaptation, to explain the different types of geographical variation in morphology, which he had observed.

7.2. SEASONAL VARIATION IN SIZE

To study seasonal variation in size, fairly large samples are required. Unfortunately, these were only available for the three common species C. pulicaris, punctatus and impunctatus. Samples were based on specimens collected in Britain and were divided into two groups, those from southern England and those from Scotland. This ensured that seasonal variation in size would not be confused with any geographical variation in size. To ensure that samples did not contain more than one species, only specimens which could be reliably identified (as 'typical forms') were used.

One of the most commonly used measurements of overall size in Culicoides is wing length. It is not an ideal measure because of possible proportional changes with increase in body size (i.e., allometric changes), but was considered adequate to demonstrate any overt seasonal variation in size (i.e., the variation in size for any given week of the year, due to environmental factors, was expected to be considerably greater than proportional changes in wing length with increase in body size). Wing length is the distance from the basal arculus to the wing tip.

Wing length was plotted against date of collection (divided into periods of one week) for the subsample of each species and linear regression carried out. The significance of the regression coefficient (i.e., whether the slope of the line is significantly different from zero) was also calculated, using a 't' test:

$t = \frac{b-0}{sd}(b)$ where b = regression coefficient and sd = standard deviation.

Results

In C. punctatus, the size of the midge clearly decreases during the summer (i.e., as temperature rises) in both southern England (Fig.25a)and Scotland (Fig.25b).Oecrease in size is more marked in southern England (b = 0.0597) than Scotland (b = 0.0122), where the midges decrease in size over a greater period of time(Table 2). In southern England there is evidence of a second, or autumn emergence of flies during September, which are approximately the same size as those from late June and early July. Size decreases linearly in relation to temperature during the early part of the year until a minimum wing size is reached. It then remains constant for the remainder of the year. The relationship between collection date and size is therefore not strictly linear when the whole year is considered.

An interesting departure from the general trend of the southern England data is shown by specimens from Lundy Island and the Scilly Isles. As west coast islands are usually warmer than the mainland, it would be expected that island specimens would be smaller than average for the time of year, not larger. Either, the remote and exposed nature of Lundy Island depresses the temperature significantly in some years, or other environmental variables may be responsible, as Lane (1978) suggests, when in relation to other Diptera.

C. pulicaris shows a similar trend in the seasonal variation in size to punctatus, with a more overt decrease in southern England than Scotland. For brevity, the pooled data for pulicaris is presented in Fig.28 and Table 2.

In contrast to punctatus and pulicaris, C. impunctatus does not show the expected decrease in body size during the warmer months. In both southern England (Fig.26a)and Scotland (Fig.26b) and Table 2, the calculated values of 'b' are not significantly different from zero.

One possible explanation for this is that impunctatus is very sensitive to small variations in temperature, and that collection date is a poor estimate of ambient temperature. Consequently, the data were replotted using the mean temperature for the locality and month of collection. Again, the regression coefficients were not significantly different from zero in both samples(Table 3).The pooled

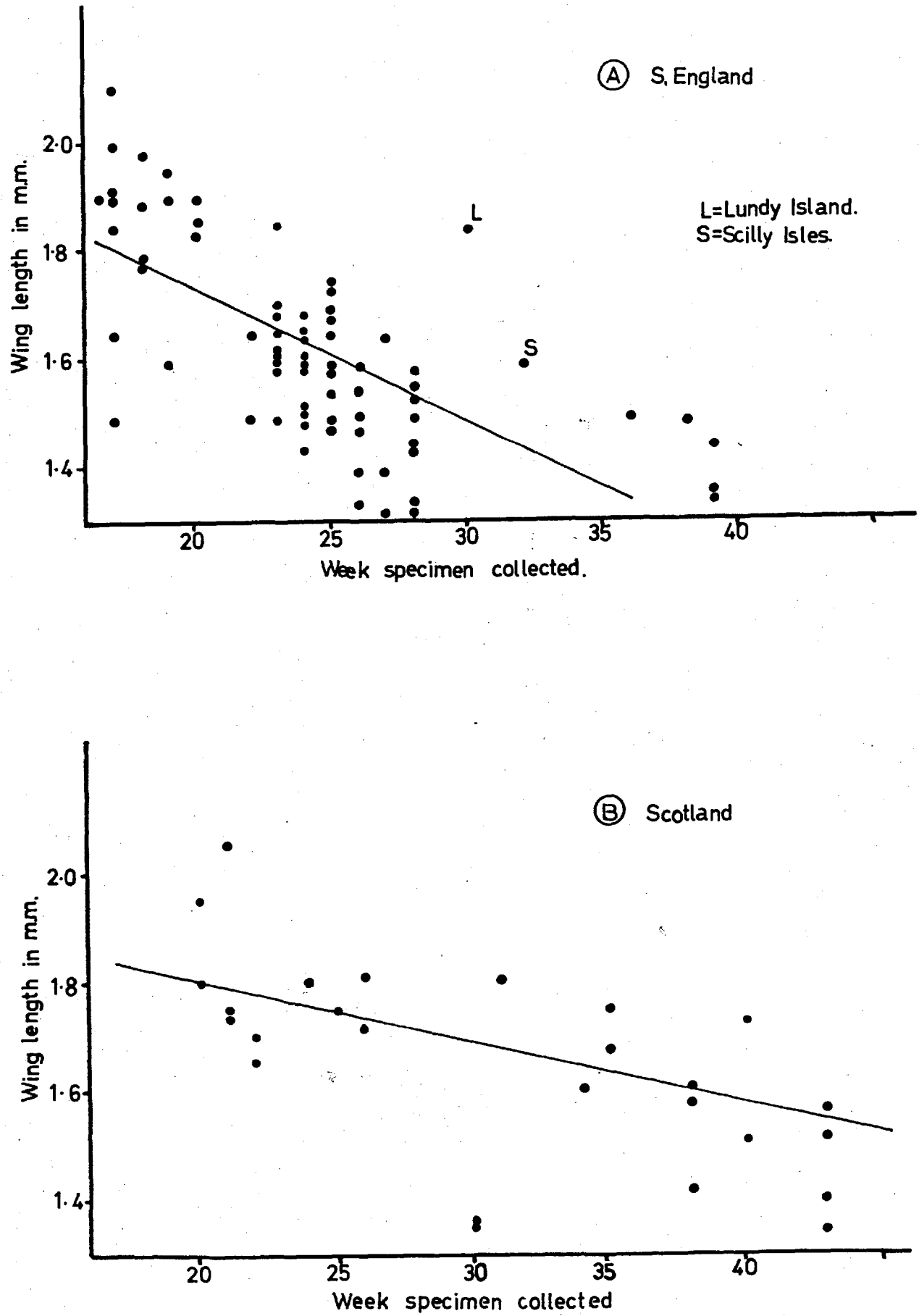


FIG. 25
INFLUENCE OF COLLECTION-DATE ON SIZE OF C. PUNCTATUS.

TABLE 2 SUMMARY OF REGRESSION OF WING LENGTH ON COLLECTION DATE

Species	Region	Regression coefficient 'b'	Standard deviation of 'b'	Significance of calculated 'b' from zero		
				value of 't'	degrees of freedom	significance
punctatus	Scotland	-0.01222	0.00479	2.550	20	$P > 0.01 < 0.02$
punctatus	S.England	-0.02436	0.00309	7.868	76	$P < 0.01$
pulicaris	Pooled	-0.02806	0.00567	4.953	40	$P < 0.01$
impunctatus	Scotland	-0.00615	0.02229	0.276	72	not sig.
impunctatus	S.England	-0.00321	0.00607	0.523	51	not sig.

TABLE 3 SUMMARY OF REGRESSION OF WING LENGTH, ON MEAN TEMPERATURE, FOR LOCALITY, FOR MONTH OF COLLECTION

impunctatus	Scotland	-0.00541	0.00323	1.673	72	not sig.
impunctatus	S.England	-0.00559	0.00431	1.296	51	not sig.
impunctatus	Pooled	-0.00262	0.00238	1.099	125	not sig.

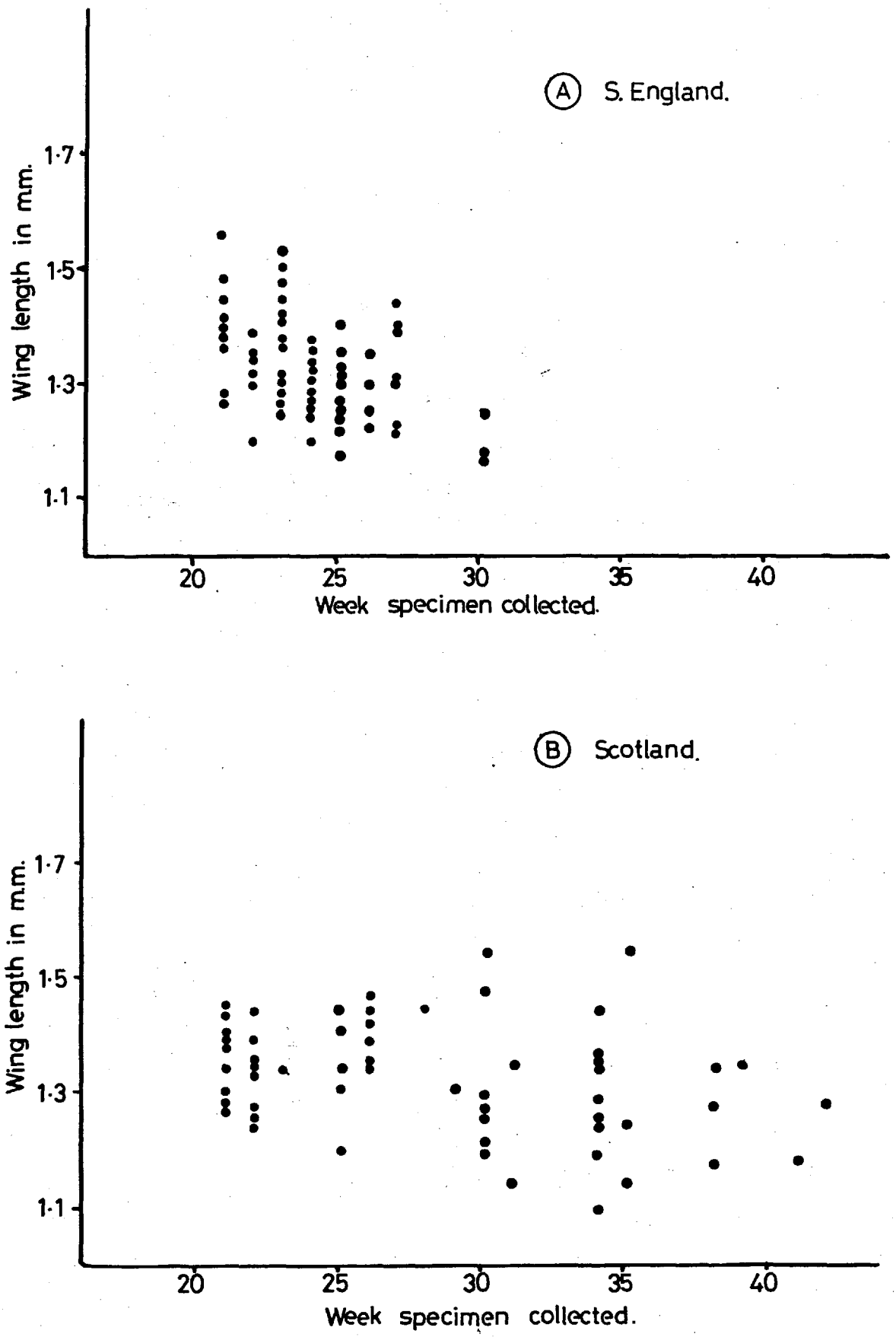


FIG. 26
INFLUENCE OF COLLECTION-DATE ON SIZE OF C. IMPUNCTATUS.

data are presented in Fig.27. Quite why impunctatus does not show the expected variation (seasonal) in size is difficult to understand. It may be related to the contradictory evidence in the literature concerning the number of generations per year.

C. impunctatus was reported as univoltine by Hill (1947) at Liverpool, Parker (1949) at Loch Lomond, Service (1969b)at Brownsea Island, Boorman (1970) at Woking, Onyiah (1971) at Ascot and Nielsen (1963) in Denmark. It was recorded as bivoltine by Kettle (1950) at Loch Lomond, Reuben (1963) at Lephinmore, and Onyiah (1971) at Ascot. Kettle (1950) monitored the abundance of midges at two sites on Loch Lomond, and suggested that the species was composed of two biological races. He rejected the hypothesis that impunctatus is basically univoltine, but may be induced to become bivoltine by climatic conditions. Kettle also suggests that the bimodal seasonal distribution shows the presence of two discrete races, distinct in three biological characteristics: seasonal abundance; sex ratios; and vertical distribution. However, Onyiah (1971) found that impunctatus was unimodal in three sites during 1969, but bimodal in the same sites in 1970, and attributes this difference to the development of immature stages under favourable climatic conditions.

If the coefficient of variation for wing length is calculated for each of the three species studied, it is clear that impunctatus is less variable (CV = 7.5%) than either pulicaris (CV = 11.1%) or punctatus (CV = 11.2%). It may be postulated therefore that because impunctatus larvae develop in peaty soil, which is poor in nutrients, the larvae may take different lengths of time to reach a minimum size, or nutritional state, prior to pupation. Hill (1947) showed that impunctatus overwintered as a fourth instar larva and that pupation took place between April and the middle of July. Time of pupation and subsequent emergence may therefore be controlled by nutrition as well as temperature, although the possibility of a long developmental period (11 - 12 months according to Hill) would lessen the effect of any temperature changes over a relatively short period during larval life. In the more variable species, pulicaris and punctatus (which breed in marshes and mud), nutrition may not be such a limiting factor, and the larvae can reach a nutritional threshold much more readily before pupation. Consequently their final size may be more dependent on temperature. This pattern

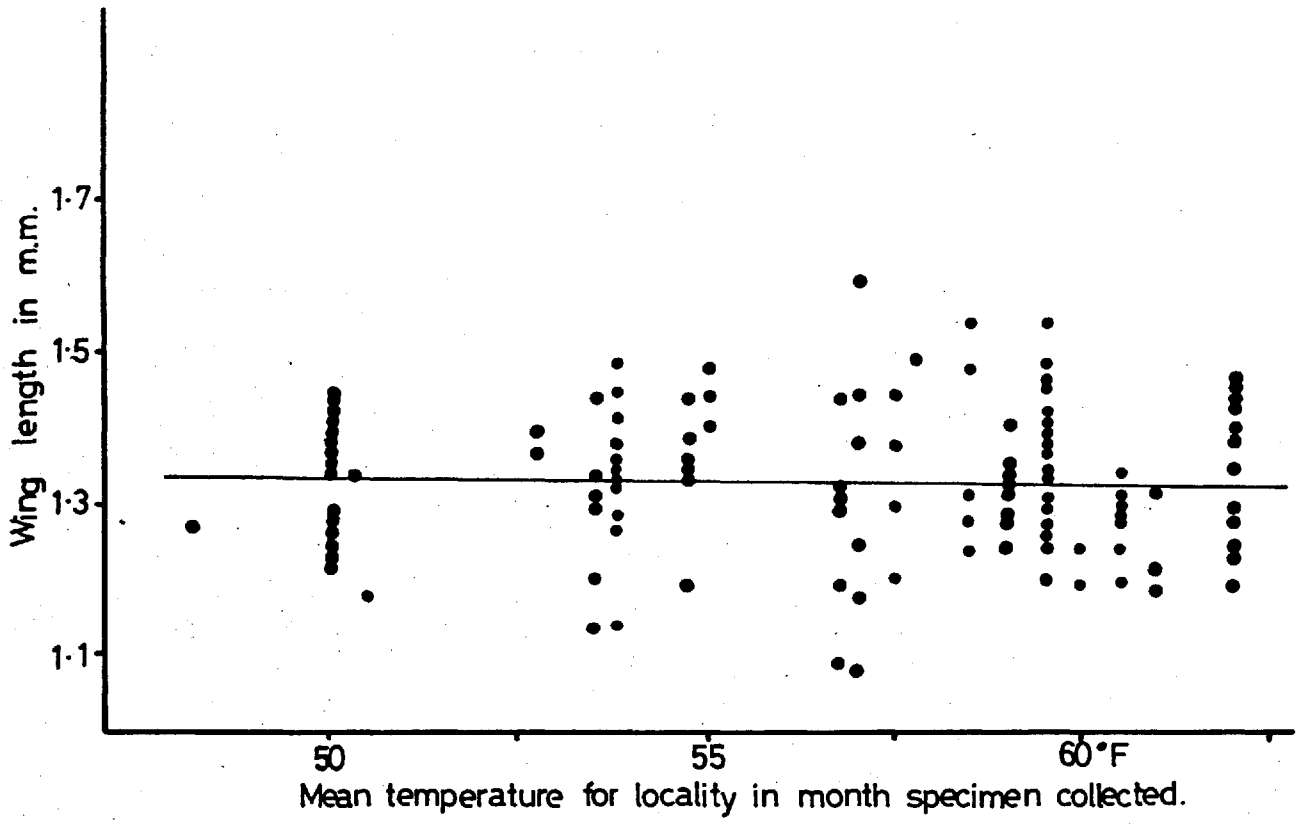


FIG. 27
INFLUENCE OF TEMPERATURE ON SIZE OF C. IMPUNCTATUS.

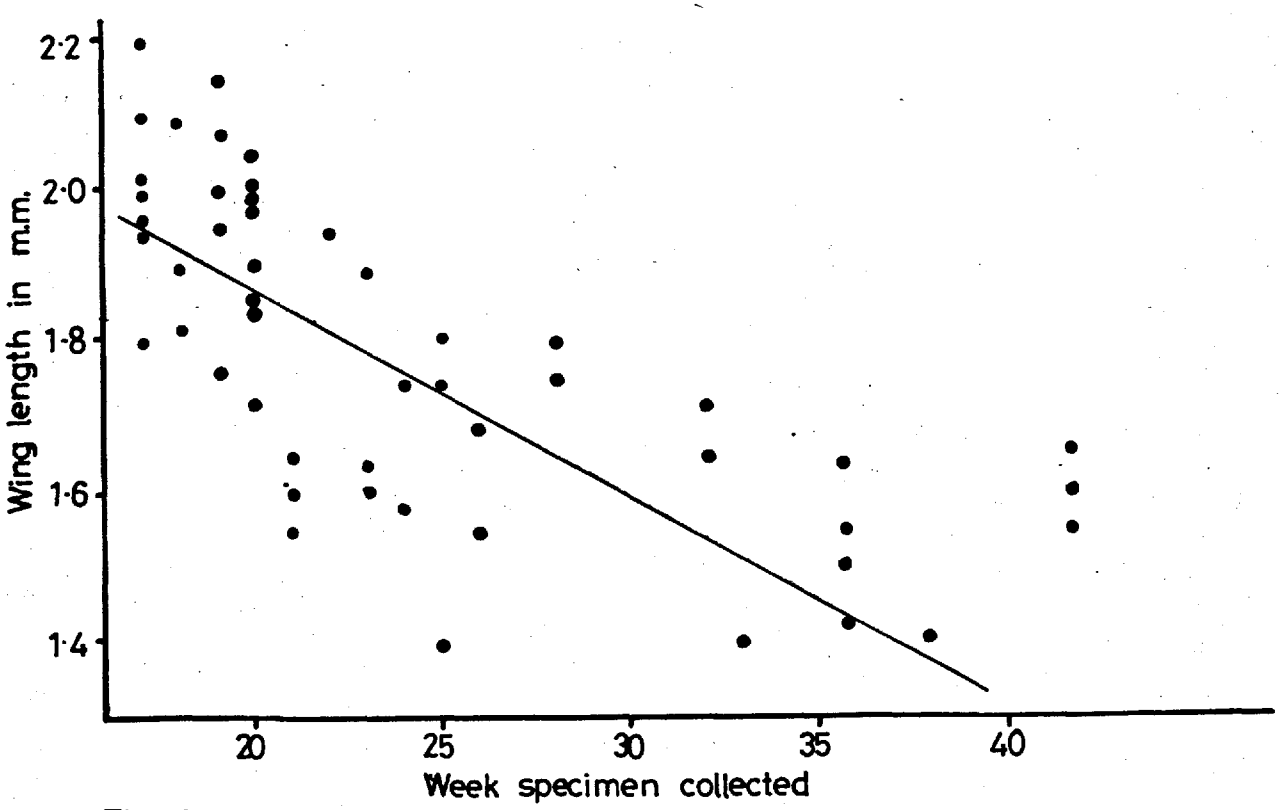


FIG. 28
INFLUENCE OF COLLECTION-DATE ON SIZE OF C. PULICARIS.

of variation is common in other Diptera, e.g., Calliphora and Musca, which are able to pupate at a wide range of nutritional states (as measured by size).

In a taxonomic context, the results indicate that the use of absolute size to discriminate between species is undesirable, because specimens of one species collected early in the year may be the same size as those of another species collected at a later date. (cf. punctatus Fig.25 with impunctatus Fig.26). It is possible that absolute size may be used to distinguish between species subjected to similar environmental variables, i.e., collected at the same locality and time of year, but this is not a common problem, as usually specimens are compared from distant localities. Under these conditions, size would be unreliable or even misleading.

7.3. GEOGRAPHICAL VARIATION

It was anticipated (in retrospect, rather ambitiously) that a geographical variation study would be made to determine whether any clinical effects existed in the species of the pulicaris complex. Since results of the previous section show that individuals (of some species) vary in size throughout the year, a comprehensive study has been rendered impractical with the material available.

For any such study of geographical variation in these insects, it is necessary to partition size variation (and other correlated characters) into a temporal and a spatial component. This requires that either (i) specimens are collected from a number of sites within a comparable period of time, or preferably, (ii) a series of specimens are collected throughout the year in a number of localities, and then the seasonal profiles of these localities are compared.

Any study that does not distinguish these components may lead to the demonstration of a false cline. An indication of the potential for this type of enquiry in the pulicaris complex may be demonstrated by comparing collections of C. punctatus made in Norway (Kautokieno) and England (Hampshire). Both samples were collected in mid-July (Table 4). The specimens from Hampshire have a mean antennal ratio of 1.07 and those of Norway have a mean of 1.12. Using a 't' test for small samples, these means are significantly different at a probability level of less than 0.01. The variances of the two samples do not differ significantly (F= 1.22).

Table 4. Comparison of two populations of C. punctatus

Variable	Locality	
	Norway (Kautokieno)	England (Hampshire)
Antennal ratio	$\bar{x}$ = 1.124 sd. = 0.042	$\bar{x}$ = 1.072 sd. = 0.038
Palp length µm	$\bar{x}$ = 198.6 sd. = 18.4	$\bar{x}$ = 244.5 sd. = 29.1

Sample size = 10 for both samples

If the size of individuals from the two populations is compared (using palp length as an estimate of size), they are found to be significantly different ($p<0.2>0.1$). If the data on palp length of Scottish specimens given by Campbell & Pelham-Clinton (1960) are used as criteria for identification, then the Norwegian specimens are more typical of newsteadi (= halophilus) than punctatus. This point underlines the disadvantages of identification based on size when specimens of different localities are compared.

The small size of the Norwegian specimens is interesting in connection with the preditions of Bergmann's rule. It suggests that the smaller individuals of a species are found in the warmer parts of its geographical range, and larger sized races in the cooler districts. This rule is usually applied when populations from various lattitudes are compared and suggests that the Norwegian speciemns would be larger than the Hampshire specimens, not smaller. Petersen (1952) found a similar situation in wing lengths of Scandinavian butterflies, where smaller individuals were collected from more northerly localities. However, Bergmann's rule relates temperature to body size and the apparent deviations from this rule may be that lattitude is a crude measure of ambient temperature. In a laboratory study, Ray (1960) investigated the application of Bergmann's rule to a number of poikilotherms and concluded that when temperature was carefully measured, body size followed Bergmann's rule.

These examples reveal that significant differences in Culicoides morphology do, in fact, occur between geographically separated localities. Furthermore, without providing the type of

detailed analysis outlined above, it is difficult to obtain both an accurate description,and the biological significance,of the geographical variation.

7.4. VARIATION IN SEGMENTS OF THE ANTENNAE

The antennae play an important rôle in the taxonomy of Culicoides, as they do in the life of the midge. The most significant taxonomic characters on the antennae are the distribution of sensilla, and quantitative differences in segment lengths.

The objective of this section is to investigate the nature of the variation in quantitative characters, i.e., differences in proportion of segments and ratios, and to suggest their possible functional significance.

The variation in the distribution of sensilla in a laboratory colony of C. nubeculosus is described by Kremer & Delécolle (1974).

7.4.1. Relative Lengths of Individual Segments

A conspicuous component of many species descriptions in Culicoides and other Ceratopogonidae is the detailed listing of the lengths of individual antennal segments. The measurements are usually given in arbitrary units since relative, rather than absolute lengths are regarded as significant. Their usage stems from Winnertz, 1852, (Campbell & Pelham-Clinton, 1960).

The observed variability in the individual segment lengths prompted a study to compare the degree of symmetry in right and left antennae and to determine whether measurement of only one antenna (right or left) is sufficient to describe a specimen accurately.

7.4.2. Sampling and Data

Because of the possibility that differences between right and left antennae might be very small, perfect specimens were required, in which the segments of neither antennae had been deformed during the process of slide mounting. Such stringent requirements of the material, together with the necessity for homogenous samples

(i.e., specimens from different localities could not be pooled), meant that samples (n = 20) were only available for the three common species - pulicaris, punctatus and impunctatus.

To reduce error to a minimum, measurements were made on segments of the right and left antennae at X400. The summed length of segments iii - x (proximal length), xi - xv (distal length), iii - xv (total length), and the antennal ratio (distal length divided by proximal length) were calculated for each antenna.

The coefficient of variation was calculated for each segment, to facilitate comparisons between the segments of the antennae of any one individual. This eliminates difference in size so that variation in short segments could be compared with the larger segments, or with total antennal length. The mean coefficient of variation for each antennal segment of a species was also calculated. The data for the three species are given in Tables 5 to 7, and histograms of the mean coefficient of variation for each segment in Fig. 29.

7.4.3. Discussion

Using the absolute lengths of segments, the coefficient of variation for any given segment varies from 0.3% to over 6.0%, in each of the three species. Fig.29 shows the mean coefficient of variation between the segments of the antennae, plotted as a histogram for each species. This facilitates a comparison between species and the different sections of the antennae. There is no general pattern in the variation between the segments of complementary antennae in any of the species, but each one shows a clear trend in the compound measurements - distal, proximal and total lengths. These three measurements exhibit a much smaller variability between the right and left antennae than any single segment. Total length shows the least variation of all. This difference between the variation of individual segments and the variation of compound measurements is very interesting, and requires further discussion.

The results would indicate that there is greater control of the various sections of the antennae (= compound measurements) than of individual segments, i.e., segments may vary in length within each section (proximal length for example), but some

TABLE 5

ASYMMETRY IN THE ANTENNAE OF C.PULICARIS. COEFFICIENTS OF VARIATION BETWEEN RIGHT AND LEFT ANTENNAE.

Segments	3	4	5	6	7	8	9	10	11	12	13	14	15	PL	DL	TL	AR
	0.00	0.00	1.83	0.00	1.79	1.83	3.72	0.00	1.29	0.00	3.34	1.11	0.00	1.10	0.66	0.88	0.43
	0.00	1.99	1.79	1.70	3.28	3.36	0.00	0.00	5.23	0.00	0.00	2.11	3.82	0.21	0.21	0.00	0.42
	1.42	1.93	1.83	0.00	1.74	1.79	1.79	3.62	1.27	0.00	2.17	0.00	0.00	0.20	0.43	0.10	0.64
	0.00	1.99	1.79	1.70	0.00	3.62	1.79	3.62	0.00	0.00	0.00	3.28	2.39	1.27	0.86	0.21	2.13
	0.00	0.00	1.70	0.00	1.58	0.00	1.66	0.00	2.43	1.20	3.28	2.08	0.79	1.04	0.60	0.82	0.44
	1.37	0.00	3.72	0.00	1.74	1.79	0.00	1.79	0.00	0.00	3.00	0.00	1.48	0.20	0.86	0.32	1.07
	3.01	0.00	1.88	1.74	1.62	0.00	1.74	1.74	1.16	0.00	1.08	1.91	1.55	0.40	0.21	0.10	0.62
	2.88	3.92	1.74	0.00	0.00	0.00	0.00	1.83	1.25	2.57	1.08	3.92	0.72	1.02	0.00	0.52	1.02
	2.94	0.00	0.00	0.00	1.88	1.99	1.99	0.00	4.28	0.00	2.48	5.94	0.00	0.00	0.71	0.35	0.71
	1.37	3.62	0.00	0.00	1.58	0.00	0.00	1.79	0.00	0.00	2.11	1.07	0.00	0.20	1.04	0.41	1.23
	0.00	3.92	0.00	0.00	1.79	1.83	1.93	3.44	4.11	1.27	0.00	3.39	1.70	0.69	0.66	0.68	0.02
	1.48	0.00	0.00	0.00	1.74	0.00	1.83	3.53	1.32	2.35	1.03	0.98	1.45	0.60	0.22	0.43	0.38
	4.37	4.04	1.99	3.62	0.00	1.79	1.70	1.79	1.25	2.70	3.34	0.00	0.73	0.62	1.10	0.85	0.48
	4.12	3.82	1.79	0.00	3.21	0.00	1.79	1.79	1.21	2.39	1.06	2.14	0.00	0.82	0.84	0.83	0.03
	1.55	0.00	1.93	3.62	0.00	1.79	1.79	1.79	2.66	1.27	0.00	1.05	1.47	0.00	0.22	0.10	0.22
	1.66	2.24	2.11	1.99	3.92	1.99	2.05	3.92	1.55	5.89	1.29	2.35	1.66	0.00	0.99	0.48	0.99
	0.00	0.00	7.64	1.79	1.74	1.79	0.00	0.00	0.00	0.00	1.18	1.09	0.79	0.21	1.12	0.66	0.90
	0.00	0.00	4.87	1.62	4.76	0.00	1.66	1.70	3.56	1.18	0.00	1.00	2.91	1.79	0.20	0.80	1.99
	0.00	2.05	3.62	1.70	0.00	1.70	1.70	1.66	2.48	0.00	1.08	1.99	2.24	1.63	0.21	0.73	1.84
	2.66	3.72	0.00	1.66	1.55	1.66	0.00	1.70	2.32	2.35	1.02	0.93	0.00	0.00	0.40	0.19	0.41
mean	1.44	1.66	2.01	1.05	1.70	1.34	1.35	1.78	1.87	1.16	1.43	1.82	1.18	0.60	0.58	0.47	0.80

PL = proximal length
DL = distal length
TL = total length
AR = antennal ratio

TABLE 6

ASYMMETRY IN THE ANTENNAE OF C.PUNCTATUS. COEFFICIENTS OF VARIATION BETWEEN RIGHT AND LEFT ANTENNAE.

Segments	3	4	5	6	7	8	9	10	11	12	13	14	15	PL	DL	TL	AR
	1.70	2.24	0.00	0.00	0.00	0.00	0.00	0.00	1.40	2.66	2.43	3.44	1.62	0.91	0.50	0.23	1.41
	3.14	2.17	1.93	1.79	1.79	1.83	0.00	0.00	1.32	1.40	1.25	2.35	2.45	0.92	0.69	0.80	0.23
	1.66	0.00	1.99	0.00	1.93	1.93	1.99	2.05	1.34	0.00	3.75	0.00	5.37	1.91	0.48	1.20	1.42
	1.66	2.24	2.11	2.11	0.00	0.00	2.11	2.05	1.45	1.37	1.25	0.00	0.00	0.23	0.00	0.12	0.23
	4.87	2.17	2.11	0.00	1.93	1.99	2.11	0.00	7.14	1.37	4.96	4.42	0.81	0.68	0.99	0.83	0.31
	1.62	0.00	0.00	0.00	2.11	2.11	2.17	0.00	1.45	2.94	4.96	1.09	0.75	1.13	0.51	0.84	0.62
	0.00	1.88	5.23	1.66	1.58	0.00	1.62	1.62	4.42	1.13	0.96	0.00	3.04	1.16	0.41	0.39	1.57
	1.42	2.05	1.93	0.00	1.70	3.62	3.62	0.00	1.29	1.25	1.03	1.03	1.45	0.82	0.44	0.21	1.26
	5.89	2.05	0.00	3.82	3.72	1.93	1.99	0.00	0.00	0.00	0.00	0.00	1.52	0.43	1.64	0.56	2.08
	1.62	2.31	0.00	2.05	2.05	0.00	2.17	2.24	1.48	4.28	1.29	1.23	0.81	0.23	0.00	0.12	0.23
	2.94	4.15	0.00	1.88	3.62	0.00	3.62	0.00	2.48	1.27	1.16	1.04	0.73	0.00	0.22	0.10	0.22
	1.66	0.00	0.00	2.11	2.17	0.00	2.17	0.00	6.44	0.00	1.29	2.48	0.93	2.03	0.00	1.03	2.03
	1.52	1.99	0.00	1.88	1.79	1.88	0.00	0.00	0.00	0.00	1.18	0.00	0.74	0.43	0.68	0.55	0.25
	1.55	4.15	0.00	0.00	1.99	2.11	4.28	4.28	4.37	1.40	1.27	0.00	1.62	0.71	0.74	0.00	1.45
	0.00	2.17	4.04	1.93	1.93	2.05	2.11	0.00	0.00	0.00	2.57	0.00	2.39	0.22	1.23	0.71	1.00
	1.55	0.00	6.14	0.00	0.00	0.00	2.11	2.11	2.77	0.00	3.75	2.39	1.55	1.13	0.97	0.11	2.11
	1.79	2.39	0.00	0.00	2.11	4.14	2.32	0.00	1.58	3.14	0.00	2.61	0.00	0.21	0.00	0.13	0.25
	1.58	2.17	0.00	1.93	3.82	0.00	0.00	2.11	0.00	4.11	2.57	3.68	1.81	0.00	0.49	0.24	0.49
	0.00	2.24	4.14	0.00	0.00	2.11	4.56	0.00	2.88	3.89	2.31	1.58	0.23	0.32	0.57	0.37	0.85
mean	1.81	1.82	1.49	1.06	1.71	1.30	1.94	0.82	2.44	1.46	1.98	1.40	1.46	0.67	0.52	0.43	0.87

PL = proximal length
DL = distal length
TL = total length
AR = antennal ratio

TABLE 7

ASYMMETRY IN THE ANTENNAE OF C.IMPUNCTATUS. COEFFICIENTS OF VARIATION BETWEEN RIGHT AND LEFT ANTENNAE.

Segments	3	4	5	6	7	8	9	10	11	12	13	14	15	PL	DL	TL	AR
	3.82	4.87	0.00	0.00	0.00	0.00	4.42	4.42	0.00	4.87	0.00	0.00	3.77	1.84	0.00	0.94	1.84
	1.79	7.71	2.48	4.87	0.00	0.00	0.00	2.48	1.94	1.66	1.45	1.37	1.01	0.28	1.17	0.72	0.89
	1.74	0.00	2.57	2.48	5.05	0.00	2.66	2.66	4.04	3.62	1.52	1.34	1.96	0.57	1.50	0.44	2.08
	1.74	0.00	2.39	2.39	2.39	2.57	0.00	2.57	1.88	1.79	2.94	0.00	0.00	0.57	0.58	0.58	0.02
	0.00	5.05	0.00	2.48	0.00	2.48	5.05	0.00	0.00	0.00	3.01	1.34	0.98	0.56	0.59	0.58	0.31
	3.82	2.57	0.00	0.00	2.48	0.00	2.57	5.05	1.88	3.53	0.00	4.04	0.00	1.10	1.49	0.14	2.59
	1.93	2.57	5.23	2.57	2.48	5.05	0.00	0.00	0.00	1.99	0.00	1.40	2.02	0.61	0.62	0.62	0.01
	0.00	2.05	0.00	0.00	2.17	2.31	2.24	2.31	1.79	4.87	0.00	2.61	1.93	0.00	0.79	0.40	0.79
	1.93	0.00	0.00	5.05	2.57	2.57	2.57	5.23	1.99	3.92	3.07	0.00	2.92	1.16	0.61	0.90	0.55
	1.79	2.39	0.00	0.00	0.00	2.39	0.00	0.00	1.93	1.74	1.40	0.00	2.96	1.11	0.28	0.42	1.40
	1.83	2.48	2.39	2.39	2.48	2.66	2.66	0.00	0.00	1.83	0.00	4.37	1.93	0.00	0.90	0.44	0.90
	1.93	0.00	2.57	0.00	0.00	0.00	2.66	2.77	2.31	4.04	0.00	3.07	3.14	1.26	0.62	0.31	1.89
	3.62	4.87	7.44	2.39	0.00	2.48	2.57	0.00	3.82	1.88	0.00	2.88	4.15	1.48	2.34	0.44	3.83
	0.00	4.87	4.71	4.71	2.24	0.00	2.39	2.48	1.79	3.28	0.00	2.48	3.49	1.83	0.28	0.81	2.12
	1.88	0.00	2.57	0.00	0.00	5.23	2.57	2.57	0.00	1.88	1.62	5.54	0.97	0.88	1.83	1.35	0.95
	3.62	0.00	2.48	0.00	0.00	0.00	2.48	2.57	1.79	1.79	1.52	1.42	0.94	0.28	0.29	0.29	0.14
	3.82	7.71	2.57	5.23	2.48	0.00	0.00	2.80	0.00	3.82	1.58	5.89	2.84	0.00	2.50	1.22	2.50
	1.99	0.00	0.00	2.48	5.05	0.00	2.66	0.00	0.00	0.00	1.58	2.94	1.93	0.91	1.57	0.31	2.48
	0.00	0.00	5.23	0.00	2.48	2.57	0.00	2.57	3.92	0.00	1.52	1.29	1.99	0.57	0.93	0.75	0.35
	0.00	2.24	0.00	2.24	2.31	0.00	2.48	2.48	3.62	1.62	2.62	2.43	1.74	1.28	0.28	0.53	1.56
mean	1.85	2.47	2.13	1.96	1.71	1.51	2.00	2.14	1.63	2.41	1.19	2.22	2.03	0.81	0.96	0.61	1.34

PL = proximal length
DL = distal length
TL = total length
AR = antennal ratio

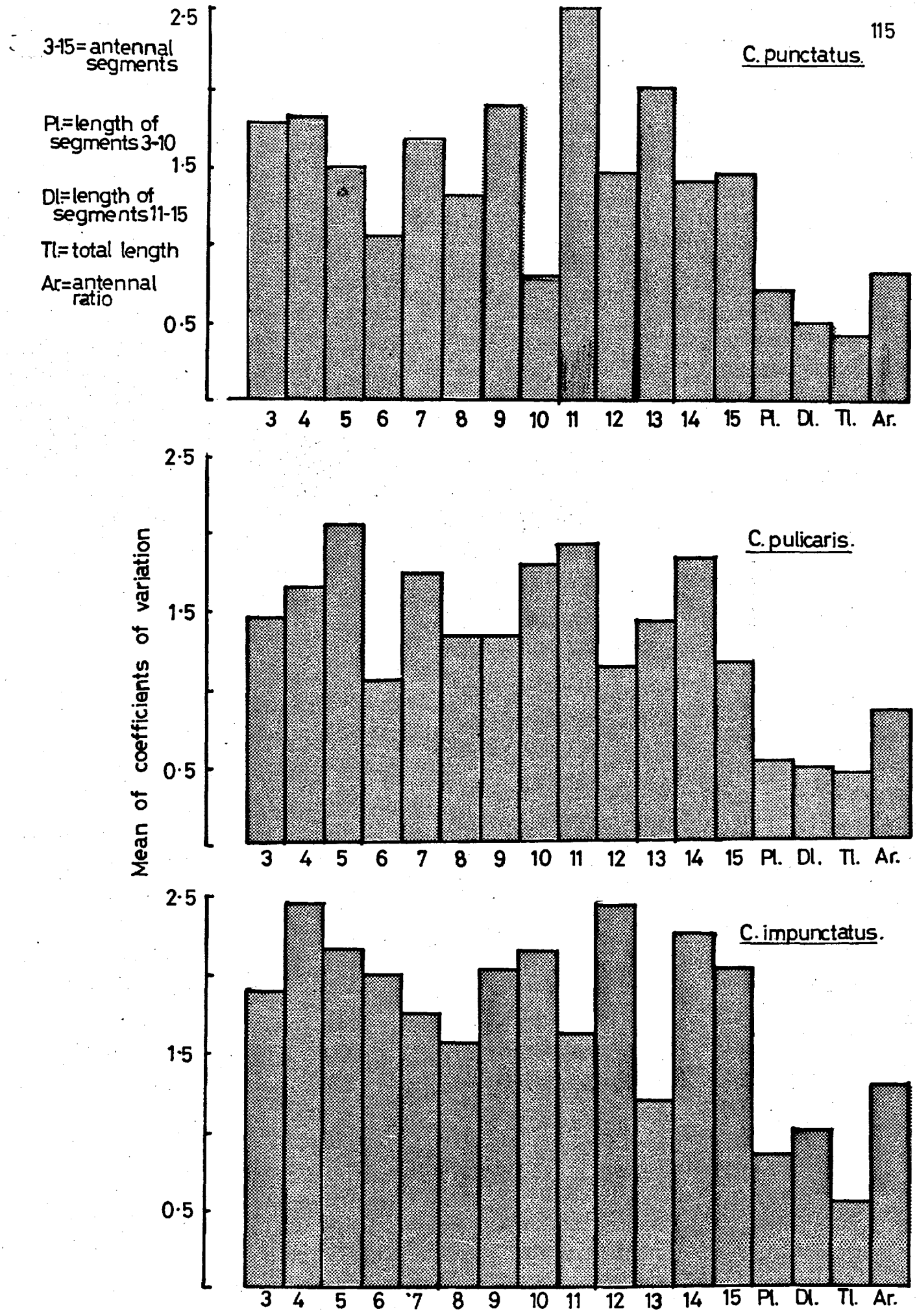


FIG. 29
VARIABILITY OF DIFFERENT SECTIONS OF THE ANTENNAE.

'homeostatic' mechanism ensures that an increase in the length of one segment is compensated by a decrease in another. This 'length homeostasis' ensures that asymmetry in the sections of the antennae is reduced to a minimum. An extreme example of this phenomenon was occasionally observed in C. newsteadi. Normally, in the pulicaris group the first segment of the flagellum (segment iii) is approximately 50% longer than the following few segments. However, in some newsteadi specimens, the segment iv of one antenna was the same size as segment iii, and therefore some 50% longer than segment iv of the complementary antenna. This large difference was compensated for, so that the proximal length of the two antennae differed very little. Two hypotheses may be advanced to explain this phenomenon. Firstly, that there is a functional basis to this, and secondly, that the observations reflect the development of the antennae in the pupa.

Although not actually demonstrated in male Culicoides, the plumed antennae of male Nematocera have clearly been shown to be important in swarming and in location of females for mating, particularly in mosquitoes (Roth (1948), Wishart & Riordan (1959), Autrum (1963) and Nijhout (1977)). There have been no equivalent studies on the function of antennae in females (which is the subject of the present study) and it is possible that the disparity between symmetry of segments and sections of the antennae may give some insight into function. Tischner (1953) developed an hypothesis for the action of Johnston's organ in mosquitoes, in which the organ responds to a narrow range of frequencies, the maximal sensitivity determined by the resonant frequency of the antennal flagellum. It is possible that the few hairs present on segments of the female antennae of Culicoides, although not as well developed as in the male, may be sufficient to receive some sound waves and generate a vibration along the main axis of the antennal shaft. If this is the case, then the accurate control of the proximal, distal, and total lengths between right and left antennae would be essential, to maintain the correct resonant frequency.

In the absence of any experimental evidence, the ability of hairs on female antennae to receive sound must remain in some doubt, but Ewing (1978) was able to show that the small and relatively hairless antennae of female Drosophila were able to respond to frequencies corresponding to the male wing beat. Furthermore, Roth (1948) showed that even after complete removal of the flagellar hairs, males still respond to sounds, apparently by vibrations of

the shaft alone. On complete removal of the flagellum, however, they gave no reaction at all. Behavioural observations in Culicoides may help to clarify this point. Downes (1955) observed and experimented with a number of Culicoides swarms, but only during the period after the female had entered the swarm, and not the location of swarms by females. He suggests "that from time to time, a female, either by accident or by some stimulus, flies into the swarm". Such a stimulus could possibly involve sound, although Downes suggests a more likely explanation of a female arriving by a visual response to a specific swarm maker.

The second, and more probable, hypothesis for the relatively low variation of antennal sections is that it may be a result of morphogenesis in the pupa. This offers a simpler explanation than the previous one. [Details of imaginal disc development in Drosophila are given by Schneiderman (1977)]. Although much is known of the embryological development of the mouthparts, little is known of antennal development in Nematocera. The difference in the variability of segments and sections may be explained in developmental terms, if morphogenesis is postulated to occur in two stages. First, the overall size of both left and right antennae is specified in the growth of the imaginal disc during early development, perhaps in response to some environmental variables (see section on wing size and temperature, p.102). Subsequently, probably during eversion of the pupa, when the tightly folded disc undergoes spectacular morphogenetic movement to form an extended adult structure, the relative lengths of the proximal and distal sections are expressed. This is most likely to be determined genetically. After this determination, the development of individual antennae is uncoupled and the secondary division into 'segments' takes place within the existing framework of sections. The general similarity of imaginal disc morphogenesis would suggest that similar homeostatic phenomena observed in the antennae also occur in the legs and palps.

Although the two hypotheses are discussed separately, they are not mutually exclusive. Morphogenetic control might be of selective advantage because of the functional significance of the antennae as sound receptors.

The phenomenon described here shows some superficial similarity with the 'oligomery' found in some Heteroptera, especially Lygaeidae, and Psocoptera (Burgess & Chetwyn, persoanl communication). However,

TABLE 8.

Asymmetry in Antennae of Culicoides pulicaris, punctatus and impunctatus, Using Proportional Lengths of Segments

Coefficients of Variation Between Right and Left Antennae

Segments	pulicaris	punctatus	impunctatus
3	1.64	1.68	1.65
4	1.77	1.77	2.71
5	2.09	1.60	2.19
6	1.37	1.21	1.99
7	1.64	1.63	1.92
8	1.47	1.34	1.72
9	1.25	2.02	2.00
10	1.89	1.18	2.10
11	1.82	2.53	1.83
12	1.23	1.71	2.33
13	1.30	1.82	1.37
14	1.66	1.62	2.35
15	1.34	1.37	1.95
PL	0.602	0.674	0.818
DL	0.581	0.528	0.964
TL	0.476	0.431	0.611
AR	0.802	0.876	1.343

PL = Proximal length (segments iii - x)

DL = Distal length (segments xi - xv)

TL = Total length

AR = Antennal ratio

the situation is different in as much as these insects are exopterygotes and the number of segments varies between antennae, usually as a result of damage during an earlier instar.

The results obtained from the use of proportional lengths of segments (Table 8) are very similar to those obtained from absolute length data. This is the result of the very low variability of total antennal length between right and left antennae. Hence, dividing each segment by total length (to give proportional length) is much the same as dividing by a constant.

In conclusion, the taxonomic implications of these observations are threefold:-

(i) Although there is some variation between segment length in complementary antennae, the difference expected between species is greater than between the antennae of the same individual. In this situation, the measurement of one antenna only from each specimen is thought to be adequate.

(ii) The use of segment lengths to distinguish closely related species should not be given undue emphasis, but may be used in conjunction with a number of other characters when inter-specific differences are large and significant.

(iii) The expression of segment lengths as a proportion of total antennal length is no more or less reliable than using absolute lengths. Therefore the use of proportional lengths has practical advantages, facilitating easier comparisons between different taxa.

7.5. ALLOMETRY OF SIZE

7.5.1. Background

Many taxonomists consider that body sizes of arthropods are not always taxonomically significant, since the size attained by an individual may be limited by environmental conditions. Greater importance is attached to differences in proportion, or to ratios expressing the 'shape' of a structure, supposedly independent of size (e.g., the palp ratio used in Culicoides taxonomy expresses the shape of the third palp segment by measuring its length relative to its width). More subtle estimates of proportional size are sometimes made by measuring the length of an appendage segment relative to another.

These quantitative characters are used extensively in the taxonomy of Culicoides, often occupying a large part of specific descriptions, and yet there have been few investigations of changes in the proportions, in relation to the size of the midges. Such a study is essential if quantitative characters are to be used as criteria for distinguishing species.

The mathematical models of allometry, developed for the study of growth and form, provide a useful tool for studying the problem of size and proportion.

7.5.2. Allometric Principles

The first quantitative analysis of differential growth concerned the change in size of individual organs (brain, heart, etc.) relative to overall body size in mammals. Early work was synthesised into a more generalised account by Huxley (1924) in which he related the size of an organ (y) to overall body size (x), by the simple function $y = x^{\alpha}$. This power function is now termed the law of simple allometry. In the function mentioned above, the allometric growth ratio, or equilibrium constant, α is the most important parameter to biologists, being a measurement of the growth ratio of y relative to x. It is a basic feature of this law that although two structures (x and y) may be increasing at different rates, the ratio (α) of these rates remains constant.

The allometric growth ratio is a pure number, having no dimensions and therefore may be compared directly from sample to sample (see discussion in Reeve & Huxley, 1945). The biological significance of the intercept on the Y axis, b, (termed the initial growth index) has been the subject of much debate, summarised in an excellent review by Gould (1966). Following his discussion of the interdependence of α and b, and the influence of measurement unit, no biological interpretation of b is advanced here.

Two structures are said to exhibit allometry (= heterogonic growth, disharmonic growth) if the growth of one structure is more or less rapid than that of the standard structure, and the exponent in the allometry function changes constantly in accordance with the law (Huxley, 1932). When growth of the two structures remains the same, their geometrical similarity is maintained with an increase in size and α is equal to unity. This is termed isometry (= isogonic growth, harmonic growth). If α is less than 1, then

the relationship between the two dimensions is termed negative allometry and if α is greater than 1, it is termed positive allometry. There is no biological difference, only a formal mathematical one, between positive and negative allometry. It is simply a matter of which quantity is taken as the dependent variable or the independent variable.

Although Teissier (1960) dismisses the use of elaborate physico-chemical explanations of allometry, he does advance a useful physiological context for allometry, as the unequal response of two organs to the same group of factors. McMahon (1973) discusses rates of physiological processes in terms of the allometry function.

In recent years there has been much interest in the use of multivariate statistics to analyse growth. There are clear reviews by Brown (1969) and Davies & Brown (1972), who render clarification of the rather confused terminology, and compare different published studies.

Since the development of the allometric function to describe growth, the concept has been applied to a wide range of problems. The evolution of relative growth in the Gerridae is discussed by Matsuda (1961) and for the Orthoptera, also by Matsuda (1963). Legay (1977) used the allometric relationships of egg length to width, to investigate the physical constraints of insect egg shape. Brown & Davies (1972) made an important study of growth and its taxonomic implications in cockroaches. Allometry has been used in a number of taxonomic investigations, notably Gould (1966, p. 610). Geographical variation of allometry is reviewed by Petersen (1952) and discussed by Thorpe (1976).

The wide range of studies in allometry led Teissier to distinguish two main categories:

(i) Allometry of growth - where the specimens compared belong to successive stages in a particular ontogeny.

(ii) Allometry of size - where the specimens compared are individuals of different size at some specific stage of development. This is termed static allometry by Gould (1966).

With reference to arthropods, Teissier states "the problem of allometry of growth is not formally different to allometry of size".

The present study concerns allometry of size because Culicoides are endopterygotes, showing profound morphological changes between immature stages and adults. Most allometry studies of insects are undertaken on exopterygotes, which show a gradual development from the immature to adult insect. Brown (1977) has investigated allometry

in Coccinellidae beetles, using homologous structures, in larva and adult. To summarise the expansion of the allometric concept to include size influence in single stages, the following definition is particularly useful: 'Allometry denotes the mathematical relationship between the size of a part and size of the whole (organ or organism) to which the part belongs'. This allows further distinction into 'growth' and 'size' allometry, and also permits extension of the concept to heterauxesis (i.e., comparison of parts to wholes in differently sized insects). It does not prejudge any issue connected with the functional, evolutionary, or taxonomic significance of allometry.

7.5.3. Allometry and Taxonomy

The rôle of allometry in taxonomy has been discussed by Simpson, Roe & Lewontin (1960) and more thoroughly by Gould (1966) who suggested that variation in the parameters α and b may be correlated with changing environments. He further suggests that it is possible to deduce the exact nature in which allometric patterns are determined, or phenotypically alterable, within the genetic system. Some features are constant when the environment is changed, others are labile. The significance of this point with reference to taxonomic characters is of the utmost importance when studying variable species, over a large geographic range.

Indeed, the failure to recognise the allometric consequences of overall size differences had led to unwarranted taxonomic distinctions between organisms, when the allometric nature of the distinguishing character is not clearly understood. Johnston (1939) found that the ratio of antennal segments was of no taxonomic value in discriminating between two species of Cimex and that allometric growth of segments during development affected this ratio. Later, in the same paper, he was able to show the allometric basis of a new and effective ratio. Further examples are furnished by Reid (1942) on Laemophloeus (Coleoptera: Cucujidae) and Boratynski (1952) on female Coccoidea.

It is possible to obtain taxonomic information by differences in α values, as shown in the present study. Differences in the values of b have been used in the past where α values do not differ, although the biological significance of this discrimination is rather dubious. Furthermore, owing to the mathematical inter-

dependence of α and b_b may only be used as a discrete character when the values of α are equal in the relationship considered.

7.5.4. Choice of a Reference Dimension

This choice is fundamental to the satisfactory interpretation of results. A convenient measure of body size, overall length for example, is commonly used and has obvious biological significance. However, the abdomen of Culicoides is only slightly sclerotised, and often distended, making accurate measurement almost impossible. The rounded nature of the head also adds problems to measuring total length. Brown & Davies (1972) used the total length of eleven accurately measured body sclerites as an approximation to body length. A single sclerite has been used by many authors as a reference dimension, but this has the disadvantage that the sclerite itself may vary in an allometric fashion. For this reason, the wing length, a commonly used measure of general body size in Culicoides, was not used. As flight in most flies is governed by the resonant frequency of the coupled thorax and wing, a size-dependent shape change in the wing is to be expected.

Multivariate statistical methods in the study of allometry has led to the use of a 'multivariate reference dimension'. The elements of the eigenvector associated with the largest eigenvalue of the variance-covariance matrix of logarithmically transformed characters provide, when suitably scaled, a set of coefficients (one for each character) that approximate to the α values. They are referred to as multivariate allometry exponents. The theory is given by Jolicoeur (1963). Brown & Davies (1972) showed empirically that there was congruence between multivariate values of α and those based on an overall body length reference dimension, in three species of Ectobius (Blattidae).

7.5.5. Fitting the Allometric Growth Function

Several methods have been used for estimating the parameters α and b, which is not surprising since this is a major concern in any allometric study. These are reviewed by Kidwell & Chase (1967) and Brown & Davies (1972).

The simplest method is to plot two variables, x and y, on

logarithmic scales and estimate the slope of the line (α), and its intercept on the Y axis (b). Williams (1972) describes a graphical method for demonstrating the presence of allometry, but it does not allow estimation of α and b. For any statistical validity, a more sophisticated method is required.

Calculating the regression Y = log y on X = log x by the method of least squares has commonly been applied, although this method assumes x is measured without error, and that the error in y is normally distributed with constant variance. The shortcomings of this method were outlined by Kermack & Haldane (1950) who proposed a new technique - the reduced major axis method - in which the products of deviations of x and y are minimised for each point on the line. This has the advantage that neither variable is treated as dependent on the other; the method is not affected by change in scale between variables, and both slope and y intercept can be calculated efficiently. Brown & Davies compared the method of least squares to a theoretically preferable method of Bartlett (1949) (as outlined by Simpson, Roe & Lewontin, 1960), and found that for their data, there was little difference in the estimates of α and b. This is not generally the case as Kermack & Haldane (1950) have shown; the least squares method and their reduced major axis method (very similar to that of Bartlett) gave appreciable differences in estimates of α and b.

The apparent disparity in the findings of the two studies may be attributable to differences in their primary data. Bartlett's method works most effectively on data from a single eliptical cluster of points. This is often encountered in allometry of size studies where a single stage is measured and the variables used are usually highly correlated. The high correlation that exists between individual segment lengths and the overall appendage length, after logarithmic transformation, will produce estimates of α and b that differ little, when calculated by the least squares method or Barlett's method.

7.5.6. Sampling Data

Sample homogeneity is essential in any taxonomic work, especially in this study, where many subspecific forms are suspected, to guard against samples being made up of two or more populations, the combined variation of which would mask the separate variation

of each population. Simpson et al. (1960) gave five categories of homogeneity that should ideally be met when sampling: locality, environment, time, age, and sex. Samples fulfilling all the criteria of 'biological homogeneity' and of sufficient size for statistical analysis, were only available for three species - C. pulicaris, punctatus and impunctatus. For C. pulicaris, the sample was taken from a single night's light trap catch. The sample of C. punctatus was also taken from a light trap catch, but on a separate occasion. For C. impunctatus, the sample was taken over a period of 30 minutes, from flies biting a human face. All specimens were preserved immediately in 70% alcohol. Random subsamples of 40 individuals of each species were then slide mounted in Berlese mounting medium.

The possibility of obtaining more uniform samples by laboratory rearing of progeny from wild caught females of known, or presumed, identity was not adopted for the following reasons:

(i) Culicoides are not easy to rear in the laboratory.

(ii) Insectory-reared specimens might not be morphologically analogous to wild specimens.

(iii) Although the morphological relationships between members of the pulicaris complex may be possible to evaluate, on material bred in the laboratory, the results would still require comparison with wild material before they could be employed for practical identification purposes.

The antennae, maxillary palps and hind leg were studied as they possess important taxonomic characters. The criteria for measuring individual segments are described in Section 5.

A summary of the data is given in Tables 9 to 11 and the estimated values of the allometric parameters are given in Table 12, for punctatus, pulicaris and impunctatus. Before discussing these results, it is desirable to examine the accuracy of the estimated values of α and b.

7.5.7. Significance Tests

Brown & Davies (1972) report that confidence limits are rarely given in allometry studies and that without them, the results are often impossible to assess. Consequently, in the present study, 95% confidence limits and significance tests were calculated as part of

TABLE 9.

SUMMARY OF OATA FOR ALLOMETRIC STUOY - CULICOIDES PULICARIS

	segment	mean	standard error	range
Antennae	3	57.24	0.579	48.72 - 64.96
	4	41.90	0.335	35.96 - 46.40
	5	44.71	0.425	38.28 - 52.20
	6	46.80	0.377	40.60 - 51.04
	7	48.37	0.470	40.60 - 53.36
	8	46.13	0.381	40.60 - 49.88
	9	46.19	0.458	39.44 - 51.04
	10	46.05	0.376	40.60 - 49.88
	11	64.40	0.717	52.21 - 71.92
	12	65.13	0.638	53.36 - 70.76
	13	74.53	0.834	62.64 - 83.52
	14	78.82	0.850	66.12 - 88.15
	15	107.15	0.966	95.11 -114.83
Maxillary palp	1+2	105.84	1.008	91.64 - 114.82
	3	83.63	1.078	70.76 - 98.60
	4	37.23	0.600	26.68 - 46.41
	5	39.78	0.492	34.80 - 46.40
Hind Leg	Femur	605.62	5.837	503.49 - 664.99
	Tibia	618.09	5.999	512.99 - 688.74
	Tarsus 1	323.58	4.211	251.74 - 370.47
	2	183.34	1.989	151.99 - 204.24
	3	103.55	1.286	85.50 - 118.75
	4	57.59	0.568	47.50 - 66.50
	5	58.90	0.717	47.50 - 66.50

TABLE 10.

SUMMARY OF DATA FOR ALLOMETRIC STUDY - CULICOIDES PUNCTATUS

	segment	mean	standard error	range
Antennae	3	51.92	0.512	45.24 - 58.00
	4	38.00	0.374	33.64 - 44.08
	5	40.08	0.458	34.80 - 48.72
	6	41.85	0.469	37.12 - 49.88
	7	42.79	0.575	37.12 - 52.20
	8	41.36	0.525	35.96 - 49.88
	9	40.44	0.620	34.80 - 51.04
	10	40.41	0.589	35.96 - 51.04
	11	60.13	1.156	51.04 - 90.48
	12	60.83	0.759	51.04 - 73.08
	13	67.12	0.969	60.32 - 85.83
	14	71.37	0.921	61.48 - 83.52
	15	101.83	1.151	87.00 -113.67
Maxillary palp	1+2	99.56	1.542	75.40 - 114.8
	3	79.18	1.224	61.48 - 92.80
	4	33.01	0.697	25.52 - 42.92
	5	38.71	0.634	31.32 - 46.40
Hind Leg	Femur	536.61	6.206	451.24 - 631.74
	Tibia	550.33	6.755	465.49 - 655.49
	Tarsus 1	273.12	4.403	218.49 - 332.50
	2	162.29	2.652	113.99 - 194.74
	3	91.56	1.485	66.50 - 109.25
	4	51.98	0.680	42.75 - 61.75
	5	54.75	0.551	47.50 - 61.75

TABLE 11.

SUMMARY OF DATA FOR ALLOMETRIC STUDY - CULICOIDES IMPUNCTATUS

	segment	mean	standard error	range
Antennae	3	44.95	0.433	40.60 - 49.88
	4	32.88	0.372	30.16 - 40.60
	5	33.20	0.351	29.00 - 38.28
	6	33.98	0.285	30.16 - 38.28
	7	34.10	0.274	31.32 - 38.28
	8	32.82	0.242	30.16 - 35.96
	9	32.42	0.307	30.16 - 38.28
	10	31.98	0.313	29.00 - 38.28
	11	42.25	0.523	34.80 - 48.72
	12	45.73	0.547	39.44 - 52.20
	13	54.46	0.508	48.72 - 63.80
	14	60.26	0.639	52.20 - 68.44
	15	84.79	0.725	76.56 - 96.27
Maxillary Palps	1+2	68.44	0.708	58.00 - 75.41
	3	58.36	0.770	49.88 - 68.44
	4	31.86	0.538	26.68 - 44.08
	5	33.18	0.549	25.52 - 39.44
Hind Leg	Femur	442.12	4.161	384.74 - 489.24
	Tibia	463.12	4.076	403.74 - 498.74
	Tarsus 1	218.62	2.439	189.99 - 251.74
	2	129.49	1.527	104.50 - 147.24
	3	77.62	1.002	71.25 - 90.25
	4	50.75	0.696	42.75 - 57.00
	5	52.25	0.674	42.75 - 57.00

TABLE 12.

ESTIMATED VALUES OF ALLOMETRIC PARAMETERS SHOWN IN CULICOIDES PULICARIS, PUNCTATUS AND IMPUNCTATUS.

	segment	pulicaris		punctatus		impunctatus	
		∝	b	∝	b	∝	b
Antennae	3	1.055	0.052	0.830	0.220	0.854	0.197
	4	0.601	0.728	0.830	0.162	0.976	0.067
	5	0.959	0.076	0.977	0.067	0.978	0.068
	6	0.943	0.088	0.996	0.061	0.888	0.121
	7	1.235	0.014	1.161	0.022	0.884	0.124
	8	0.912	0.016	1.098	0.032	0.873	0.128
	9	1.126	0.026	1.316	0.008	1.190	0.018
	10	0.674	0.499	1.160	0.021	0.967	0.069
	11	1.280	0.014	0.817	0.276	1.382	0.007
	12	1.079	0.051	0.979	0.100	1.427	0.006
	13	1.243	0.020	1.004	0.094	1.065	0.064
	14	1.081	0.061	1.080	0.061	1.136	0.046
	15	0.760	0.666	0.965	0.182	0.592	1.879
Maxillary palp	1+2	0.962	0.488	0.951	0.517	0.805	0.968
	3	1.229	0.090	1.108	0.177	1.162	0.133
	4	1.135	0.067	1.239	0.036	0.811	0.431
	5	0.507	2.177	0.717	0.706	1.342	0.030
Hind leg	Femur	0.960	0.394	0.888	0.602	0.995	0.317
	Tibia	1.003	0.311	1.009	0.304	0.948	0.437
	Tarsus 1	1.200	0.050	1.254	0.034	1.040	0.121
	2	1.040	0.074	1.323	0.014	1.122	0.045
	3	1.049	0.039	1.169	0.020	1.049	0.041
	4	0.591	0.345	0.535	0.466	0.975	0.041
	5	0.579	0.380	0.042	8.966	1.025	0.031

an allometric growth program written by R. G. Davies (Imperial College), based on Bartlett (1949). Significance tests were carried out to:

(i) Test whether α differed significantly from zero.

(ii) Test whether α differed significantly from unity.i.e., size induced variation of x and y are geometrically similar.

(iii) Test whether there are significant deviations of x and y from linearity, i.e., whether a higher order equation than $y = x^{\alpha}$ is required to explain the size induced variation in shape. This test may only be carried out when using Bartlett's method and not (with these data) for the method of least squares.

I Significance of the Slope of the Line

In the equation for simple allometry, if $\alpha = 0$, then x and y are unrelated. The test is carried out by calculating the statistic described by Simpson et al. (1960, p. 233-237), which has Student's 't' distribution, with n-3 degrees of freedom. As the present study is being made on size variation within a single developmental stage, it is to be expected that the total length of an appendage is highly correlated with each component segment. This is certainly the case for all segments examined, except for hind tarsus v of punctatus (Table 13), and antennal segment xv of impunctatus and palp segment v in pulicaris. Although the remaining values of α all suggest that α is significantly greater than zero, there is a tendency for a low value of α in the apical segments of the leg and palps of both pulicaris and punctatus. C. impunctatus differs in that the only location of a low value of α is in the tip of the antennae.

II Significance of the Deviation from α Equal to Unity

When $\alpha = 1$, the structures are termed isometric. Although this term is more commonly applied to growth studies, where two structures compared have the ratio of growth rates constant, it is a useful concept in size allometry to describe the constancy of 'shape' with change in absolute size. In a taxonomic context, isometric structures may be considered reliable taxonomic characters, in that they are not influenced by size.

In their study of Ectobius, Brown & Davies found that relatively few structures grew isometrically and consequently few simple

TABLE 13.

RESULTS OF SIGNIFICANCE TESTS: STRUCTURES FOR WHICH THE CALCULATED VALUE OF α DIFFERS LITTLE FROM ZERO

Appendage	Segment	punctatus	pulicaris	impunctatus
Antenna	xv			$P>0.05<0.10$
Maxillary Palp	v		$P>0.05<0.10$	
Hind Leg	Tarsus iv	$P>0.01<0.02$	$P>0.02<0.05$	
	Tarsus v	$P>0.10<0.20$		

comparisons of shape were possible between the sexes or species examined. In contrast, the results of the present study show that many structures have α values which do not differ significantly from 1 (at $p<0.05$). These are denoted by 'I' in Table 14. C. pulicaris and punctatus are similar in having about three quarters of the structures examined isometric, compared to C. impunctatus with nearly 90% isometric. All the palp segments are isometric for the three species, as are the leg segments of pulicaris and impunctatus. The majority of the deviations from isometry occur in the antennae but no common pattern of positive or negative allometry emerges among the species. A few comparisons may nevertheless be made. There is an area of negative allometry at the base of the flagellum in pulicaris and punctatus, and an area of positive allometry around the junction of the two parts of the antennae (segments x and xi) in pulicaris and impunctatus. The apical section of the antennae, with more elongate segments, has a higher incidence of positive or negative allometry than the basal section. Before further discussion and biological interpretation of allometry, it is necessary to ascertain whether the differential size relationships shown, constitute simple allometry, or more complex patterns, i.e., whether the relation between log x and log y differs from linearity.

III Significance of Deviations from Linearity

A structure shows simple allometry if α remains constant over the entire range of size (or growth). Brown & Davies outline three types of deviation:

(i) A progressive change of α with size.

(ii) A discrete change in the slope of a line plotted on double logarithmic coordinates.

(iii) Rhythmic fluctuations of α about an average value.

All of these yield significant deviations from the linear relationship as calculated by Bartlett's method. The double logarithmic plots given by Buxton (1938) for Pediculus show a clear departure from linearity in the final moult (type ii above). The deviation is different for ♂ and ♀ and demonstrates 'critical points' (points of inflection) in the sense of Reeve & Huxley (1945). Gould (1966, p. 589-600) discusses the significance and occurrence of deviations from linearity.

TABLE 14.

TYPES OF ALLOMETRY IN ANTENNA, PALP AND HIND LEG

Structure	Segment	pulicaris		punctatus		impunctatus	
Antenna	3	I	L	I	L	I	NL
	4	-ve	NL	-ve	L	I	L
	5	I	L	I	L	I	NL
	6	I	L	I	L	I	L
	7	+ve	NL	I	L	I	NL
	8	I	NL	I	L	I	L
	9	I	L	+ve	NL	I	NL
	10	-ve	L	I	NL	I	L
	11	+ve	L	I	L	+ve	NL
	12	I	L	I	L	+ve	NL
	13	+ve	NL	I	L	I	L
	14	I	L	I	L	I	L
	15	I	NL	I	NL	-ve	L
Palp	1+2	I	L	I	L	I	L
	3	I	L	I	L	I	L
	4	I	NL	I	L	I	L
	5	I	L	I	NL	I	L
Hind Leg	Femur	I	L	I	L	I	L
	Tibia	I	L	I	L	I	L
	Tarsus 1	I	L	+ve	L	I	L
	2	I	L	+ve	L	I	L
	3	I	L	I	L	I	L
	4	I	L	-ve	L	I	L
	5	I	L	-ve	L	I	L
% Isometry		79.16%		79.16%		87.5%	
% Linearity		75%		83%		75%	

I = Isometry

L = Linearity

(Simple allometry)

NL = Non linearity

-ve = negative allometry

+ve = positive allometry

Those structures which show a linear relationship with overall appendage length at P = 0.05% or less, are indicated by 'L' in Table 14. The symbol 'NL' denotes non-linearity.

It can be seen that the majority of structures, approximately 75%, exhibit allometry of size. This is very different from the results of Brown & Davies, where most of the differential growth was non-allometric. The high incidence of non-linearity found by these authors may be attributed to an abrupt change in shape, resulting from the moult of nymph to adult cockroach. Such an abrupt change in the slope of the regression line would not be expected in a study of size allometry, unless the dimensions of the structures studied are highly adaptive, or very intimately related to the environmental changes that affect overall size.

In Culicoides, the antennae show the most cases of non-linearity. Although such cases may require a higher order equation, or more complex function, than y = bx to describe the variation, there are problems associated with this in as much as any increase in number of parameters in the equation would make biological interpretation more difficult. It is of considerable taxonomic interest that the antenna has the highest incidence of non-linearity, as it contains many important taxonomic characters. However, the rather non-systematic and relatively scanty occurrence of allometry in antennal segments does not completely rule out their taxonomic use, provided sufficient attention is paid to the possible adverse effect of gross size. To paraphrase: the significance of different proportions of individual segments should be used with great caution as taxonomic criteria at the species level, preferably, the proportional differences of a number of segments should be used.

For an accurate assessment of allometry, discussion should ideally be confined to those segments that exhibit simple allometry (linearity, discussed above). Under these restrictions, only 50% of the observed deviations from isometry are pertinent, lying in the antennae of all species, and in the leg of C. punctatus. The taxonomic significance of allometry in the antennae has already been discussed, although it is difficult to advance any hypothesis for relating the structure to function, because there is no systematic occurrence of allometry. The negative allometric properties of over half the segments of the hind leg of C. punctatus

makes the use of tarsal ratio (tarsus i divided by tarsus ii) rather dangerous.

At this juncture, it is worth considering hypotheses for the physiological basis of size allometry. In this type, the 'growth' referred to in other studies may be loosely interpreted in terms of the expansion of the adult, following emergence from the pupa. Evidence for this may be taken from some studies of mosquitoes by van Heuvel (1963), in which it was shown that, at lower temperatures, expansion of adult Aedes aegypti was slower, but more prolonged, than at higher temperatures, resulting in a larger insect. There may, therefore, be a real growth component which may best be termed an 'expansion component'. Although the expression of size allometry may be viewed in terms of environmental conditions, it is only analogous to, and not homologous with, the post embryonic growth described in allometric growth studies.

7.5.8. Allometric Gradients

Of particular interest to the taxonomist are interspecific differences in the pattern of allometry, shown by any body appendage or axis. These patterns are most easily shown by plotting the values of α against successive segments of the structure studied, and then connecting them together by a line. The line is traditionally termed the growth gradient (Huxley, 1932), but allometric gradient may be a more acceptable term in the context of the present study. In studies of growth, the highest value of α corresponds to a 'growth centre', whereas in allometry of size, a high or low value of α would reflect differential response to genes (or polygenes) to specify size, including the influence of environmental variables such as temperature. Mather & Jinks (1971) gave an outline of the genetic basis of quantitative inheritance, including the polygene hypothesis of Mather (1943).

Examples of growth gradients along the main axis of the body are given by Blackith, Davies & Moy (1963), and Brown & Davies (1972). Among those for body appendages, Matsuda (1961) and Clark & Hersh (1939) are useful.

Allometry gradients were constructed for the antennae, palps and hind legs. Included in these diagrams are the 95% confidence intervals of α shown by a vertical bar.

Generally, the 95% confidence limits are rather large, and many of the considerable differences in α values are not significant in view of the great amount of individual variation. This is a good illustration of the need to use confidence limits in work on allometry. Based on these limits, the following discussion outlines the significant differences and general trends in the allometric gradients of the species studied. In the antennae (Figs.30a-c) of all the species, the allometry gradient oscillates around $\alpha = 1$ and does not show a general trend. This is not as expected, as most allometric analyses of insect appendages have demonstrated a marked tendency towards high negative values of α at the extremities of limbs. This trend is shown clearly in the hind legs of C. pulicaris (Fig.31a) and C. punctatus (Fig.31b) but, curiously enough, not in C. impunctatus (Fig.31c). In the hind legs of the former two species, the allometry gradient rises to a maximum in tarsal segments i and ii, and then declines to a value of $\alpha = 0.6$ in pulicaris and $\alpha = 1.5$ in punctatus. In these two species, the apical tarsal segments are proportionately very much smaller in the larger insects than the smaller ones. In most previous reports, the area of maximum α values (growth centre) lies in the femur and tibia, although Brown (1969) found the growth centre for Ectobius in the first two tarsal segments. The allometry gradient for the hind leg of impunctatus is similar in shape to that of the antennae.

A similar situation to that of the hind legs is shown by the palps. Again, pulicaris (Fig.32a) and punctatus (Fig.32b) show negative values of α at the distal region and the allometry gradient of impunctatus oscillates around $\alpha = 1$ (Fig.32c).

It is possible that the general stability of the segment proportions in the antennae with change in size, shown by the allometry gradients, may be due to sampling. In an attempt to obtain a taxonomically homogenous sample (for reasons given above) specimens of each species were collected at a single sample site. Insects exposed to much the same environmental effects were collected and consequently the size range of specmens obtained should not be as great as the species exhibits when development has taken place in variable conditions. The presence of different shaped allometry gradients for the antennae, palps and hind legs would suggest that sampling may not be significant in all of the species. It would be most interesting to take larvae from a single gene-pool and then

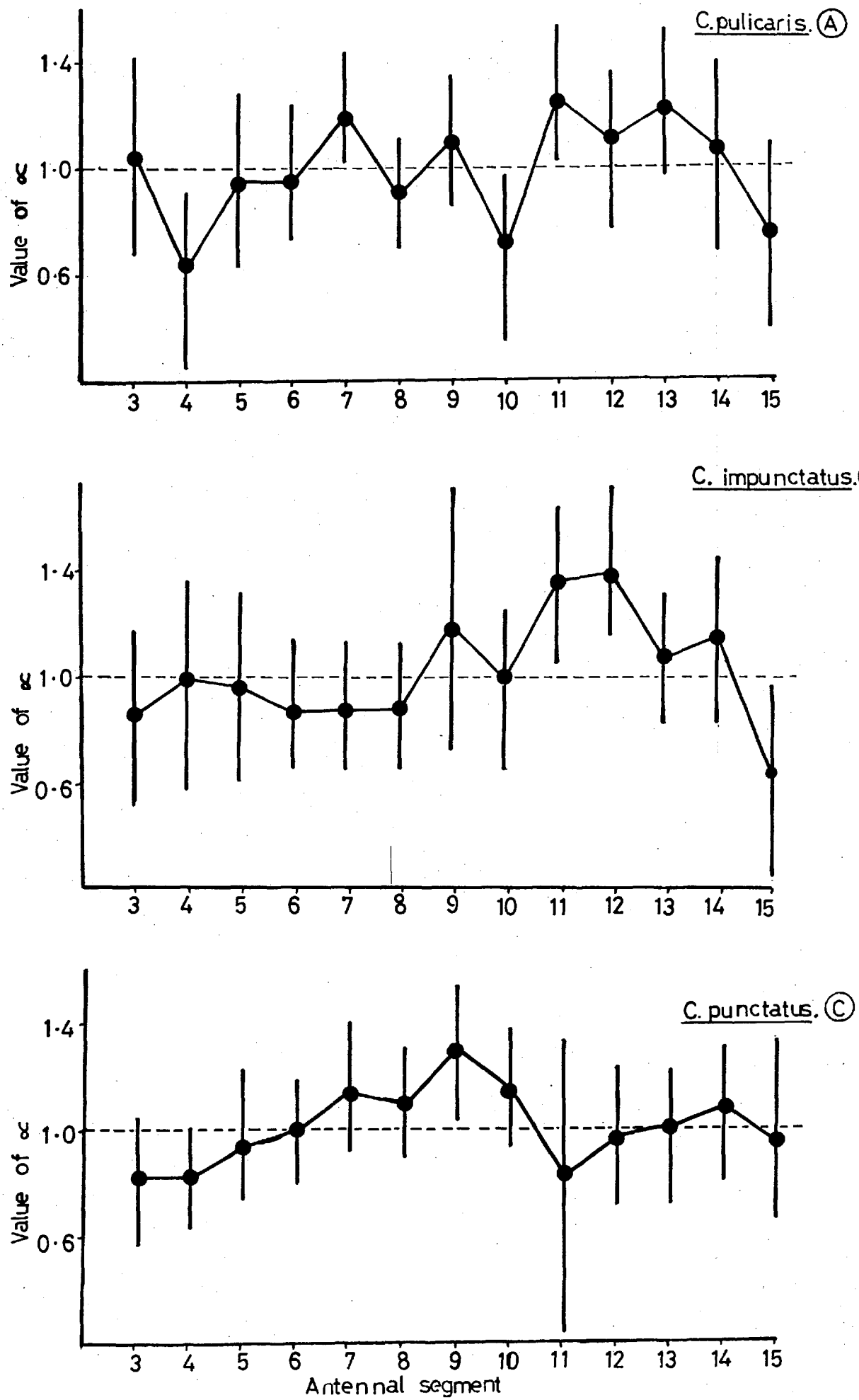


FIG. 30
ALLOMETRIC GRADIENTS IN THE ANTENNAE

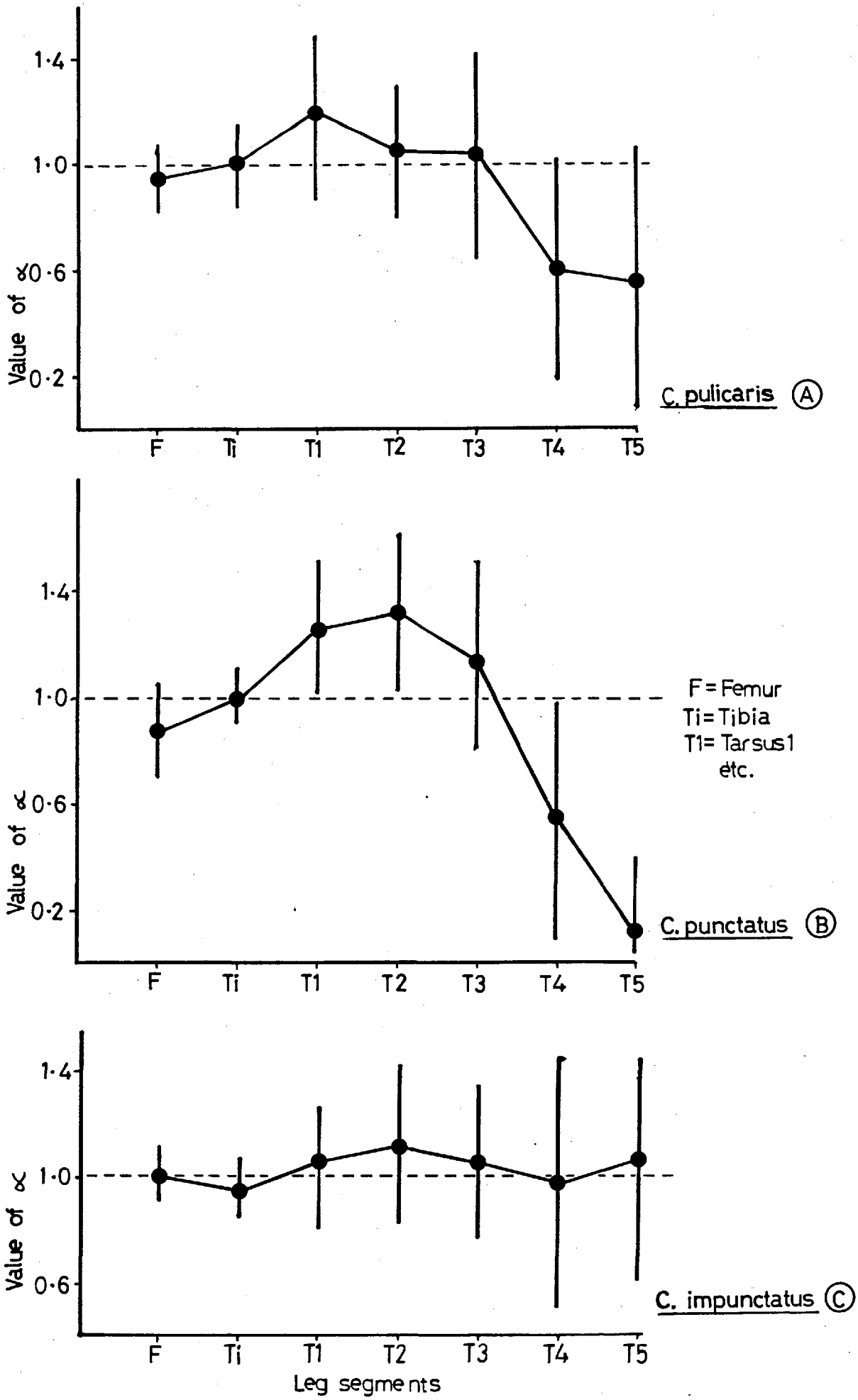


FIG. 31
ALLOMETRY GRADIENTS IN THE LEGS

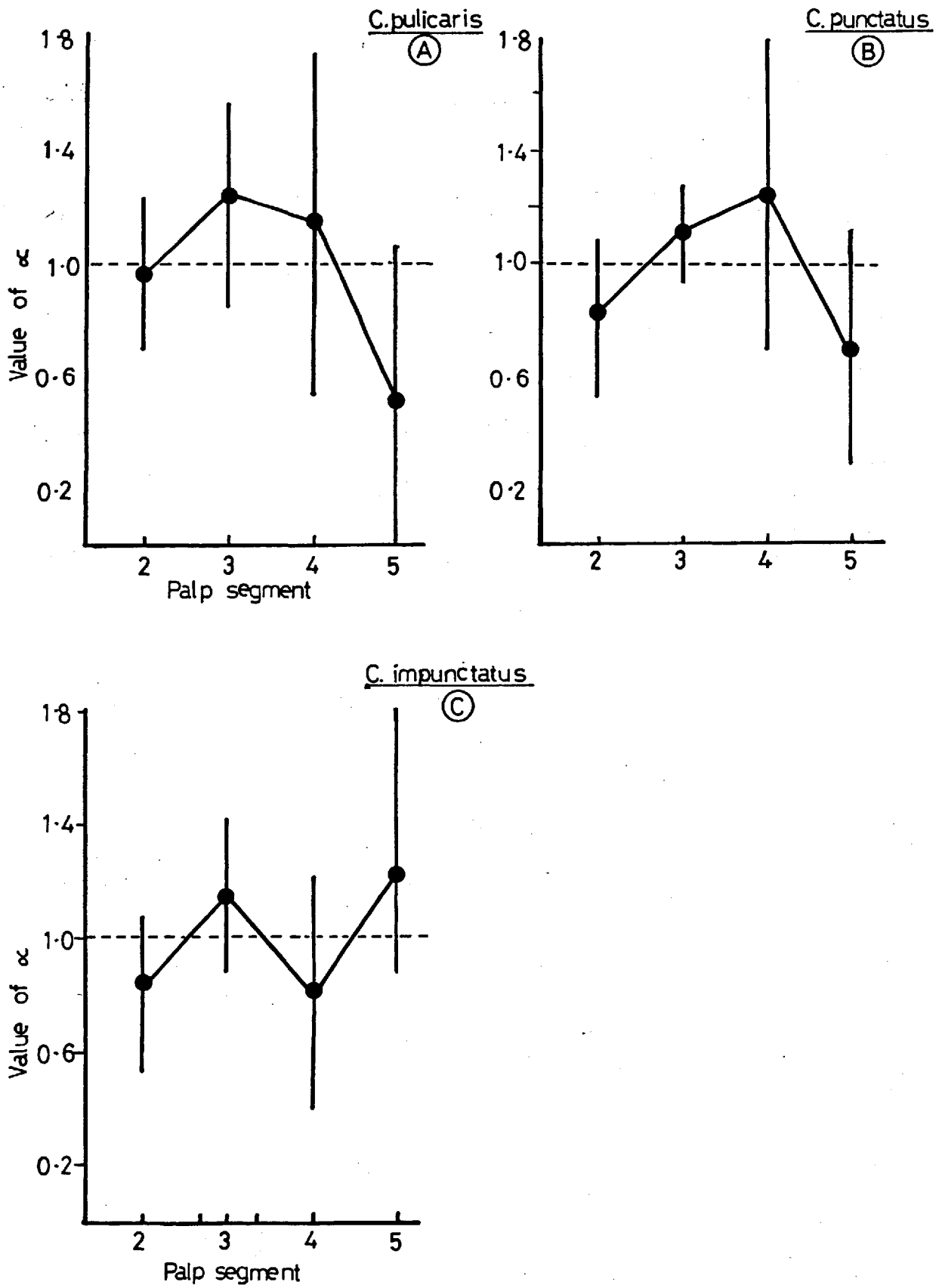


FIG. 32
ALLOMETRY GRADIENTS IN THE PALPS.

rear them under different temperature regimes to measure the effect, if any, of environmental conditions on the expression of allometry. Such a study would also help to show the relationship between quantitative inheritance and environment, and their joint influence on the shape of an insect. It should be noted that pooling of specimens from different geographical localities, to increase the range of environmental conditions sampled, would not help the problem but only complicate it further. If there is geographic variation in the allometry exponent, α , then the regression line calculated from the pooled data will not in general have the same slope as those calculated from the different localities. If the α values differ considerably, the slope for the pooled data will have little biological significance and may lead to erroneous taxonomic conclusions. This gives a clear hypothesis to test quantitatively and promises to be an interesting line for future work.

The lack of any regular occurrence of allometry in individual antennal segments may be explained in the light of the findings of the previous section. It was shown that the overall shape of the antennae was less variable than the lengths of individual segments, i.e., the total length, combined length of segments iii - x and combined length of segments xi - xv, and that they were probably specified early in development. The subsequent division of these antennal sections into individual segments was secondary, and therefore their individual lengths are less critical. If this is correct, then it is not surprising that there is no consistent occurrence of allometry in individual antennal segments.

Because all three species have similar habits, the adaptive significance of the observed differences in allometry gradients is not clear. All three species will bite readily, although impunctatus is a very common species and therefore more often recorded biting humans and other domestic animals. The taxonomic difference between the species lies in a few structural characters, (of impunctatus particularly) and difference in behaviour (mating swarms for example). It is common amongst species complexes of biting flies that the morphology of the group is remarkably consistant, with the principal evolutionary divergence in their larval habitats, behaviour and physiology. It is therefore not possible to advance any realistic causal hypothesis to account for the observed allometric properties in the sense of Gould (1966). In Culicoides it appears that a size

increase permits the expression of shape changes, rather than requiring them. Brown (1969) was able to relate allometric differences in legs of two species of Ectobius to different locomotory requirements in sandy and woodland habitats. Generally, very little attempt has been made to account for allometry parameters functionally. McMahon (1973, 1977) has recently emphasised this aspect of the subject, mainly for vertebrate skeletal systems and in respect of mass/metabolic relationships. It is an obvious field for development but probably not started in such similar insects as Culicoides species complexes.

Whilst there is difficulty in establishing the significance of different allometry gradients, their taxonomic value is much clearer. C. pulicaris and C. punctatus exhibit very similar patterns of variation for the palps and legs (shown by allometry gradients), both of which are distinct from impunctatus. The taxonomic affinity of pulicaris and punctatus is so marked that their validity as different species may be doubted. If the specific status of these two taxa is to be upheld by the use of behavioural data, for example, then it seems likely that the establishment of growth patterns preceded their speciation.

Section 8. THE ESTABLISHMENT OF NEW CHARACTERS

Prompted by the general lack of absolute diagnostic characters for separating members of the C. pulicaris complex, attempts were made to find new and useful characters. This section describes the characters tested and reports on their taxonomic value.

Owing to the importance given to the interspecific variation of wing patterns in most studies of Culicoides, it was necessary to incorporate these into this essentially numerical study. Consequently, a series of new characters were developed to quantify these patterns.

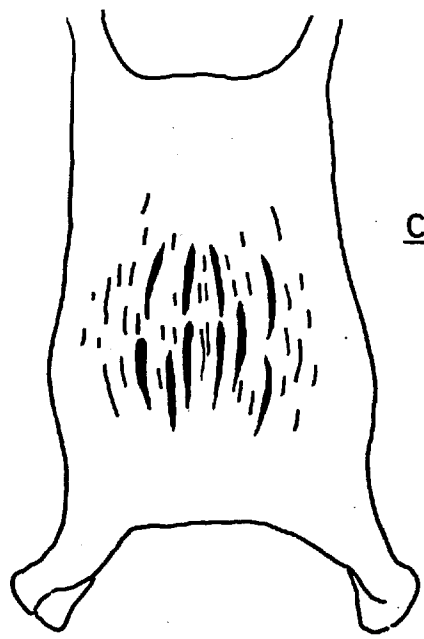
8.1. CIBARIUM

Unlike the Phlebotominae, the cibarium has not been used in the taxonomy of the Ceratopogonidae. Two aspects of this structure at the anterior end of the gut were therefore tested for possible taxonomic use; the presence of teeth on the cibarium and a new ratio, the cibarium/pharynx ratio.

Wirth & Blanton (1969) suggested that the cibarium has"yet to be found to have any taxonomic value", but detailed work by Callot, Kremer & Geiss (1972), and in the present study, shows this not to be the case. Within the pulicaris group, there is some variation in the shape and surface sculpturing. The cibarium of C. pulicaris was first described by Leon (1924).

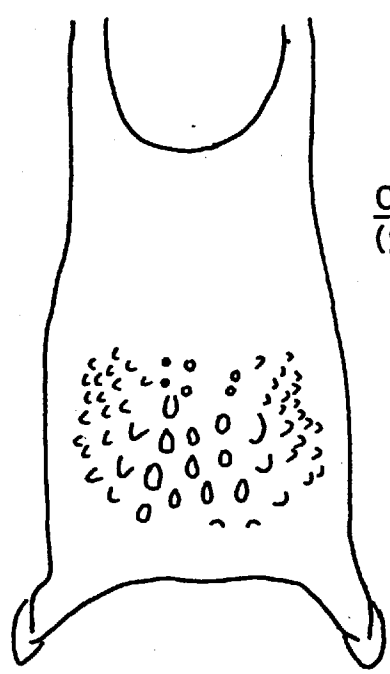
In most species of the pulicaris group, there is a series of longitudinal striations (Fig. 33), but in C. fagineus a distinct series of vertical teeth are present (Fig. 34). This character distinguishes fagineus from all other members of the pulicaris complex.

The function of these teeth is not known, but may be associated with the physical rupturing of ingested blood corpuscles, prior to chemical digestion. In the Phlebotominae, where the cibarial armature is well developed, Lewis (1974) found evidence that they acted as a comb, filtering out large parasites, and thus may influence the transmission of a parasite by the biting fly (Lewis, 1975). Similar work has been done by McGreevy et al. (1978) with mosquitoes. This aspect of morphology could be of particular importance in predicting which species of Culicoides are possible vectors of protozoan and nematode parasites. The structure of the



(33)

C. pulicaris



(34)

C. fagineus
(paratype)

FIG. 33-34
VARIATION IN SCULPTURING OF CIBARIUM

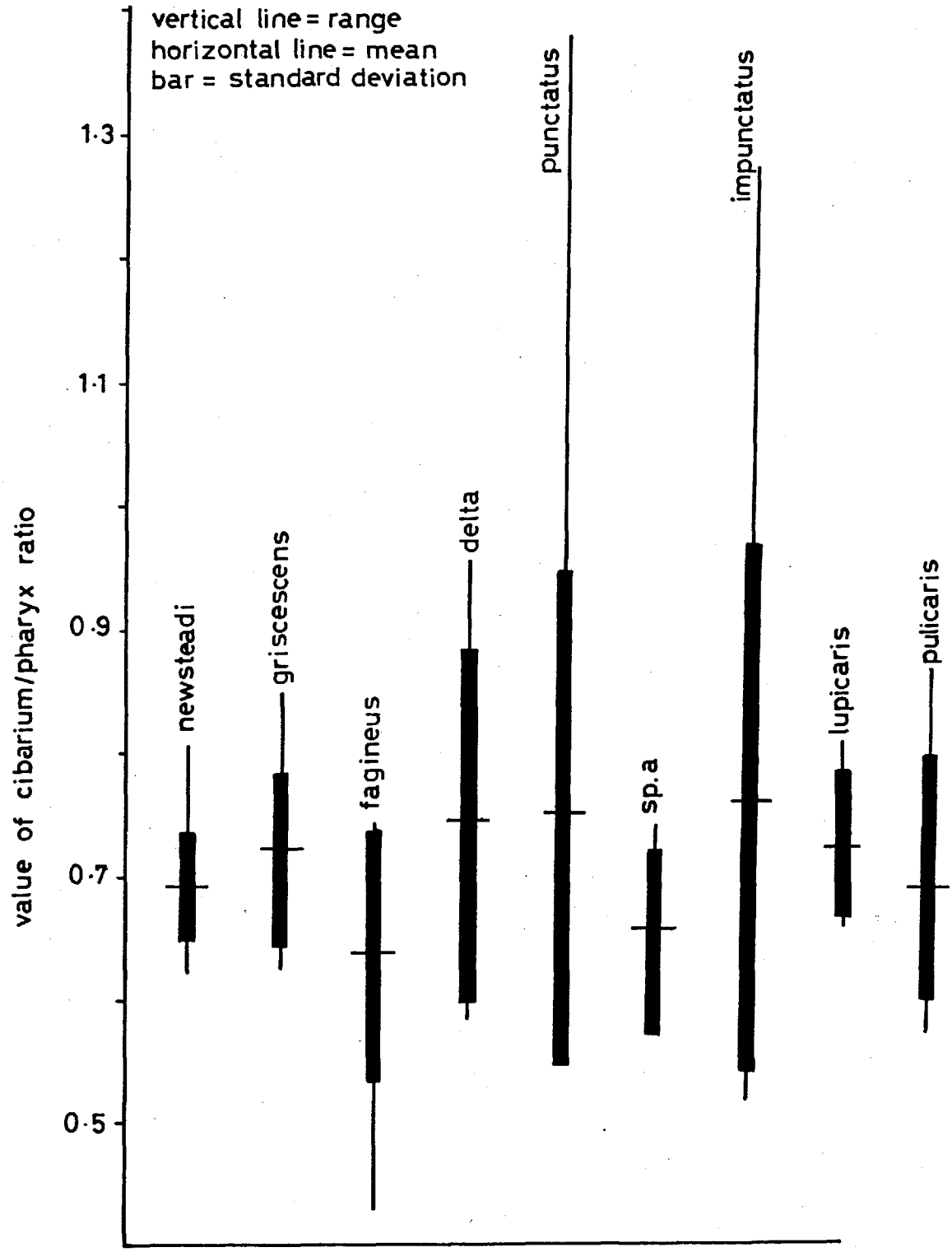


FIG. 35
VARIATION OF CIBARIUM/PHARYNX RATIO IN THE C.PULICARIS COMPLEX

cibarial sensilla of mosquitoes has been described by Lee & Davies (1978) and related to possible function by Lee (1974).

There may be some relation between the structure of the cibarium and feeding preferences in the C. pulicaris complex. All the species of the complex have been observed feeding on ungulates and man, except C. fagineus. Although there is no positive evidence, the presence of the teeth on the cibarium of fagineus may be correlated with a different host preference, e.g., birds, or even reptiles.

The cibarium/pharynx ratio is the length of the cibarium divided by the length of the pharynx. A summary of the variation in the ratio is presented in the form of a Dice-Leraas diagram Fig. 35). It is clear that no one sample has a mean outside one standard deviation from the mean of any other sample. By this criterion, it can be concluded that the means are not significantly different. These data eliminate the necessity to carry out 't' tests for significance (Simpson, Roe & Lewontin, 1960, p. 353). There is a marked difference in the observed ranges of the species. They vary from a range of 0.18 in newsteadi to 0.83 in punctatus. The variation does not appear to be the result of different sample sizes or the fact that each sample is based on specimens from different localities, e.g., in punctatus, a series of specimens from Norway exhibit a range in values from 0.66 - 1.38.

In conclusion, this ratio is of little taxonomic use in the C. pulicaris complex.

8.2. CHAETOTAXY

The arrangement of setae (chaetotaxy) is used extensively in the taxonomy of many groups of Diptera, but has been little used in the Ceratopogonidae, owing to the lack of large setae.

In Culicoides, though, a few well developed setae are present on the hind legs, forming the hind tibial combs. In some species groups, differences in the number of these spines is sufficiently great to be of taxonomic value. Unfortunately, this is not the case in the pulicaris group, where each species has a variable number of spines, usually from five to seven. Occasionally, the number of spines differs between the two tibial combs of one individual.

Culicoides also possess small setae on the vertex of the head,

which have not so far been used in taxonomic studies of the genus. The arrangement of these setae was therefore examined and an attempt made to develop a system of nomenclature to facilitate the evaluation of their taxonomic significance. The setae are short and scattered. They are often broken off during collection and mounting, but as with all true setae, the sockets remain. The sockets were used in this study.

Accurate drawings were made of a sample of five specimens of each species (except lupicaris, where only two specimens were available), with a drawing tube attached to a Wild M11 microscope.

The general pattern of setae was the same for all the species examined, with large setae along the margin of the eye and smaller setae, half their size, on the main surface of the vertex. Those on the eye margins were numbered, starting from where the eyes are contiguous. A series of arcs were then superimposed on the drawings, to connect the setae on the eye margins to those on the surface of the vertex (Fig. 36). Each arc was then numbered after the large setae at its ends.

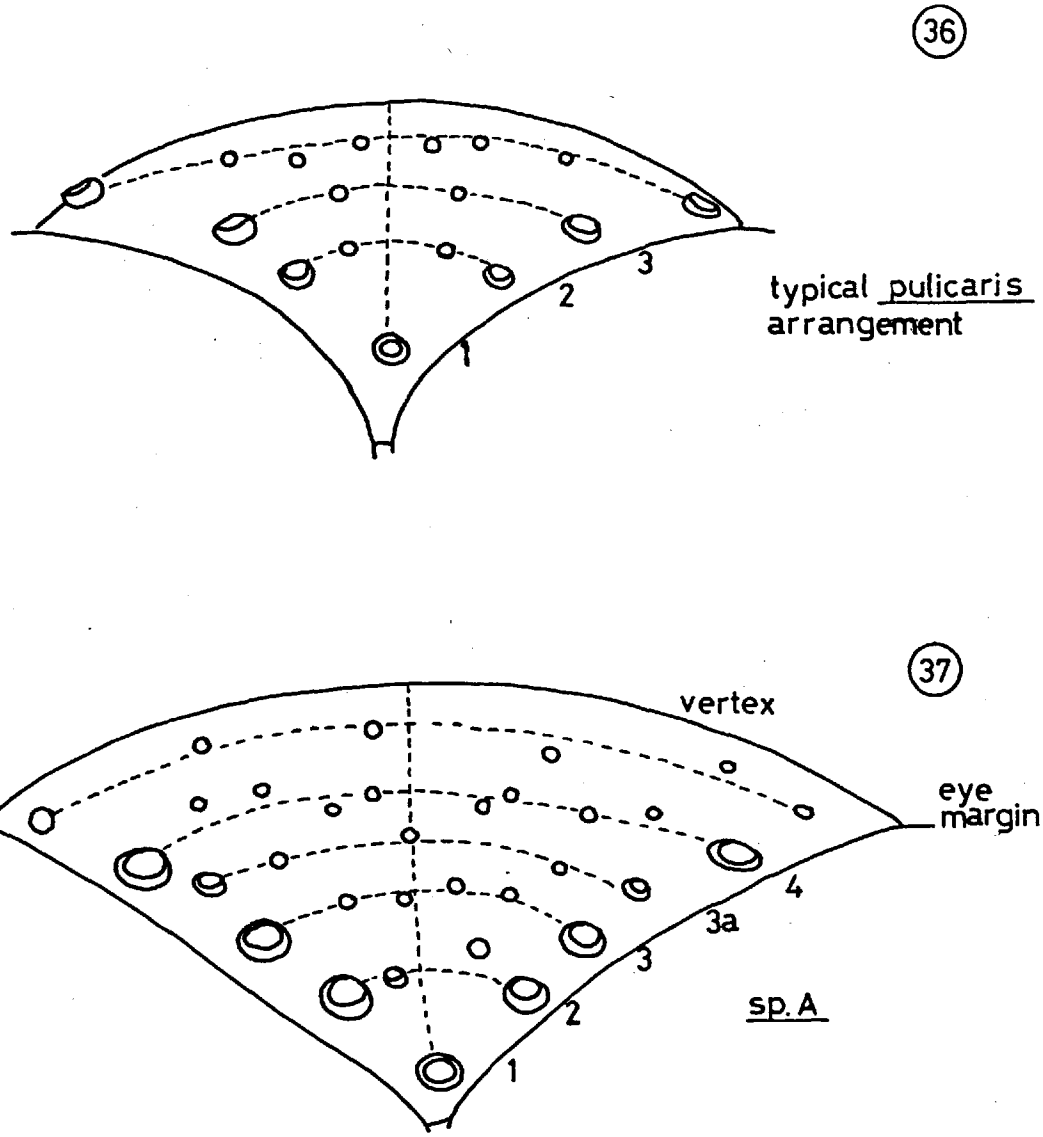
Within the pulicaris complex, the usual arrangement of setae was as follows:

Row 1 : 2 spines
Row 2 : 4 spines
Row 3 : 4 spines
Row 4 : 6 spines.

A sample from Japan, referred to as sp. A, had many more setae than the other species (Fig. 37). This sample had an additional row, bounded by two small setae at the eye margins. Furthermore, row 4 had ten setae, compared with the usual six in the rest of the complex.

Although some differences emerged between the species, chaetotaxy did not render itself a practical technique, for the following reasons:

1. Variation within most species was as great as that between species. The only exception was the sample from Japan (sp. A).
2. The frequent lack of bilateral symmetry made interpretation of the drawings and homology of some setae difficult. In the higher Diptera, (e.g., Muscidae), where setae are used extensively in taxonomy, there is considerable bilateral symmetry.
3. Although specimens were slide mounted, they were not consistantly orientated in an anterior/posterior plane, making observation of



FIGS 36-37
CHAETOTAXY OF VERTEX IN CULICOIDES

setae towards the top of the head (row 4) very difficult. It was not practical to remount many of the specimens, as this often resulted in obscuring other characters, such as mandibles, maxillae and cibarium.

8.3. WING PATTERN

8.3.1. Introduction

Wing pattern has played a major rôle in the taxonomy of Culicoides and no study would be complete without discussing it. Much of the initial separation of species in the pulicaris group was based on wing pattern features, and although considerable emphasis is placed on patterns in defining species groups within the genus Culicoides, there has been little investigation into these characters. The nature of the intrinsic variation of wing patterns is occasionally discussed in the description of species, but not on a quantitative basis. The subject of homologies in the pattern has also been rather neglected. It is the object of this section to develop a reliable and adaptable method for coding wing pattern and incorporating it into a quantitative study.

The quantitative expression of shape presents many problems in numerical taxonomy and morphometrics. Compared to the large number of papers developing and refining sophisticated statistical techniques for the analysis of numerical data, the coding of shape has lagged far behind, and still remains an area where much work needs to be done. Though there have been attempts to describe shape in purely mathematical terms, these have resorted to complex mathematical functions. Sneath (1967) attempted to describe variation in the overall form of hominid skulls by trend-surface analysis, based on the transformation grids of D'Arcy Thompson. The two dimensional outlines of Molluscs were described and compared by Younker & Ehrlich (1977), using Fourier analysis of polar coordinates. Meltzer, Searle & Brown (1967) used mathematical functions (called Walsh functions) to study leaf shape. The last two of these studies applied only to outlines and relied on operational homology. Such an approach to homology is based on the phenetic similarity of structures on different organisms (often simply the position relative to some named point) and does not incorporate any of the

phylogenetic inferences used to establish homologies in other studies. Operational homology was the basis of a semi-automatic method for recording data by Moss & Power (1975) in which a coordinate digitiser was used to record the position of a number of intersections of structures (these intersections were used as characters). The study was based on hypothetical 'caminalcules'. Despite the shortcomings of operational homology, Moss & Power found their method gave a similar result to those based on an intuitive approach.

In this section, two contrasting methods are developed for coding wing pattern, one based on scanning the wing mechanically and the other based on extracting pattern elements, with some claim to be regarded as biologically meaningful. Each coding method was used to produce a classification of a sample of wings. By comparing the classifications, the two coding methods could be evaluated.

Previous Studies of Wing Pattern

When patterned wings have been used in the past, for taxonomic purposes, a number of approaches to the method of description have been developed. They have been applied to a wide range of insect orders, and may be divided into two broad categories: 'descriptive methods' and 'homology methods'.

The desriptive methods are by far the most common in the literature and attempt to express the relationship between taxa in a purely phenetic sense. The first, and generally less elegant, use a series of 'types' selected from the overall range, or commonest patterns, in the group being studied. For example, Munroe (1947) recognised a series of five pattern types in the African Trypetidae (Diptera). When subsequently discussing a given species, its wing pattern was stated as being of a basic type and any significant deviation from the reference was described. This was a commonly used method of describing patterns for use in taxonomic discussions.

Perhaps the simplest application of this approach illustrates (graphically and verbally) the wing pattern of each species, without any attempt to synthesis 'types' at all.

The second group of descriptive methods includes those in which a generalised scheme of pattern elements is given. For example, Khamala & Kettle (1971) named the various spots on a Culicoides wing.

In this group of methods, the homology of pattern elements is implied by the construction of a 'generalised wing'. It relies on associating a particular spot on a specimen with a spot on the generalised wing diagram. The exact determination of a homology is therefore delegated to the reader. The work of Oldroyd (1952) is a rather sophisticated example of the descriptive methods, in which he proposes a 'dynamic' explanation for the wing pattern of Haematopota (Tabanidae). His hypothesis suggests a system of diffusing pigment, or bleaching agent, radiating from specific foci on the wing.

Usually, descriptive methods are based on wing venation, i.e., the wing pattern is described relative to the venational nomenclature of Tillyard or Comstock & Needham. This approach has the implied assumption that wing veins are important physical (or physiological) agents, influencing the development of wing pattern. Pattern and venation may well be correlated to some degree, although a priori, there is no evidence to suggest why genetic specification for an essentially structural system (wing veins) should be closely linked (or associated) with specification for the pigmentation of an overlying epidermal sheet (colour pattern). Therefore, it is not necessary to restrict the description of patterns to venational terms (further evidence on this point will be discussed later).

A somewhat mechanical method for coding patterned wings - termed the scanning method - using a superimposed grid, is described below (page 151).

The homology methods - termed 'prototype' by Schwanwitsch (1924) and 'colour pattern plan' by Shelford (1917) are purely hypothetical. Such models express the relationship between fundamental pattern element and therefore, of necessity, are abstract in nature. It is the hypothetical nature of these models, and consequently the homologies they elucidate, that is the key to their use. Some workers, notably Eimer (1897) and to some extent Shelford (1917), have imputed ancestral status to these models. However, Schwanwitsch emphasised that his prototype model was developed solely for the duduction of homologies and not for that of ancestors. Phylogenies may subsequently be constructed, using evidence from homologies deduced from these models, as shown by Vane-Wright & Huggins (1972) and to a lesser degree Graham (1950).

The methods for constructing homology models vary in detail, although all utilise the central principle of a sequential arrangement of extant organisms. Schwanwitsch gives little indication of his methodology which led to such a profusion of papers. The only clue is to be found in the statement when he refers to Eimer's work; "the inadequacy of his results is quite natural, as he neglected the method of carefully constructed morphological series, which at present is the only sure way of establishing pattern homologies". In contrast, Shelford described in considerable detail how he extracted the longitudinal pattern elements of his tiger beetle elytra. These longitudinal elements were then superimposed on a series of equally stringently determined lateral pattern-elements. This resulted in a framework from which he could then extract homologies. A further method for building a homology model, termed the pattern element method, is described on page 163.

8.3.2. Materials

For this study, the wings of 22 specimens of the Culicoides pulicaris species group were selected to represent the range of pattern variation within the species group. The specimens were slide mounted and viewed under dark field illumination.

OTU's were numbered 1 - 23 as shown in Table 15. Wings were drawn on graph paper to a standard size, with the aid of a Wild drawing tube attached to a Wild M 11 microscope. This ensured that gross size of the wing was eliminated from the analysis. For pattern coding, the wings were orientated with the costa lying along the uppermost edge of the grid.(Fig. 38).

8.3.3. The Scanning Method

Perhaps one of the more obvious methods of comparing wing patterns is to scan the wings mechanically, noting the presence or absence of pigment. The object is to produce a series of matrices, which are numerical images of the specimens, and then compare them. This is easily achieved by superimposing a grid onto a wing and noting the presence or absence of pigment within each grid compartment. Such a method, operating directly on observed data, will inevitably produce a descriptive model.

TABLE 15. Source of Wings Used in Section 8 (Pattern Analysis).

Code number	Species	Origin
1	delta	Scotland: Arran
2	grisescens	Scotland: Cromarty
3	fagineus	England: Hampshire (pt)
4	lupicaris	Scotland: Lanarkshire (pt)
5	impunctatus	England: Northumberland
6	pulicaris	England: Hampshire
7	pulicaris	England: Surrey
8	punctatus	England: Surrey
9	punctatus	England: Surrey
10	pulicaris	England: Surrey
11	impunctatus	Scotland: Inverness
12	impunctatus	Scotland: Inverness
13	impunctatus	Scotland: Inverness
14	impunctatus	England: Surrey
15	fagineus	Israel: Tel Aviv
16	grisescens	Wales: Montgomeryshire
17	fagineus	England: Hampshire (pt)
18	fagineus	England: Hampshire (pt)
19	grisescens	Wales: Montgomeryshire
20	grisescens	Wales: Cardiganshire
21	grisescens	Wales: Montgomeryshire
22	fagineus	England: Hampshire (pt)
23	fagineus	as OTU 18, with 17 coding differences

pt = paratype

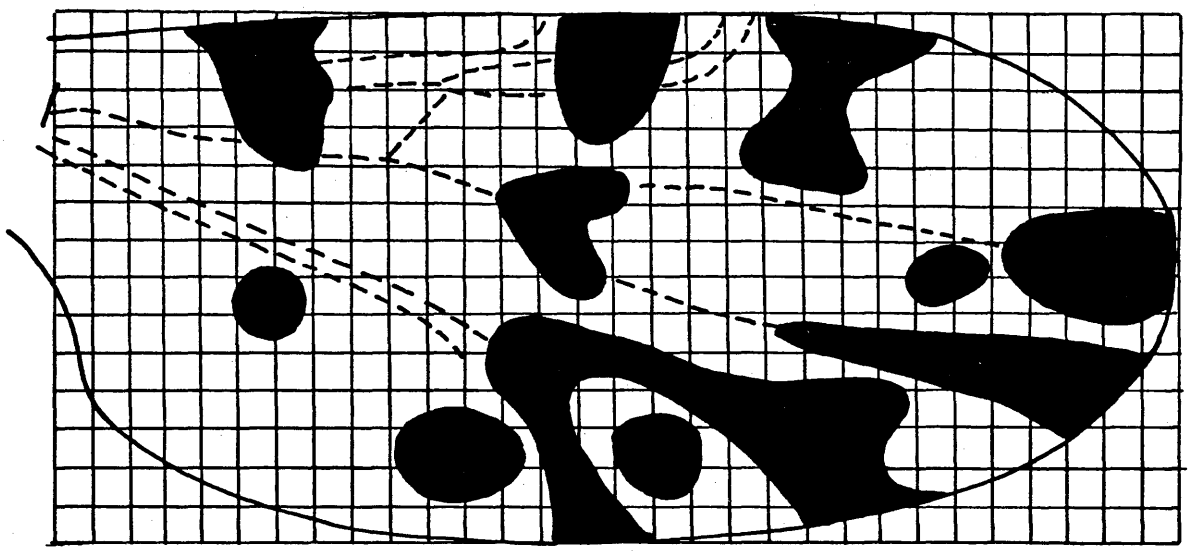


FIG. 38
GRID SUPERIMPOSED ON A CULICOIDES WING

```
0 0 0 0 1 1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 1 1 1 X X X X X X X
0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 0 0 0 1 1 0 0 0 0 X X X X
0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 X X
0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 X
0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 X
0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1
0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 1 1 1
X 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 X
X 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 0 0 0 1 1 1 1 1 1 1 1 X
X 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 1 1 1 1 1 1 0 0 0 1 1 X X
X 0 0 0 0 0 0 0 0 0 1 1 1 1 1 0 1 1 0 1 1 1 1 0 0 0 0 0 X X X
X X X X 0 0 0 0 0 1 1 1 0 1 1 0 1 0 0 1 1 1 1 0 X X X X X X
X X X X X X X X 0 0 0 0 0 1 1 1 0 0 0 0 X X X X X X X X X X
```

FIG. 39
CODED MATRIX CORRESPONDING TO THE WING ABOVE

Rohlf & Sokal (1967) have investigated some properties of scanning images using a sample of 25 "caminalcules". They found that the taxonomic relationships between the caminalcules resulting from scanned data and 'conventional numerical taxonomic methods' were 'acceptably similar'.

In the present study, a grid of 30 x 14 units was superimposed on the wing drawings (Fig. 38). Each square of the grid was coded for the presence or absence of pigmentation. Pigmentation was deemed present when the area of pattern within each square exceeded 50%. Presence of pigment was coded 1; absence was coded 0; and no-wing (at the corners of the grid) coded X. This produced a series of matrices of 420 characters (Fig. 39). Because of the difficulties occasionally encountered in deciding whether 50% of a grid was pigmented, the borderline squares between 'pigmented' and 'non-pigmented' were re-coded for one of the wings (OTU 18). This new OTU (OTU 23 in Table 15) had 17 of the 420 grid squares (4%) re-coded. OTU 23 was incorporated into the analysis, in addition to OTU 18, in order that the effect of varying the coding for a small proportion of characters could be tested.

The method employed here differs from that of Rohlf & Sokal (1967) in several ways. Firstly, they did not distinguish between the absence of a structure and the absence of an organism on the grid (i.e., non-comparable characters were not recognised). Secondly, they did not have any minimum criterion for determining the presence of a structure within a grid cell. If the structure was at all visible within the grid cell, it was recorded as present.

Once the problem of quantifying wing pattern has been resolved, the next problem is the comparison of the matrices describing each wing. Clearly, a correlation coefficient is not applicable for comparing binary data, so a distance measure has to be used. Furthermore, this measure of similarity is required to compare both binary data and ordered multistate data, thus, when classifications based on both coding methods are compared, the effect of alternative numerical treatment of the data will be reduced to a minimum. A general similarity coefficient which fits these requirements is that of Gower's general similarity coefficient (Gower, 1971a; Sneath & Sokal, 1973). When applied to binary characters, this coefficient reduces to match-mismatch coefficient, in which both joint presence of joint absence of a character state is considered a match. A similarity matrix was constructed using this coefficient

and , after conversion to a distance matrix, a principal coordinate analysis was carried out. A summary of the results is given in Table 16. In this analysis, the first eigenvector accountsfor only 23% of the total variance - a comparitively low proportion - which is increased to 38% by the inclusion of the second eigenvector. The fourth and consecutive eigenvalues are fairly similar in magnitude, suggesting that most of the taxonomically interesting variation is described in three dimensions.

When the first three eigenvectors are plotted against one another (Fig. 40), there are fewer discrete clusters of specimens than might be expected, as several of the wings were taken from specimens collected in the same locality and at the same time. In the plot of the first versus the second eigenvector (Fig. 40a), OTU's 18 and 23 appear to be very close. This suggests that a relatively small number of miscodings, or dubious codings (in this case 4%), does not unduly affect the relationships of an individual. Specimens of C. grisescens (2,16,19,20,21) form a rather diffuse but recognisable group to the right of the diagram. In contrast, wings of C. fagineus (3,15,17,18,22,23) are spread out all over the plot. The same is true for impunctatus (5,11,12,13,14), but to a lesser degree. Some pulicaris and punctatus form a tight cluster (8,9,10) in both plots of the first versus the second, and second versus the third eigenvectors.

The plot of the second versus the third eigenvector (Fig. 40b) shows much the same as the first, but with clearer groupings, e.g., specimens of grisescens.

The association of a particular apect of the wing pattern with any of the first three eigenvectors is difficult. The general degree of pigmentation of the wing is analogous to 'general size' and would be expected to be the main factor influencing the position of OTU's along the first axis.

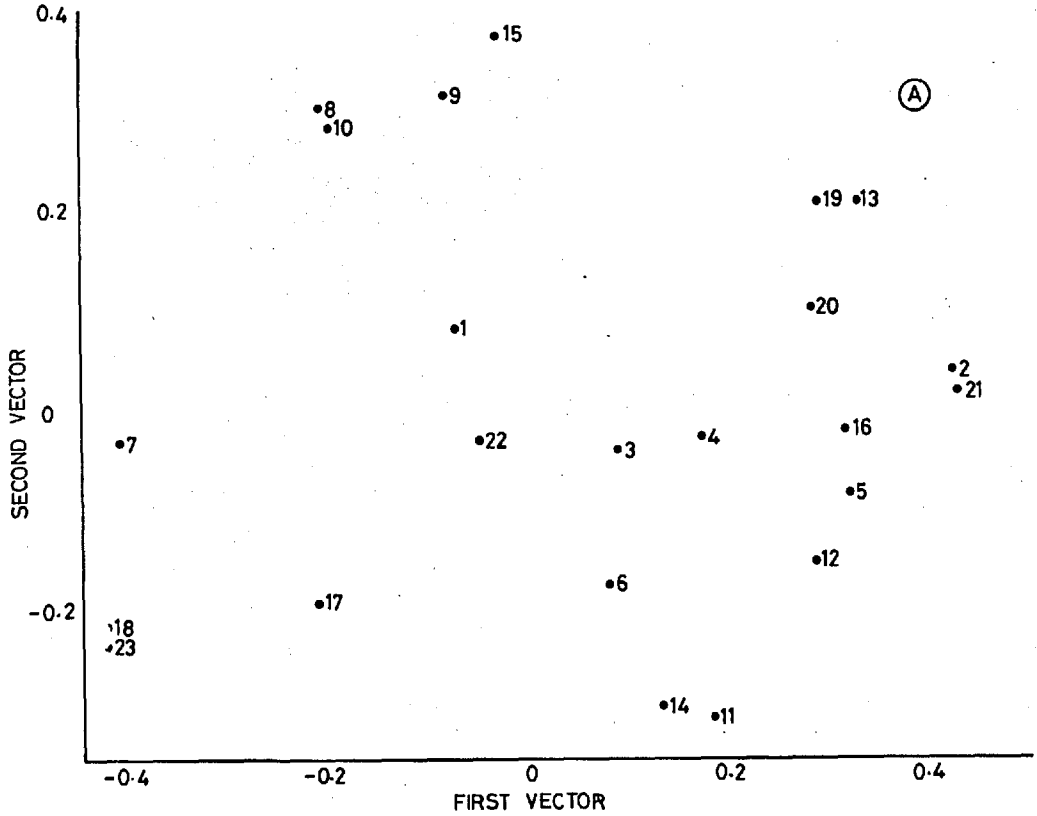
When the total number of grid squares for each wing is plotted against the value of the first principal coordinate (elements of the first eigenvector), a rather vague association between them is revealed. However, if the number of grid squares coded 1 are expressed as a proportion of the total number of squares covering the wing (i.e., 420 - the number of non-comparable squares), then the relationship between pigmentation and the first eigenvector becomes clearer (Fig. 41, Table 17). This results from the variable number of non-comparable coded squares in the matrices of the different wings. The first eigenvector may also be interpreted as

TABLE 16.

Summary of Principal Coordinate Analysis of 23 Wings Coded by the Scanning Method.

	1	2	3	4	5
Eigenvalue	1.638	0.976	0.653	0.378	0.352
Cumulative Percentage of Trace	23.83	38.03	47.54	53.04	58.16
OTU		Eigenvectors			
1	-0.124	0.088	-0.180	-0.345	0.054
2	0.377	0.036	0.108	-0.012	0.181
3	-0.042	-0.043	-0.073	-0.064	-0.205
4	0.125	-0.232	-0.144	0.129	0.169
5	0.273	-0.086	-0.078	-0.031	-0.119
6	0.033	-0.171	-0.347	-0.139	-0.051
7	-0.466	0.029	-0.081	0.015	0.269
8	-0.261	0.300	-0.165	0.129	0.084
9	-0.136	0.319	-0.095	0.109	-0.088
10	-0.259	0.302	-0.123	0.204	0.073
11	0.133	-0.307	-0.262	-0.036	0.055
12	0.242	-0.155	-0.084	0.833	-0.085
13	0.276	0.205	0.050	-0.147	-0.067
14	0.095	-0.297	-0.045	0.184	-0.099
15	-0.081	0.372	-0.001	-0.116	-0.185
16	0.266	-0.027	0.199	0.036	0.045
17	-0.267	-0.193	0.098	0.084	-0.104
18	-0.470	-0.229	0.274	-0.116	0.023
19	0.235	0.216	0.171	-0.007	0.462
20	0.231	0.101	0.224	-0.024	0.052
21	0.384	0.030	0.140	-0.017	0.167
22	-0.100	-0.031	0.151	0.187	-0.182
23	-0.467	-0.228	0.265	-0.112	0.005

The OTU's may be identified from Table 15.



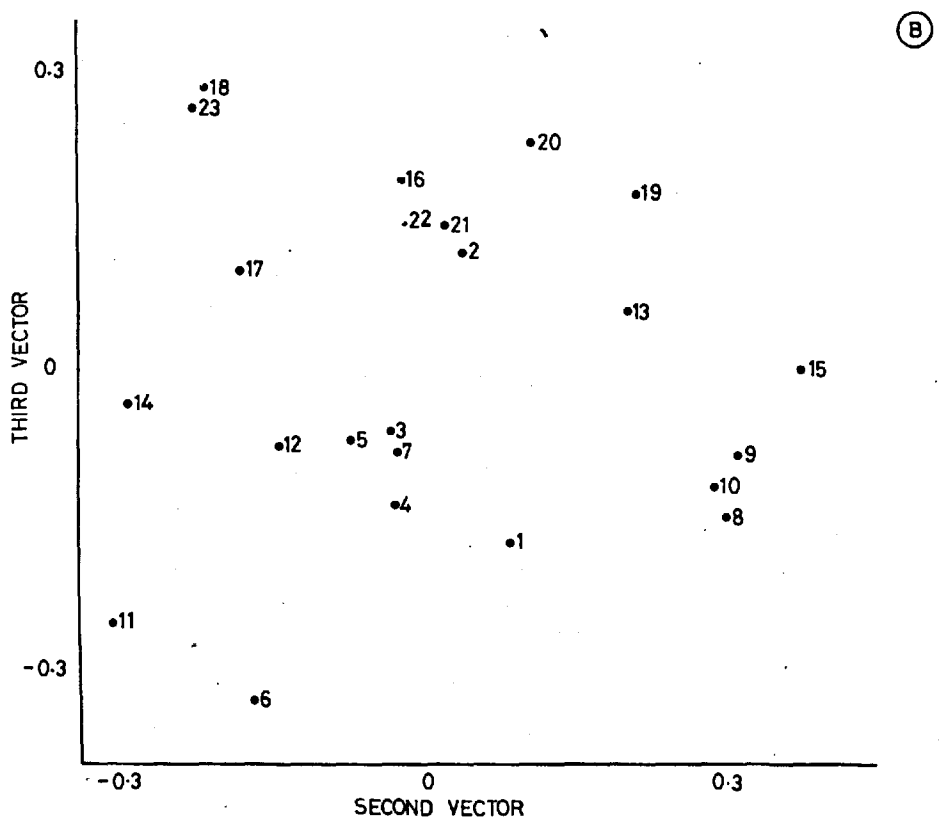


FIG. 40
PLOTS OF EIGENVECTORS FROM PRINCIPAL COORDINATE ANALYSIS OF 23 WINGS CODED BY THE SCANNING METHOD

158

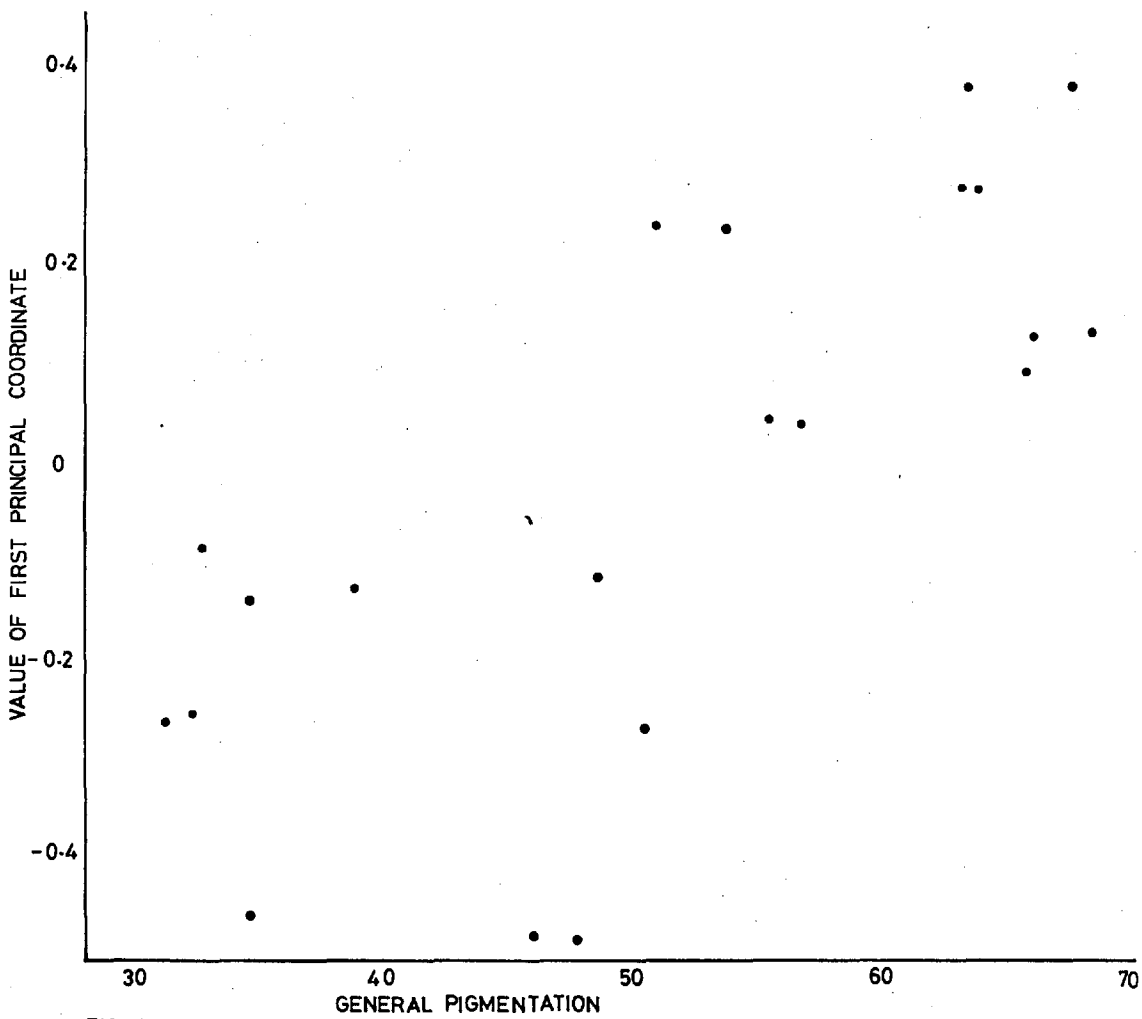


FIG. 41
ASSOCIATION OF WING PIGMENTATION WITH FIRST EIGENVECTOR
(FROM PRINCIPAL COORDINATES ANALYSIS OF WINGS CODED BY SCANNING METHOD)

TABLE 17.

Summary of Scan-Matrices for 23 Wings.

OTU	Number of Squares Coded			Total Number of Squares Covering the Wing (1+0)
	1	0	non-comparable	
1	116	186	118	302
2	220	128	72	348
3	181	145	94	326
4	223	115	82	338
5	207	122	91	329
6	191	148	81	338
7	125	236	59	361
8	108	237	75	345
9	122	232	66	354
10	110	232	78	342
11	205	95	120	300
12	216	116	88	332
13	169	167	84	336
14	234	123	63	357
15	105	227	88	332
16	230	132	58	362
17	176	172	71	349
18	167	181	72	348
19	176	175	69	351
20	187	161	72	348
21	238	115	67	353
22	163	190	67	353
23	171	181	68	352

The OTU's may be identified from Table 15.

a measure of the contrast between pale and dark areas of the wing. The somewhat loose association between pigmentation and the first axis may be a consequence of the coefficient employed here, which does not distinguish between joint presence (two 1's) or joint absence (two 0's) of a trait.

It was not possible to associate aspects of the wing pattern with any of the other axes.

Weighting of Characters

The scanning method discussed on page 151 may be extended by weighting characters according to the frequency of character states (pigmentation). The frequency of pigmentation at each grid locus on the wing, over all 23 OTU's, was calculated. The frequency of of pigmentation is graphically portrayed in Fig. 42, by means of a three-dimensional projection. The peaks show areas of the wing at which the presence of pigmentation is most frequent in the wings studied. Characters may be weighted inversely with respect to frequency of pigmentation. This assumes that greater information is contained within characters which have a low frequency of pigmentation within the sample, i.e., they may be considered 'rare characters'. Such characters contribute more when assessing the affinities of the OTU's. This is not strictly weighting for conservatism, because characters which are always pale are not weighted. The method weights against high frequency of a character state, in this case black pigment. Strict conservatism would strongly weight a grid square which is pale, when the majority of the sample is pigmented.

Two possible methods of determining the weight were considered. Firstly, 1 - p and secondly 1/p, where p = frequency of pigmentation over all OTU's. The weight 1 - p gives a lower weight for rare characters, relative to p = 50% than 1/p. 1 - p gives a higher weight for medium-common characters than for very common characters. Thus, if the frequency of pigmentation (p) is mainly in the range 0 - 50%, then 1/p is a more effective weight, as 1 - p will have very little effect. Conversely, if p is mostly in the range 50 - 100%, then 1 - p should be used. The weight 1/p was rejected as it gives greater weight to the rare occurrence of pigment at a locus and may produce a larger difference than is warranted between two identical wings, when there is an occasional miscoding or slight 'frame-shift' of the grid.

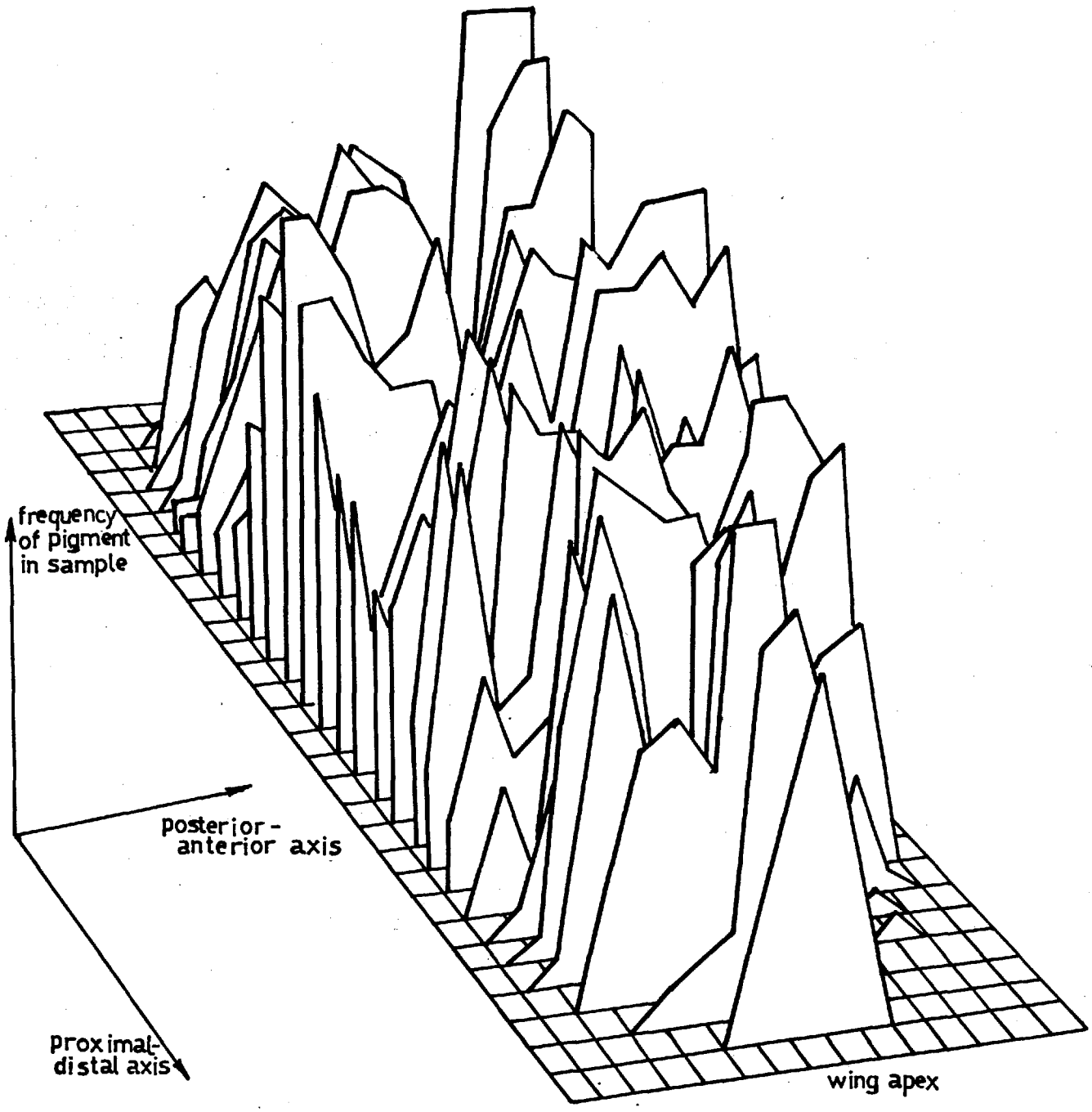


FIG. 42
FREQUENCY OF PIGMENT AT EACH GRID SQUARE

The very small difference due to error would therefore be amplified.

The data matrix transformed by weighting of 1 - p was analysed by principal coordinate analysis and found to differ only very slightly from the unweighted matrix.

Theoretical Problems with the Scanning Method

There are a number of criticisms of a practical and theoretical nature applicable to the mechanical scanning method outlined above. Firstly, when the definition of a character discussed in Section 3.1. is applied stringently to the grid elements of the scanning method, an anomaly becomes apparent. The grid elements are illogical overdivisions of some higher order character, and at the same time, rather illogical conglomerationsof a lower order character. It appears therefore, that they fall between two'logical stools'. To expand the point further, consider the series of steps which seem appropriate for characters at present used in the study of wing pattern:

- Somatic outgrowth - shape and position of wing lamina.
- Pattern elements - shape, size and distribution over wing, behavioural sign stimuli.
- Epidermal cell position- position of cell on wing relative to morphogenetic boundary.
- Epidermal cell genome

Using these 'levels of analysis', it is obvious that the shape and mesh size of the grid squares appears as an arbitrary division of pattern elements and simultaneously, conglomerations of individual epidermal cells. Hence, the status of the grid squares used above as characters is somewhat enigmatic or arbitrary. They are obviously not characters as defined above, or in any general biological sense, and therefore are best referred to as pseudocharacters. It is difficult to associate any biological significance to them. The nature of these 'characters' will consequently have a considerable influence on the use of a classification based on them. For those who adhere to a strict biological interpretation of all characters (which does not necessarily imply a knowledge of their development or genetic specification), a classification based on pseudocharacters would

be rejected.

However, it may be argued that whether or not characters of the scanning method have any biological significance, the method is still valid for quantifying a pattern. If this is the object of a study, then the use of grid squares cannot be invalidated. To summarise, the main problem concerning the use of pseudocharacters derived from scanning seems to lie in why the patterns are being compared. If the relationships of the OTU's are known using data from an independent source, the scanning method would be acceptable, on logical grounds, if it allocated an unidentified individual satisfactorily to a known group, i.e., for purposes of identification. For the present study, where characters are being sought for classification, this method is unsuitable, but it may prove a fruitful line of future research, in association with optical scanners, for automated identification.

A problem of a more practical nature with the scanning method, is the manual coding of 420 grid squares, subsequent storage in the computer and comparison of 420 binary characters. The large number of 'characters' generated by this method also presents a problem relative to the number of OTU's which it is practical to compare. The ratio n/p, in which n = number of OTU's and p = number of characters, should ideally be increased as much as possible in any multivariate study. When the number of characters is much greater than the number of OTU's, the true relationships of the OTU's becomes increasingly more difficult to resolve accurately and is more susceptible to random effects.

In the light of these considerations, an alternative approach to scanning was developed.

8.3.4. Pattern Element Method

It is generally accepted that the structure and position of a given wing element is determined during morphogenesis. Therefore, when developing a technique for interpreting wing patterns of adult flies, information concerning the mechanisms of morphogenesis are likely to be of great use. Of particular significance are the hypotheses which relate growth and shape of a structure to segmental or other important boundaries. Hence, position is considered in relative, rather than absolute terms. Lawrence (1970), and Locke (1967) discuss these hypotheses most eloquently.

Using such hypotheses, together with the assumptions of the homology models discussed in the introduction, a method has been developed to dissect a wing pattern into a number of pattern elements. This method divided the wing into 'fields', each field representing the area of the wing that may be occupied by a single pattern element. The fields are seen as logically discrete units which comply with the levels of analysis proposed above. When a complex wing pattern is analysed by this method, the variation and development of a single element may be followed throughout a series of specimens. Furthermore, the expansion, contraction and changes in shape of pattern elements within their respective fields may easily be coded for inclusion in a broader study.

It is important to note that by virtue of their origin in carefully arranged morphological series, these fields facilitate clear expression of homology statements. This, as discussed earlier, is the advantage of such an approach over a purely descriptive method.

In an attempt to rationalise the somewhat intuitive process by which the fields were constructed, the following algorithm presented.

(i) Select series of specimens, as representative of the variation within the group as possible: absolute variation and not population.

(ii) Select phase, i.e., whether the development of dark pigment on a white ground will be followed or vice versa.

(iii) Select wing showing minimally developed pattern and note spots - this is only preliminary and will not isolate all pattern elements.

(iv) Search series, following the expansion of previously noted spots. If this contains all other dark areas arising in the series, go to vi, if not, go to v.

(v) Note additional spots detected in iv, i.e., spots which appear in other OTU's and not included in the provisional series at iii. Go to iv.

(vi) Select one spot and follow its expansion (and in some cases contraction) in relation to neighbouring spots, throughout the series.

(vii) Where any two neighbouring spots coalesce, draw a line corresponding to the interface. This is usually found found by noting the states of the spots in the OTU series

immediately prior to coalescing and drawing a boundary line along the median points between the two frontiers.

(viii) If all neighbours of selected pattern element have been investigated, go to ix.

(ix) If all pattern elements have been investigated, finish, if not, go to vi.

With the aid of this algorithm, the fields shown in Fig. 43 were constructed for the Culicoides pulicaris species group. In this figure, the starting point for pattern development within a field is shown as a spot. These were deduced by finding the epicentre of each pattern element when it was a minimum in each field. The general direction of pattern radiation from these centres is also indicated by means of arrows.

One of the first decisions which has to be made when analysing a wing pattern is whether a dark pattern is to be followed on a pale background or vice versa. In the case described here, pattern on a pale ground was followed. In Culicoides, the plesiomorphic condition is most likely a completely pale wing, as in the primitive subgenus Selfia and other genera of the Culicoidini. In C. pulicaris, no statistical difference was found between the spacing of the microtrichia in dark and pale patches of the wing (see Section 3.2. for details). It was therefore concluded that the pattern of the wing was due entirely to pigmentation.

In Fig. 43, it may be seen that the fields do not occupy the whole area of the wing. This is because in this species-group, pigmentation does not occur in all parts. The condition of these unmapped peripheral areas would undoubtedly be clarified by examining species closely related to the pulicaris complex, or even the whole genus Culicoides. This underlines the point that, although the model described above was based on one species group, it may be applicable to other members of the genus without great alteration.

One aspect of the pattern element model that appears somewhat surprising, is the extent to which the fields are independent of wing veins, i.e., pattern elements traverse wing cells and veins in an irregular fashion (Fig. 44). This feature is not so extraordinary when viewed in comparison with the parallel discovery of morphogenetic compartments in Drosophila wings. Lawrence & Morata (1976) have found that the wing may be divided into a series

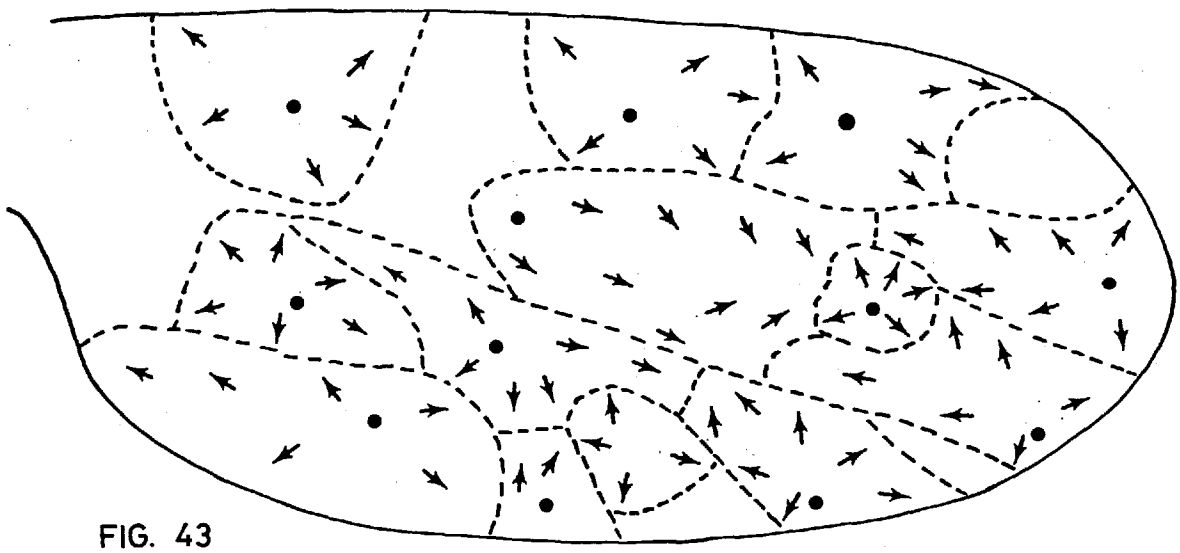


FIG. 43
RADIATION OF PATTERN ELEMENTS WITHIN THEIR RESPECTIVE FIELDS

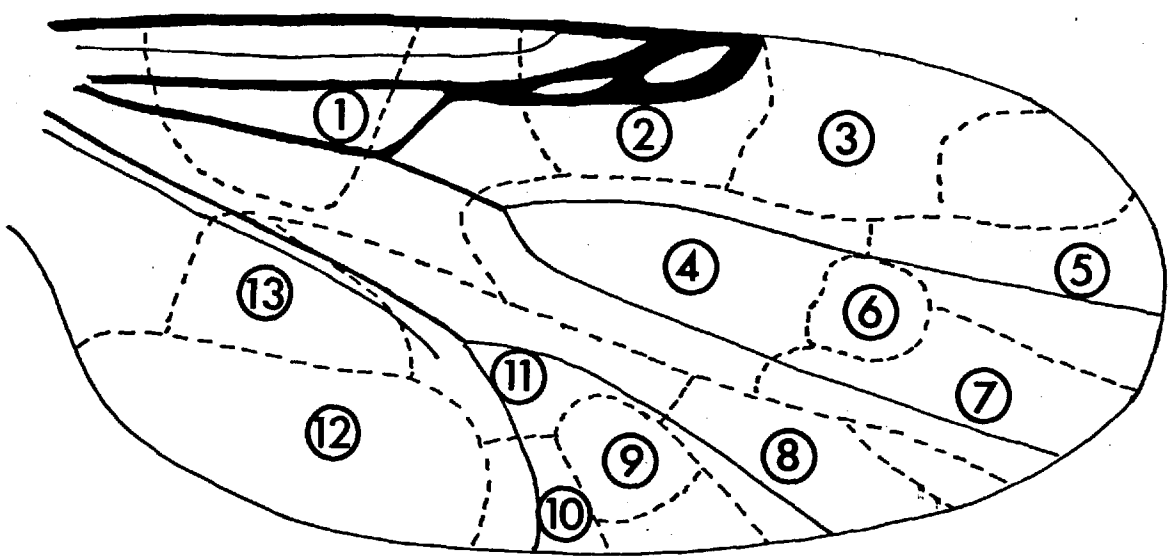


FIG. 44
DIVISION OF THE WING INTO 13 PATTERN ELEMENTS

of discrete compartments, based on their developmental background. Each compartment originates from a single cell or clone, which is specified genetically as an independent unit. Working with a mutant for wing macrotrichia, they found that the boundaries of the wing compartments ran longitudinally along the wing and were totally independent of venation.

In contrast to the prototype model advanced by Schwanwitsch, the model described here indicates the area of the wing in which pattern elements may develop, together with the modifications that these pattern elements may take. Schwanwitsch shows a ground plan to include all the basic elements that may be developed by one group of butterflies or other, but does not indicate their likely development.

In order that some of the properties and limitations of the pattern elements method could be examined, its influence on the classification of the 23 OTU's was investigated. For each of the thirteen fields of the wing, the variation the variation throughout the sample of 22 wings, although continuous, was coded in a series of states based on the behaviour of the pattern in Fig. 43. The series of states for each pattern element was divided into steps of approximately equal magnitude (Fig. 45). A clear plastic sheet on which the fields were marked, was laid over each of the wing drawings, and the state of each character recorded.

A similarity matrix was constructed, using Gower's general coefficient of similarity for 23 OTU's. After conversion to a distance matrix, analysis was carried out, as summarised in Table 18. The eigenvector associated with the largest eigenvalue absorbs 42% of the total variance, increasing to 57% by the second vector. The first three eigenvectors are plotted against one another in Fig. 46. The first plot (Fig. 46a) shows quite distinct groupings of the wings:

OTU's	17,18,22	(*fagineus*)
	7,8,9,10	(*pulicaris* & *punctatus*)
	3,6,13	(miscellaneous)
	2,4,5,11,12,14,16,19,20, 21	(*impunctatus* & *grisescens*)

The first axis described the degree of pigmentation of the wing. This is verified by Fig. 47, in which the coordinates for the first vector are plotted against the overall development of the wing pattern, measured by the sum of all the observed character

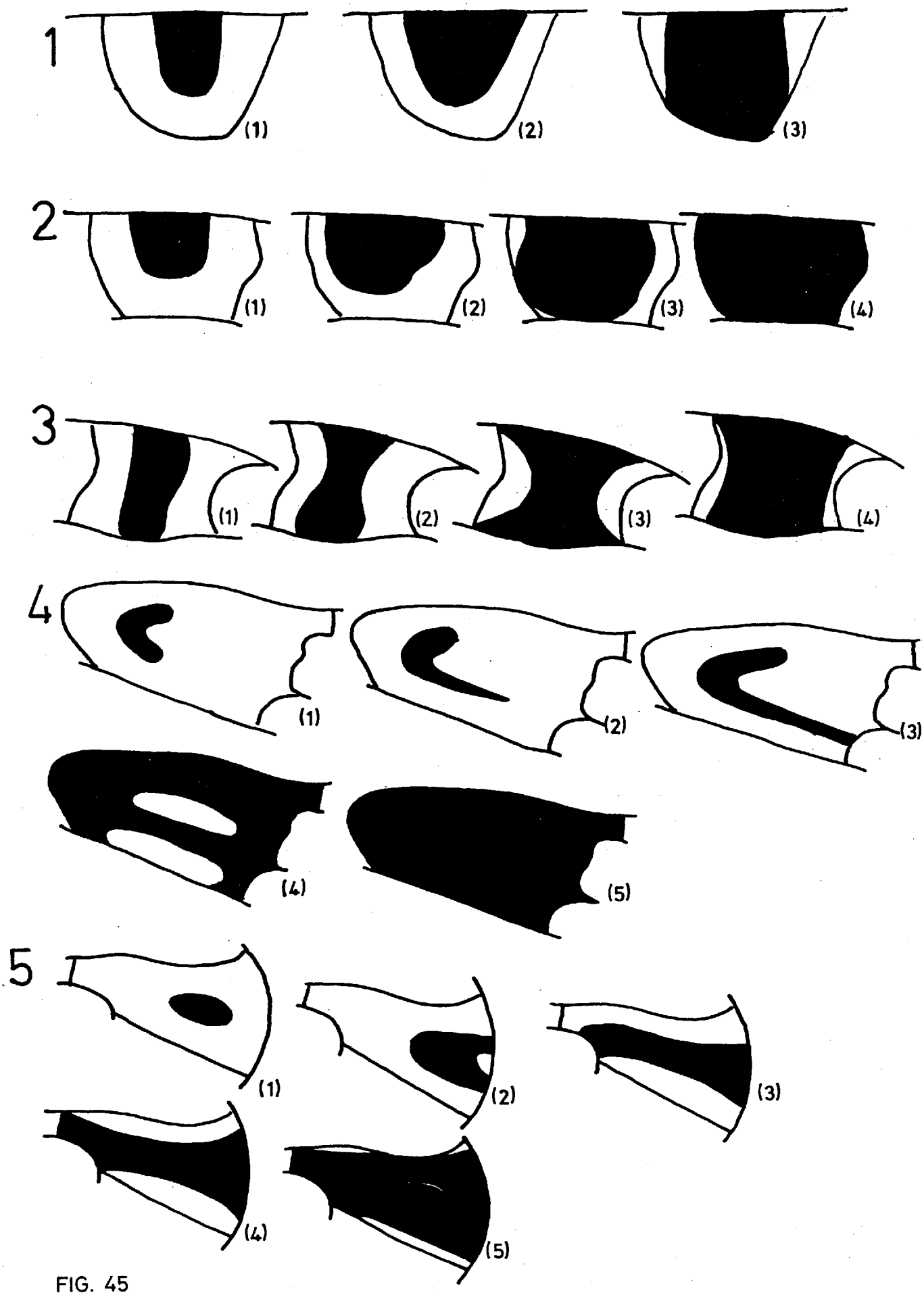


FIG. 45
CODING OF CHARACTER STATES FOR PATTERN ELEMENTS 1-5

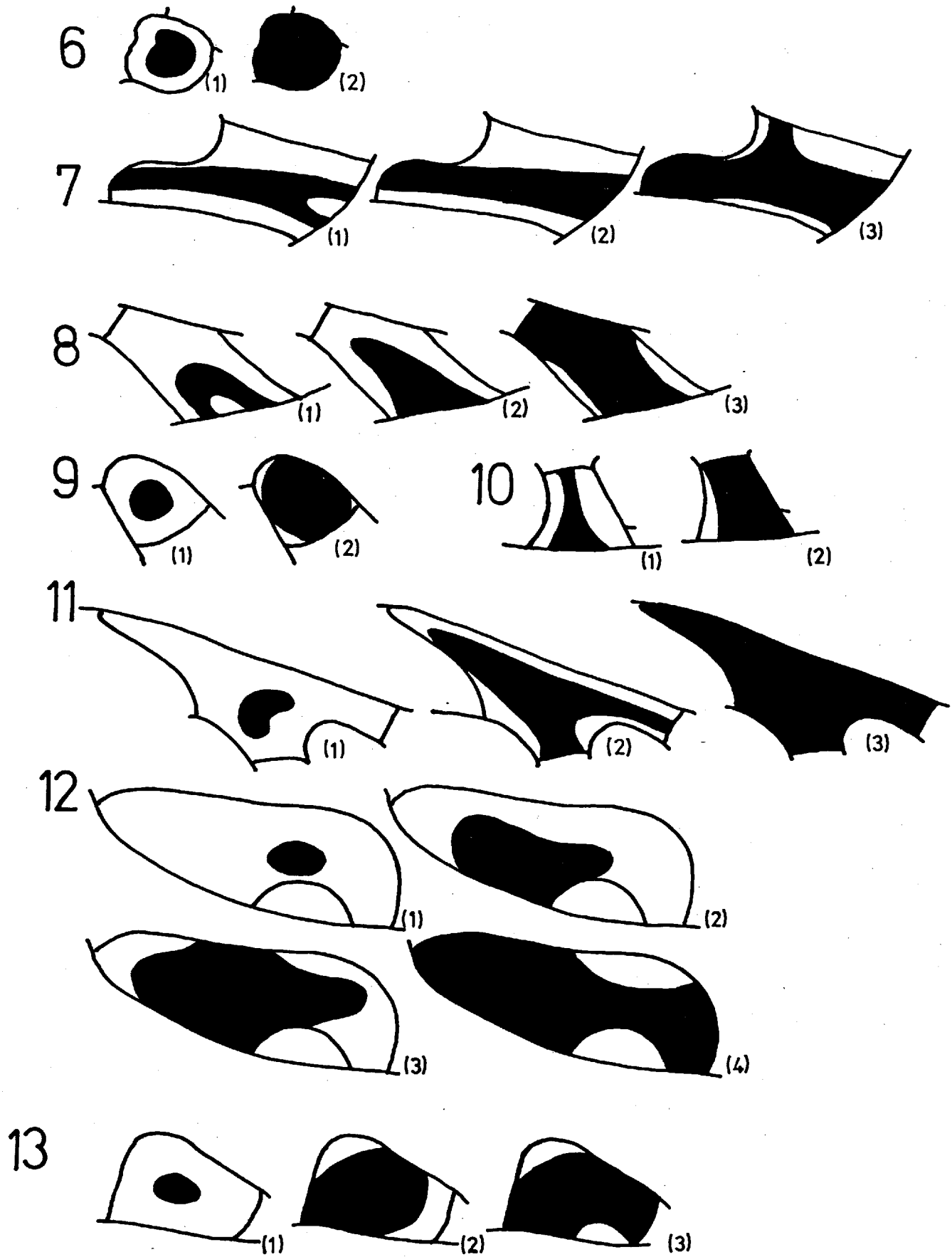


FIG. 45 CONTINUED
CODING OF CHARACTER STATES FOR PATTERN ELEMENTS 6-13

TABLE 18.

Summary of Principal Coordinate Analysis of 23 Wings Coded by the Pattern Element Method.

	1	2	3	4	5
Eigenvalue	3.142	1.112	0.651	0.417	0.357
Cumulative Percentage of Trace	42.29	57.32	66.14	71.79	76.63
OTU		Eigenvectors			
1	-0.295	-0.008	-0.225	-0.260	0.087
2	0.426	0.009	-0.125	0.032	-0.067
3	0.094	-0.082	0.379	-0.095	-0.189
4	0.315	-0.087	-0.081	0.134	-0.045
5	0.283	0.012	0.024	0.033	-0.024
6	0.102	-0.098	0.457	-0.090	0.265
7	-0.509	-0.160	-0.113	0.131	0.269
8	-0.532	-0.357	0.021	0.083	-0.162
9	-0.305	-0.461	0.126	0.134	-0.174
10	-0.575	-0.264	-0.111	-0.025	0.028
11	0.394	0.016	-0.121	0.034	0.045
12	0.383	0.016	-0.118	0.024	0.017
13	0.116	-0.078	-0.009	0.275	0.204
14	0.420	-0.007	0.093	-0.123	0.188
15	-0.418	0.010	-0.158	-0.295	0.004
16	0.349	-0.007	0.151	-0.104	-0.055
17	-0.426	0.284	-0.034	-0.069	-0.042
18	-0.289	0.461	0.111	0.118	-0.045
19	0.391	0.017	-0.165	0.058	-0.079
20	0.356	0.059	-0.202	0.059	-0.058
21	0.372	-0.091	-0.040	-0.197	-0.084
22	-0.359	0.355	0.029	0.024	-0.031
23	-0.289	0.461	0.111	0.119	-0.044

The OTU's may be identified from Table 15.

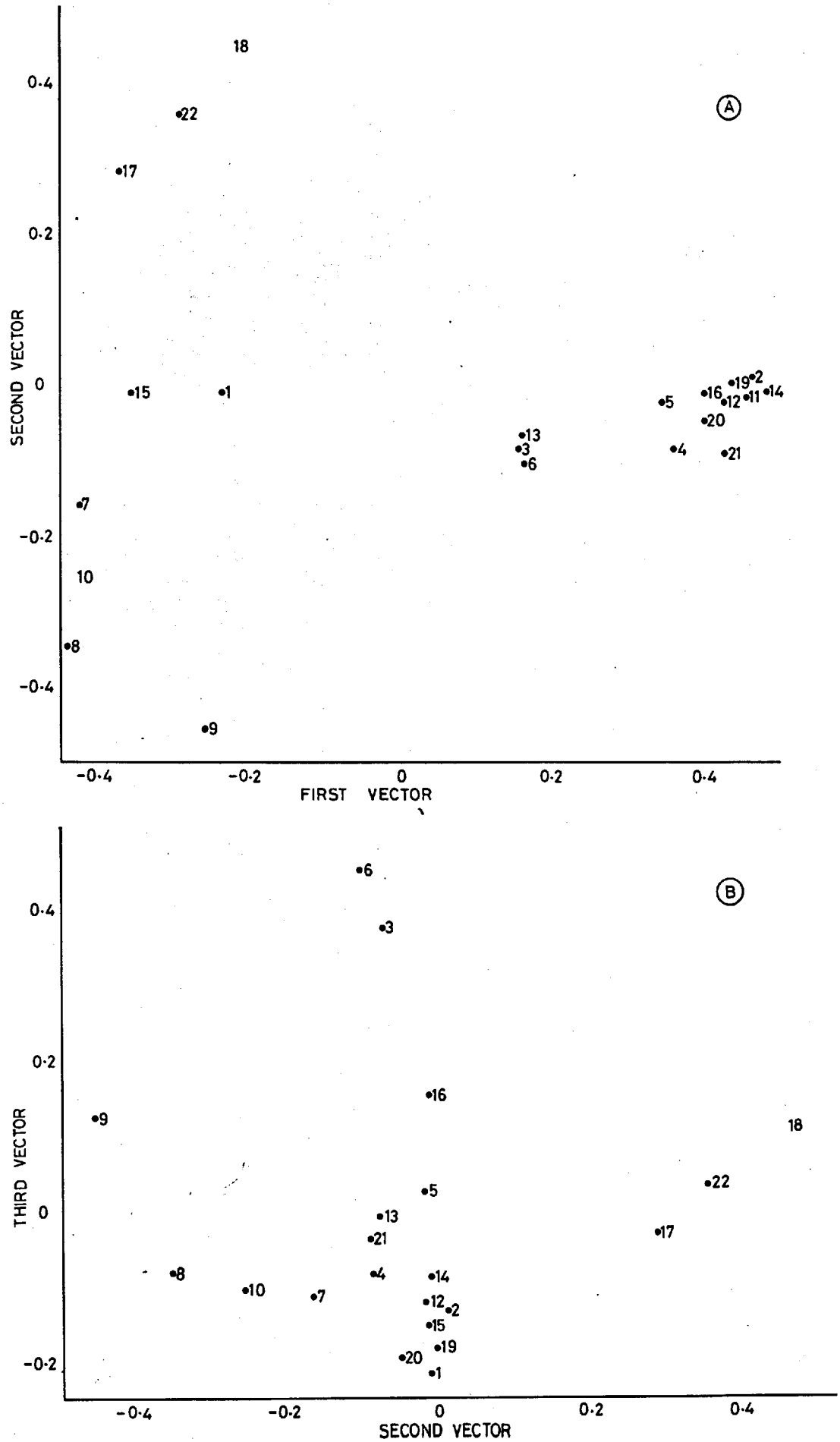


FIG. 46
PLOT OF EIGENVECTORS FROM PRINCIPAL COORDINATE ANALYSIS OF 22 WINGS CODED BY THE PATTERN ELEMENT METHOD

172

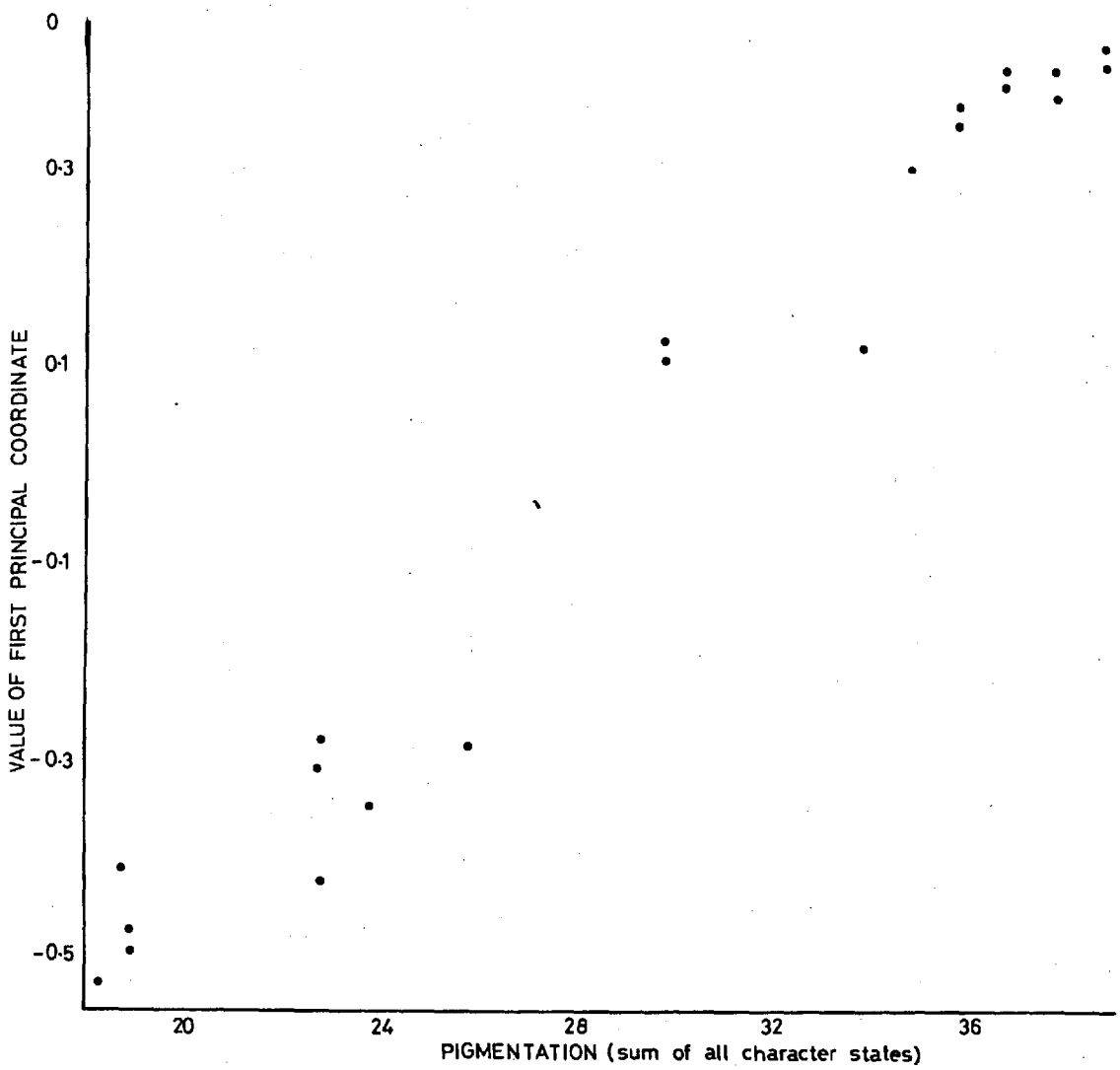


FIG. 47
ASSOCIATION OF GENERAL PIGMENTATION OF WING WITH FIRST EIGENVECTOR
(FROM PRINCIPAL COORDINATE ANALYSIS OF WINGS CODED BY PATTERN ELEMENT METHOD)

states for each wing. The characters influencing the second axis are more difficult to determine by interpolation from the relative position of the wings.

Since the wing pattern elements are treated as ordered multistate characters, it is possible to use a correlation coefficient to measure similarities between OTU's. From a correlation matrix an R-mode eigenanalysis (principal component analysis in part) may be carried out to determine the loadings corresponding to each character and hence the contribution that character makes to the axes (especially axes two and three). This only required eigenvectors to be extracted from a small matrix (of order 13 x 13). A summary of the analysis is given in Table 19.

Comparable percentages of the trace in the first five dimensions are to be found in the results of both the principal coordinate analysis and the eigenanalysis of the correlation matrix. Most of the characters receive a large negative loading in the first eigenvector, reinforcing the earlier conclusion that this axis is a measure of overall development of wing pattern (equivalent to size). In the second eigenvector, four characters are considerably larger than the others and have loadings with absolute values greater than 0.3. The four characters are wing pattern elements 3,6,9 and 10, and combined, they account for most of the variance in this axis. The third eigenvector has five characters with loadings greater than 0.3, they are: 5,7,9,11, and 13. The value of 0.3 was chosen in both cases as a convenient figure to separate the loadings, which fall into two fairly coherent groups in both eigenvectors. It is noticeable that wing pattern element 9 has the largest loading on the third eigenvector and the third largest loading on the second. This character describes the spot in the cubital cell and has been used as an important feature to separate species of the C. pulicaris complex.

8.3.5. Comparison of the Classifications Produced by the Two Coding Methods

The relationships of the wings coded by the scanning method and pattern element method are shown in Figs 40 and 46.

Visual inspection of the plots shows the grouping based on pattern element data to be more discrete than that based on the scanning data. The relative positions of the taxa do not differ much

TABLE 19.

Summary of Principal Component Analysis of 23 wings Coded by the Pattern Element Method. Details of the First Five Eigenvectors.

	1	2	3	4	5
Eigenvalue	6.351	2.269	1.498	0.783	0.599
Cumulative Percentage of Trace	48.86	66.31	77.84	83.86	88.48
Variable (Wing pattern element)	Eigenvectors				
1	-0.343	0.167	-0.147	-0.230	0.152
2	-0.315	0.148	0.128	-0.413	0.049
3	-0.299	0.341	0.054	-0.216	0.003
4	-0.307	0.119	0.262	-0.190	-0.114
5	-0.326	-0.019	-0.341	0.126	0.259
6	-0.075	-0.615	0.078	0.009	-0.091
7	-0.294	-0.104	-0.398	0.276	0.041
8	-0.327	-0.194	0.061	0.241	-0.006
9	0.136	0.349	0.420	0.321	0.558
10	0.004	-0.498	0.264	-0.469	0.465
11	-0.255	-0.005	0.466	0.115	-0.549
12	-0.370	-0.037	-0.112	0.060	0.119
13	-0.268	-0.140	0.357	0.449	0.195

between the two analyses. There are no major discrepancies in the placement of individual specimens. The wings of C. fagineus, for example, are more distinct when coded by the pattern element method, but the arrangement of specimens within this taxon do not differ significantly between classifications.

The factors influencing the distribution of OTU's on the plots of eigenvectors have already been discussed (see p. 167 and p. 173).

The rather subjective evaluation of the similarity between classifications by visual inspection of the plots becomes more objective by the use of matrix correlation sensu Sneath & Sokal (1973), particularly that termed r_{s1} r_{s2}. By this technique, the agreement of two or more classifications may be determined by calculating the correlation coefficient between the similarity matrices, on which the two classifications were based. These may be summarised briefly for this study as:

Scanning data v. pattern element data $r = 0.202$

When the weighting of the scanning data is considered, the correlations are:

Scanning data v. weighted scanning data $r = 0.979$

Pattern element data v. weighted scanning data $r = 0.199$

As noted above, from visual inspection of principal coordinate plots, the weighting of the scan data has not greatly altered or improved the classification of OTU's. This has been confirmed by matrix correlation.

The extent of the discordance between classifications produced by the two coding methods is unexpected, and therefore rather interesting , and in need of further discussion. Two hypotheses may be advanced to explain the apparent lack of congruence suggested by the correlation matrix. Firstly, there is a real and substantial difference between classifications, or secondly, the matrix correlation technique is an inefficient method for measuring the concordance of two classifications, and has therefore given misleading results.

The two coding methods contrast markedly in the extent of character redundancy, which may be the cause of a real difference between classifications. The pattern element methods shows very little redundancy, as most pattern elements, or more precisely, the limits of the fields in which the pattern elements vary, have been defined by the observed variation between OTU's. In the

scanning method, characters (= pseudocharacters) have been created irrespective of whether they show any variation between wings. They measure all that there is and do not distinguish between varying and non-varying characters. The difference between coding methods may account, in part, for the ease with which particular characters can be associated with each of the eigenvectors. In this respect, a more prominent factor is the number of characters involved (420 in scanning, compared to only 13 pattern elements) and the number of states of each of the characters (binary or multistate). All of the effects contribute to the way in which either coding method measures the shape and position of wing pigmentation.

As noted above, there is no substantial difference between the classifications in terms of the nearest neighbours of the OTU's. In the classification based on the scanning data, the OTU's are rather evenly distributed, whereas the pattern element method of coding produces a classification of tight and discrete clusters. For comparison of ordinations, it might be more appropriate to employ methods involving the rotation of the axes, minimising the sums of squares of the distances between corresponding OTU's. Although these methods appear to reproduce the visual comparisons of ordination plots more clearly than matrix correlations, Davies & Boratynski (1979) made a detailed comparison of these two methods of comparing classifications and found they agreed well.

To summarise these points; the lack of concordance of the two classifications, implied by matrix correlation, is probably the result of a real difference in the relationships of the OTU's, in addition to the shortcomings of the technique of matrix correlation.

Also of note is the extent to which each of the two coding methods used here to quantify wing pattern, measures shape and location of pigmentation. The pattern element method is the more labile with respect to location. It allows for the displacement of a spot within a given area (with a degree of tolerance) without recording the displacement. In contrast, the scanning method registers any slight variation in the location of a spot as a change in the character state for a number of characters. For a small number, it has been shown that this does not unduly affect the position of an individual within an ordination. The

pattern element method however, does not critically quantify shape beyond the sequential coding of an observed set of character states. The scanning method would seem to measure the location of pigmentation more effectively than the shape of any pigmented area.

In conclusion, the pattern element method is considered the most useful of the two coding methods in classification for the following reasons: it operates on logically acceptable characters; uses a relatively small number of characters; reduces character redundancy; and facilitates easier and faster coding. Some of the disadvantages of scanning may be overcome in the future, with the introduction of automatic methods of data capture by scanning images. Under these circumstances, it would be of more use in identification, rather than classification.

8.3.6. Wing Pattern Classification of the pulicaris Complex

Wing pattern has been used extensively in the taxonomy of the genus Culicoides, and the pulicaris group is no exception. Indeed, many of the taxa described in the pulicaris complex were first recognised by differences in wing pattern.
The section above compares and contrasts two methods for coding wing pattern. It is the objective of this section to take the more 'successful' of the two methods and further test its effectiveness by coding the wings of 84 specimens and observing the resultant classification. The 84 OTU's represent a wide range of the variation in wing pattern, within the pulicaris complex.

There are a number of points which complicate this experiment, the main one being that it is not possible to separate completely the limitations of a coding method, from the predictions of the non-specificity hypothesis. This hypothesis was proposed by Sokal & Sneath (1963) and predicted that a classification based on characters from one organ (e.g., wing pattern) is the same as that based on all organs or a differing organ. The hypothesis was originally couched in terms of classes of genes. If the hypothesis was completely true, then a classification based on wing pattern alone would be the same as that based on a number of characters. Under these circumstances, any difference in the classification would reflect the merits, or otherwise, of the wing coding method.

Unfortunately, the non-specificity hypothesis is only

partially correct (Rohlf, 1962), particularly at low taxonomic levels (see Section 9.1). Therefore it is not possible to evaluate thoroughly the pattern element method of coding wings, only to investigate its effects on grouping a large sample of wings.

Details of the 84 OTU's are given in Section 9 (Table 27), and were provisionally identified using a range of characters. All of the seven nominal taxa were represented, as well as a morphologically distinct sample from Japan (Sp. A). The condition of the 84 wings was coded by the pattern element method, using the character states outlined above. Principal component analysis was carried out on the data matrix to show the relationships of the wings. The first three eigenvalues and eigenvectors are given in Table 20, and a plot of the first two principal components in Fig. 48. The first two axes account for 65% of the total variance increasing to 74% by the third axis. The elements (loadings) of the first eigenvector are mostly in the range -0.33 to -0.21. Only one character - pattern element 9 - has a value of 0.013. This pattern element describes the spot in the cubital cell and has the highest loading on the second eigenvector. This is clearly shown in Fig. 49, in which a sample of wings have been superimposed onto the components analysis of all 84 wings (Fig.48).

OTU's are distributed along the first principal component according to an overall estimate of pigmentation in the wings. This is comparable to the'overall size' factor frequently associated with the first axis of a principal component analysis.

Superficial examination of the principal components plot (Fig. 48) shows that the separation of some species is rather indistinct, a situation for which the complex is renowned and parallels that found with quantitative characters. Closer inspection reveals that the straggly distribution of some species (e.g., punctatus) reflects differences between populations. For example, the eight OTU's of punctatus in the lower left of the diagram are quite separate from the other punctatus and overlap with a cluster of newsteadi. These specimens represent a sample from Norway in which, using a variety of characters, it was very difficult to identify them as either punctatus or newsteadi.

A similar separation of species into two groups is found in C. delta. Those in the centre of the diagram represent a sample of specimens from Kent whilst the remainder (top left) are from a

TABLE 20.

Summary of Principal Component Analysis of 84 Wings Coded by the Pattern Element Method. Details of the First Five Eigenvectors.

	1	2	3	4	5
Eigenvalue	7.082	1.352	1.087	0.712	0.555
Cumulative Percentage of Trace	54.48	64.88	73.24	78.72	82.98
Variable (Wing pattern element)		Eigenvectors			
1	-0.276	0.269	-0.156	0.198	0.295
2	-0.283	0.098	-0.337	-0.078	-0.234
3	-0.253	0.257	-0.451	-0.210	0.209
4	-0.311	-0.150	-0.175	-0.219	-0.001
5	-0.330	0.101	0.107	0.279	-0.023
6	-0.214	-0.414	0.477	-0.064	0.336
7	-0.336	0.115	0.128	0.257	0.095
8	-0.296	0.062	0.209	0.397	0.195
9	0.013	-0.595	-0.477	0.588	-0.125
10	-0.258	0.069	0.305	0.054	-0.743
11	-0.245	-0.514	-0.026	-0.388	0.073
12	-0.333	0.007	0.177	-0.153	0.088
13	-0.296	-0.080	-0.068	-0.182	-0.268

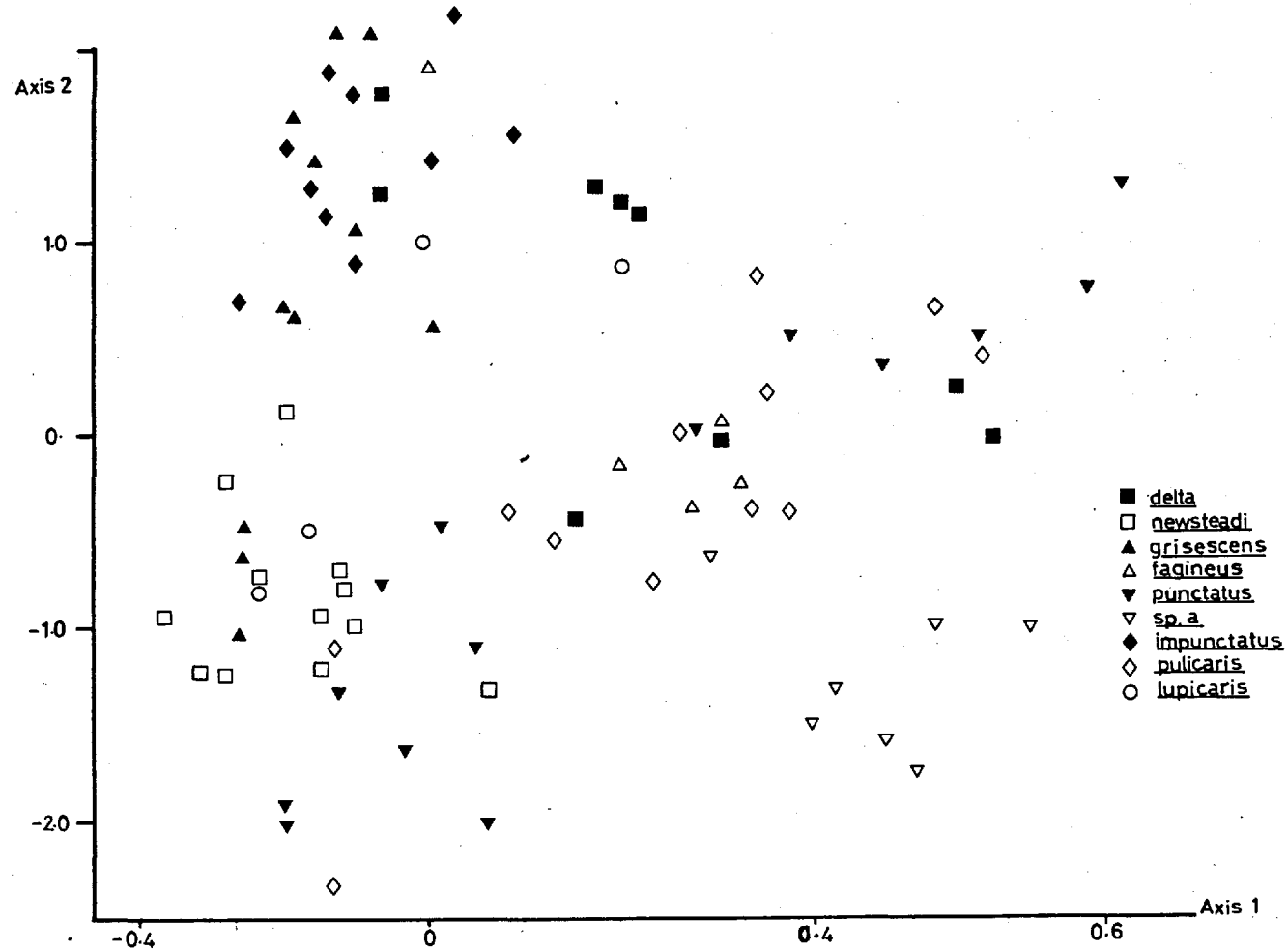


FIG. 48
PRINCIPAL COMPONENTS PLOT OF 84 WINGS CODED BY THE PATTERN-ELEMENT METHOD

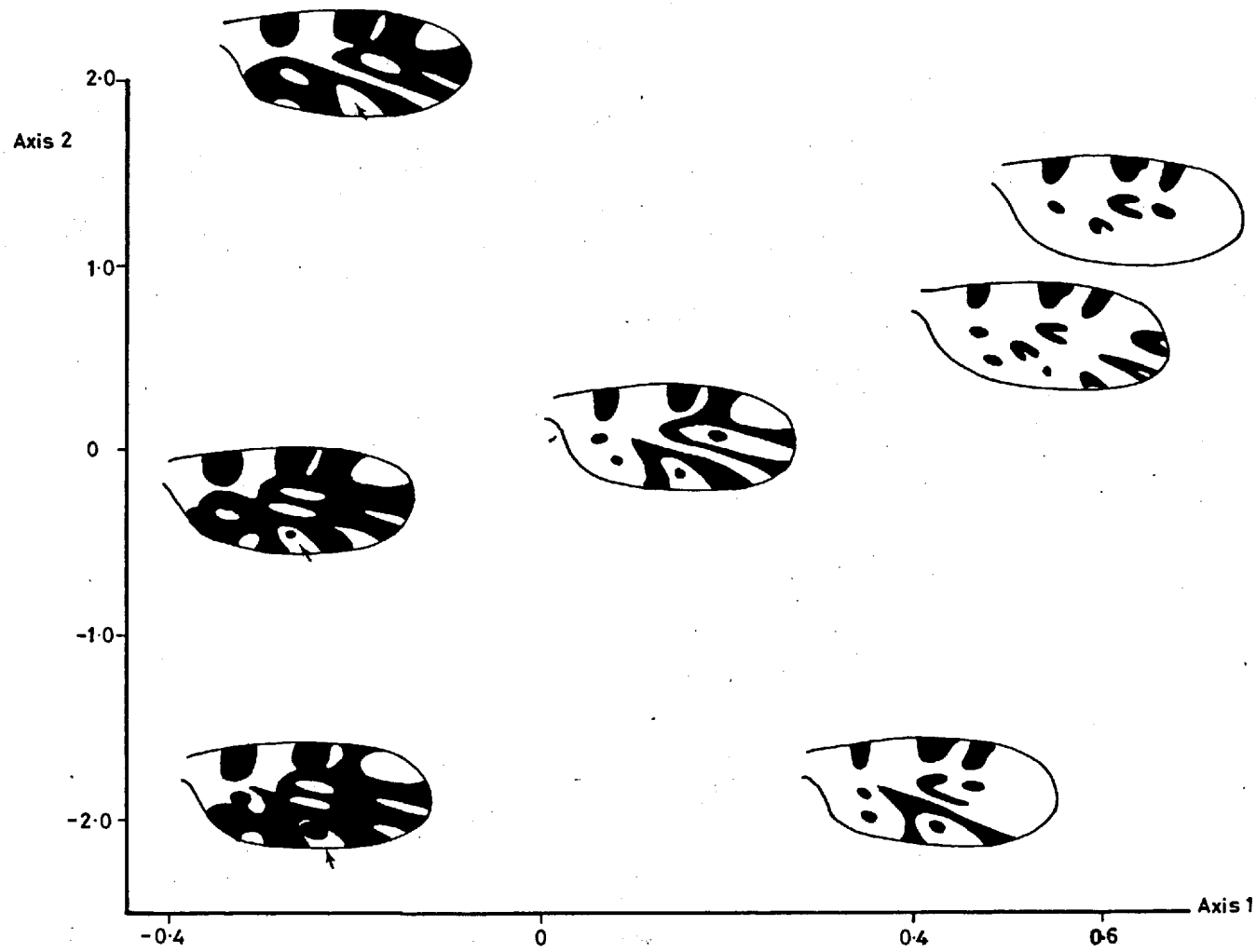


FIG. 49
PRINCIPAL COMPONENTS PLOT OF ANALYSIS BASED ON 84 WINGS.
REPRESENTATIVE SAMPLE PLOTTED TO SHOW PATTERN ELEMENTS WEIGHTED
ON FIRST TWO AXES.

181

number of localities ranging from Scotland to southern England. Four specimens of C. fagineus are tightly grouped in the centre of the diagram and are composed of specimens from localities as far apart as Israel and Hampshire (type series). However, one specimen from Surrey is placed well away from this group, at the top of the plot, by virtue of its lower character states (particularly pattern elements 9,11 and 6). The identity of this specimen as C. fagineus was confirmed by the presence of cibarial teeth.

The overall distribution of C. pulicaris and punctatus is very similar, paralleling that found previously with quantitative characters. Some species, e.g., newsteadi, impunctatus and sp. A are well grouped on both the first and second axes and thus form a coherent cluster of points. The cluster of sp. A is distinct from all other taxa. Specimens of C. grisescens are tightly grouped on the first axis, showing they have the same overall pattern, but vary on the second axis, reflecting the difference in expression of the spot in the cubital cell (pattern element 9). This species overlaps considerably with impunctatus, as the wings of the two species are very much alike. The principal differences between the species lie in other quantitative characters. It is interesting to note that, like gisescens, the spot in the cubital cell of impunctatus is variable (considerably more so than is conceded in the literature). In a sample of impunctatus from Surrey, almost 10% lacked a spot in the cubital cell.

In conclusion, the method of coding wing patterns by pattern elements appears to be quite efficient. However, the extent of its efficiency cannot be fully determined as, like quantitative characters, wing pattern does not provide an absolute means of distinguishing between species. The ability of the pattern element method to describe small differences between samples suggests that, when used in conjuction with quantitative characters, it will prove most useful in taxonomic research.

Section 9. NUMERICAL CLASSIFICATION OF THE C. PULICARIS COMPLEX

9.1. INTRODUCTION

Previous sections have been concerned with studies on individual characters, either describing them or studying their variability. The results of these studies may now be put to taxonomic use.

The main objective of this section is to produce a classification of the taxa in the pulicaris complex. Two secondary objectives were:

(i) To determine whether the large number of characters used is necessary for an effective classification and, if not, which may be discarded to preserve or enhance the classification.

(ii) To determine whether the recognised 'species' are homogeneous.

Only when a reliable classification has been produced (including the definition of taxa), can discrimination between the taxa proceed. Consequently, this section has two main parts. The first part, sections 9.2 - 9.5, concerns the elimination of variables and the generation of several alternative classifications. The second part, section 9.6, discusses these classifications, with particular reference to the homogeneity of the taxa. It establishes which taxa will be used in Section 10 for discrimination.

When attempting to classify several taxa as well as reduce the number of variables, the first major problem encountered in the experimental design of such a study, is the balance between using subjective and objective methods. The emphasis given to these aspects will inevitably govern the experimental procedure and techniques used. For example, one strategy would be to group the specimens into taxa on subjective grounds and then investigate the minimum number of characters necessary to either classify or discriminate between them. This approach moves from subjectivity to objectivity. An alternative scheme would be to establish which taxa are reliable, by treating all the specimens as individuals, and then using numerical methods, to group them. It is not essential to define groups before attempts are made to eliminate redundant variables. The effectiveness of any subset of variables may be gauged by the arrangement of specimens, rather than taxa. This

second approach is more protracted and requires considerably more computation. Although certainly more objective than the first approach, subjective decisions may still have to be made on which individuals form a recognisable group. After consideration of these two designs, the second was adopted, because it is procedurally more rigorous. Therefore, throughout this section, individual specimens are treated as OTU's. This obviously limited the number of specimens which could be used in practice, as the use of sample means would have prejudged which taxa were homogeneous.

It is not until the final part of this section that the homogeneity of the taxa is discussed and taxonomic decisions are made. It is mainly concerned with establishing groups (taxa) and the following section attempts to discriminate between them.

Briefly, the procedure adopted is as follows:-
The first stage tests whether it is practical to reduce the number of variables by simply selecting them from only one region of the body. Essentially, this is a test of the non-specificity hypothesis. During the course of the experiment, it transpired that although the data had been standardised, general body size dominated the analysis, possibly obscuring any interspecific relationships. Size was shown to be taxonomically unreliable in Section 7,p.101 and therefore attempts were made to reduce its influence and, it is hoped, to concentrate on interspecific differences in 'shape'. Following the study of size characters, more direct procedures to eliminate variables were attempted. The first used mathematical criteria for selecting important variables, and the second used 'subjective' or intuitive criteria. The mathematical methods (ostensibly more objective) selected the best subset of variables, by either looking at the relationships between the variables themselves, or by selecting those variables associated with the largest loadings in a principal component analysis. The subjective methods selected the 'best' subset of variables by trial and error (or intuition). The final part of this section reviews the several classifications generated in the experiments to eliminate variables, and then summarises and defines the taxa. These taxa will be used in Section 10..

The appropriate number of variables to be used in a multivariate study is difficult to establish, and is often decided on grounds that are only partly statistical (Beale et al., 1967).

Neither Beale et al., nor Jolliffe (1972, 1973), who both made detailed studies on the methods for discarding variables in principal component analysis, were able to suggest any objective criteria. Sneath & Sokal (1973) believe that large numbers of variables should be employed in numerical taxonomy. They suggest that "at least 60 characters should be used", although, they also concede that there is no practical or theoretical evidence to support this figure. In contrast to these views, Jolliffe (1972) maintains that often, many more characters are used than are necessary, and many of these extra variables may be removed without any significant change in the results. Some characters are present which complicate the data (by adding 'noise') and yet do not give any extra information. Furthermore, time and money are saved if some of the variables are discarded, computing time is reduced and in further analyses, fewer variables need to be measured. There is also a conceptual simplification of what may otherwise become a problem of daunting magnitude.

The results of R-mode techniques such as principal component analysis become more reliable when the ratio $\frac{n}{p}$ (n = number of OTU's, p = number of variates) is maximised. If the number of variables is reduced without any adverse affect, more confidence may be placed in the results. Reducing the number of variables in a principal component analysis while still classifying, is an attempt to maximise this ratio. In discrimination analysis, practical considerations are often more important reasons for reducing the number of variables, to ensure that the minimum number of characters have to be observed, for the identification of each specimen. A further advantage of using as small a number of characters is that the new axes are simpler to interpret in a biological sense, and the important factors are easily recognised.

Each character reduction technique produced a subset of variables and these were used in a principal component analysis of a sample of OTU's. For each of these new classifications, it was important to distinguish between the evaluation of the method itself, and the taxonomic merit of the classification. When testing the value of the reduction method, the arrangement of the specimens based on the reduced set of variables was compared with that based on the complete set. It is possible that the reduced set produces a classification which is taxonomically more acceptable than one

using all the variables. Such a result would contradict Sneath & Sokal's (1973) argument that the numbers of variables used should be as high as possible, i.e., the new set does not mimic the classification based on a large number of variables, but improves upon it in a taxonomic sense.

Davies & Boratynski's (1979) approach to this difficult point was to take the classification based on the entire character set as the acceptable reference. The divergence of classification based on reduced subsets from this reference was calculated using the rotational fit method of Gower (1971b). Subsets which showed the least divergence were regarded as most acceptable. The reasoning behind this approach was that the most effective classifications were stable in respect of added characters - if 25 characters yield a classification that is little changed by adding another 76 characters (as they found), it is surely stable and this ought, perhaps, to be a taxonomic virtue! (Davies, personal communication).

9.2. SELECTING CHARACTERS FROM A SINGLE BODY REGION

Several methods are available for reducing the number of variables in a multivariate analysis. One of the simplest is to select characters from one region of the body.

Many multivariate studies have been carried out using characters from a fairly restricted part of the organism under investigation. In the study of insects, the measurement of wings has often been used, e.g., in Bombus (Dupraw, 1965a), parasitic bees (Plowright & Stephen, 1973) blowflies (Brown & Shipp, 1977) and fleshflies (Brown & Shipp, 1978). In mammals, skull measurements have been often used, e.g., Rees (1969) and Rostron (1972). Studies using an anatomically restricted set of characters rely to a considerable extent on the non-specificity hypothesis. This was first proposed by Sokal & Sneath (1963) and states "there are no distinct large classes of genes affecting exclusively one class of characters such as morphological, physiological or ethological characters, or affecting special regions of the organism". If this hypothesis was correct, then obtaining a disproportionately large number of characters from one body region, or of one special kind, would not restrict the information to one class of genes, and would

produce an acceptable classification as one based on many different types of characters.Furthermore, there would be no a priori grounds for favouring one character over another (Sneath & Sokal, 1973). In his review of the assumptions and applications of the hypothesis, Farris (1971) indicates that congruent character sets do not necessarily support the non-specificity hypothesis, but rather that different character-sets of the same group of OTU's should give the same taxonomic structure. Although Lidicker (1973) studied only one small structure (the penis) of New Guinea rodents, he maintained that the 65 characters used were not as restricted as might first be supposed, because several different tissues were involved.

To test whether choosing characters from one body region is an acceptable method for reducing the number of variables in a multivariate study, the following experiment was carried out:

A sample of 53 specimens, representing seven taxa of the C. pulicaris group, were classified first using 13 wing pattern characters, and secondly, using a total of 72 morphological characters distributed over the body. The two classifications were then compared. Any significant difference between them would throw doubt on the prediction of the non-specificity hypothesis, and suggest that this method of reducing variables was impractical (for these data at least).

The 13 wing pattern characters have been described in detail in Section 8,p.167, and the primary data matrix is given in the appendix (p.364). A complete list of the 72 variables are given in Table 21 , and the primary data matrix in the appendix (p.364), where the OTU's used in this study are indicated by an asterisk*.

Both single linkage cluster analysis and principal coordinate analysis were carried out on the data. The latter method was chosen in preference to principal component analysis because it required eigenvalues and eigenvectors to be extracted from a smaller association matrix (53 x 53 rather than 72 x 72), thus requiring less computational effort. Furthermore, details of the eigenvector elements, which may only be obtained directly in principal component analysis, were not required because only the general grouping of the OTU's was of interest.

A plot of the first two principal axes for the 'wing pattern data' is given in Fig. 50 and for the complete set of variables in Fig.51 .

TABLE 21. Description and Code Numbers of 72 Variables

Code Number	
1	Contiguity of eyes
2	Proportional length of antennal segment iii
3	" iv
4	" v
5	" vi
6	" vii
7	" viii
8	" ix
9	" x
10	" xi
11	" xii
12	" xiii
13	" xiv
14	" xv
15	Number of sensilla on antennal segment iii
16	" xi
17	" xii
18	" xiii
19	" xiv
20	" xv
21	Length of palp segment i+ii
22	" iii
23	" iv
24	" v
25	Width of palp segment iii
26	Head length
27	Proboscis length
28	Cibarium length
29	Pharynx length
30	Number of maxillary teeth
31	Number of mandibular teeth
32	Length of wing
33	Width of wing
34	Length of costa
35	Wing pattern element 1
36	" 2

TABLE 21 contin...

Code Number	
37	Wing pattern element 3
38	" 4
39	" 5
40	" 6
41	" 7
42	" 8
43	" 9
44	" 10
45	" 11
46	" 12
47	" 13
48	Fore leg : length of femur
49	tibia
50	tarsus i
51	tarsus ii
52	tarsus iii
53	Mid leg : length of femur
54	tibia
55	tarsus i
56	tarsus ii
57	tarsus iii
58	Hind leg : length of femur
59	tibia
60	tarsus i
61	tarsus ii
62	tarsus iii
63	Number of setae on hind tibial comb
64	Presence or absence of cibarial teeth
65	Antennal ratio
66	Palp ratio
67	Head length/ proboscis length
68	Palp length/ proboscis length
69	Antennal length/ proboscis length
70	Cibarium length/ pharynx length
71	Number of mandibular teeth/ maxillary teeth
72	Costal ratio

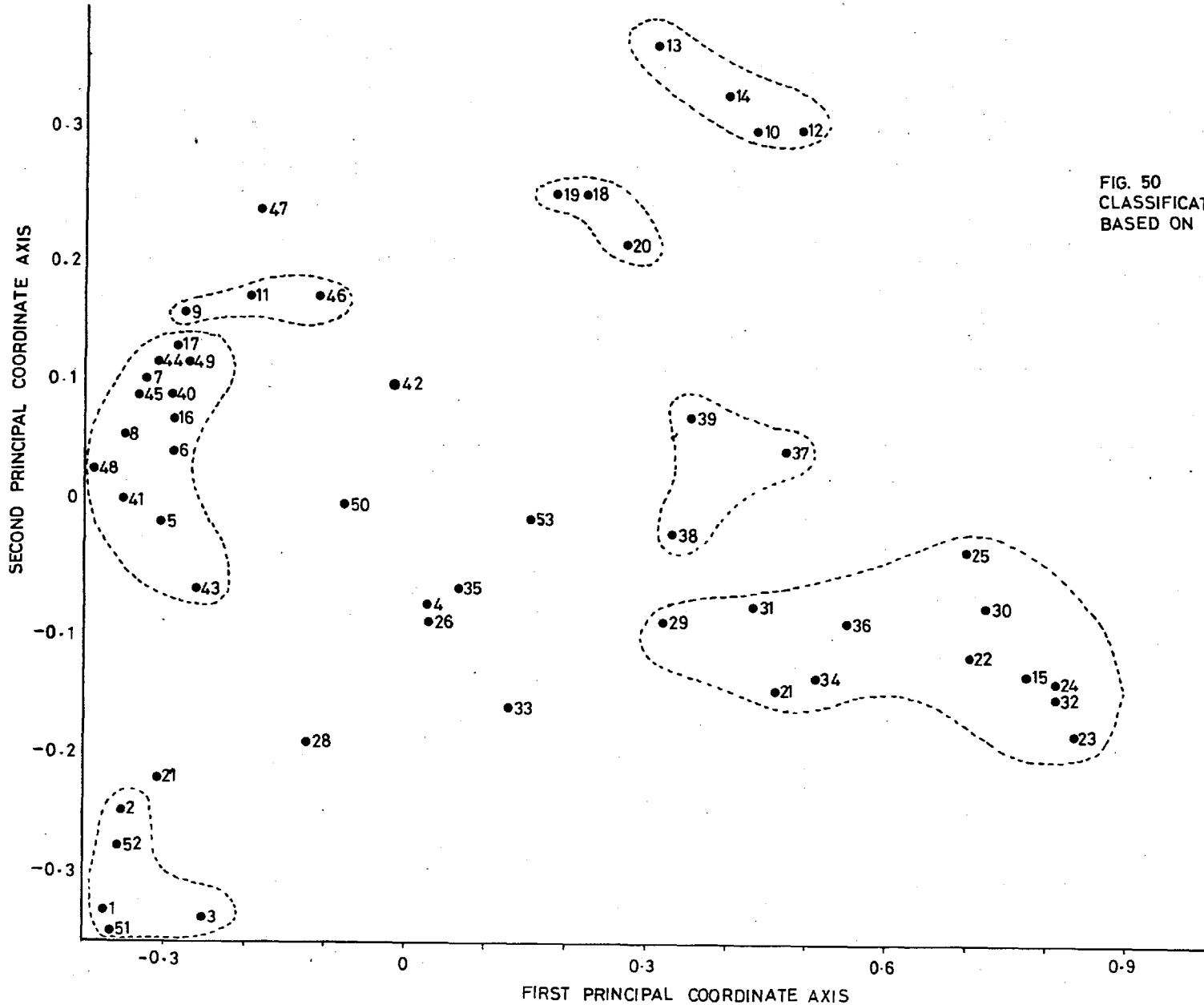


FIG. 50
CLASSIFICATION OF 53 SPECIMENS
BASED ON WING PATTERN

Results of the principal coordinate analysis using 72 variables

The percentage of total variation described in the first five dimensions is given in Table 22.

Only 19% of the total variance is described in the first dimension and the second dimension adds only a further 12%. In comparison, some 50% of the variance is described in the first two dimensions for the analysis based on wing patterns. This would suggest that intraspecific variation in the comparatively few wing pattern characters may be summarised more effectively than that for all 72 characters. Comparison of classifications using the percentage variance described by each axis has to be regarded with caution, because the total number of dimensions produced by each analysis is a function of the number of variables or characters (number of dimensions = min. ((n-1),c) where n is the number of OTU's and c is the number of characters). The wing pattern classification requires only 13 axes to describe it fully, whereas the other needs 52. The proportions on the first few axes reflect this difference, although they may also be affected by the choice of characters. In addition, the experiment seeks to establish whether an anatomically restricted subset of characters are taxonomically more useful, or as useful, as a large set of characters. To test this, the arrangement of specimens, representing different species, is of more immediate importance. The specific identification of specimens is provisional at this stage of the study, homogeneity of the species is studied later. At present, only broad trends are being investigated.

Visual inspection of the two ordination plots shows considerable difference in the grouping of the OTU's.

In Fig. 51, four basic groups may be recognised:

1. OTU's 1,2,3,4,26,27,28,40-48
2. OTU's 5,6,7,8,9,51,52
3. OTU's 11,16,17,18,19,20,29,50,53
4. OTU's 10,13,15,12,14,20,21,22,23,24,30-37

The first group comprises specimens of C. newsteadi and C. impunctatus. These are the two smallest species in the complex. Three small specimens of punctatus from Norway (26,27,28) are situated at the periphery of this group. Specimens of C. grisescens

TABLE 22.

Table to Show Percentage Variance Described in First Five Dimensions of a Principal Coordinate Analysis, of Two Sets of Data

		1	2	3	4	5
Data based on 72 variables	eigenvalue	2.37	1.50	0.75	0.62	0.48
	% variance	19.28	12.20	6.10	5.04	3.91
Data based on 13 wing pattern variables	eigenvalue	6.64	1.59	1.26	0.93	0.71
	% variance	40.51	9.70	7.68	5.67	4.33

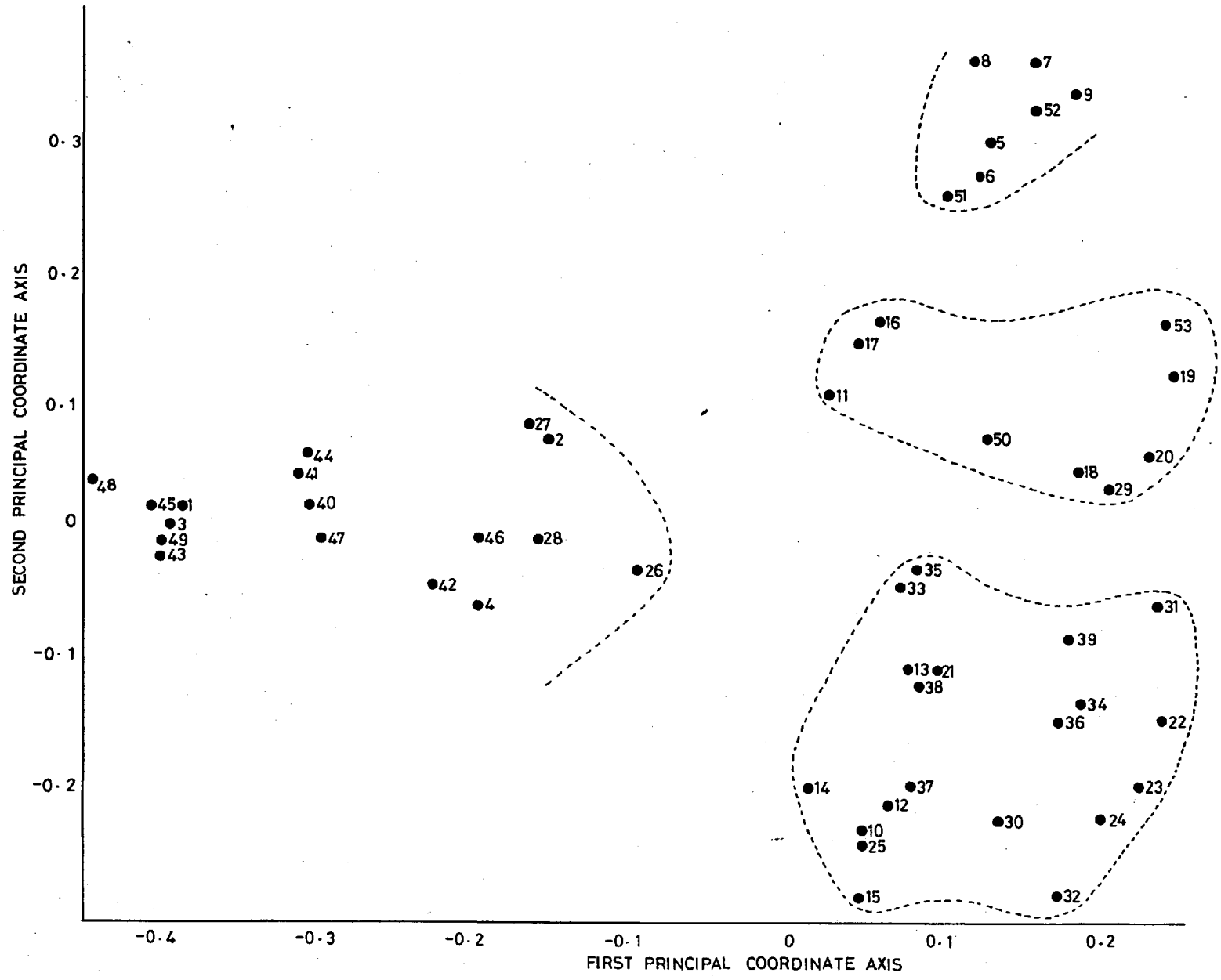


FIG. 51
CLASSIFICATION OF 53 SPECIMENS
BASED ON 72 VARIABLES

and C. lupicaris (in part) form the second group. The third group comprises specimens of C. delta, C. lupicaris (in part) and a specimen of punctatus from Lundy Island (2♀). The latter specimen exhibits a wing pattern typical of punctatus and is not a misidentification. The last group is more amorphous than the other three and contains specimens of pulicaris, punctatus and fagineus. When the second coordinate axis is plotted against the third (not figured), no distinct groupings are revealed, other than the five specimens of fagineus, which are well separated from the rest. These five specimens were confirmed as fagineus by the presence of the cibarial teeth, so typical of this species.

Results of Principal Coordinate Analysis Based on Wing Characters

In the analysis based on wing pattern characters (Fig. 5●), the groupings of specimens are more diffuse than shown by the analysis based on 72 variables. The principal groups revealed by inspecting plots of axes 1 versus 2, and 2 versus 3, are as follows:

1. OTU's 5,6,7,8,16,17,40,41,43,44,45,48,49... impunctatus (in part)
   grisescens (in part)
   delta (in part)
2. OTU's 10,12,13,14 ..... fagineus (in part)
3. OTU's 9,11,46 ..... mixture
4. OTU's 37,38,39 ..... pulicaris (in part)
5. OTU's 27,28,33 ..... pulicaris (in part)
   punctatus (in part)
6. OTU's 18,19,20 ..... delta (in part)
7. OTU's 1,2,3,51,52 ..... newsteadi
   lupicaris (in part)
8. OTU's 15,21,22,24,25,29,30,31,32,34,36 .... mainly pulicaris,
   and punctatus

Clearly, these loose groups do not reflect any useful taxonomic grouping, except perhaps for C. fagineus (group 2). Similar results to the ordinations were found by the cluster analysis of both data sets but these are not presented here for economy of space.

Comparison of the classifications show little congruence, implying that the predictions of the non-specificity hypothesis

do not apply in this case. This is, of course, rather a harsh test of the non-specificity hypothesis. A classification based on 13 characters might be expected to have a much greater 'sampling error' than that based on 72. It should also be noted that, since the original data are random variables, so are the principal coordinates from which the classification was constructed. The position is, unfortunately, further complicated because the non-specificity hypothesis is not a statistical one, as Jardine & Sibson (1971) have pointed out.

The use of variables from only one part of the body (such as wing pattern) is therefore not a practical method for reducing the number of variables in this study. It is possible that the subset of wing pattern characters is a particularly unfortunate example to choose, and that the non-specificity hypothesis does in fact hold. This is unlikely though, because Sneath & Sokal (1973) have found that the hypothesis is true only part of the time, thus the results of this experiment concur with those of other studies (summary in Sneath & Sokal, 1973, p. 100-102).

9.3. REMOVAL OF THE INFLUENCE OF SIZE FROM THE ANALYSIS

Sneath & Sokal (1973) have stressed that consideration be given to whether size should be removed from an analysis because of its 'all pervasive effect'.

In this study, an attempt was made to remove size from the analysis, for the following reasons.

1. In the previous section, a principal coordinate analysis was carried out on a set of 53 OTU's based on 72 variables. The plot of the first two principal coordinates (Fig. 51) showed size to be the major factor influencing the position of specimens along the first axis. The smallest specimens are to the left of the diagram (specimen 48) and the largest to the far right (specimen 31). This result is typical of many studies in which the first axis of an ordination study is found to be the 'size axis'. For example, Brown & Shipp (1977) found that the first principal component accounted for 84% of the trace and separated the taxa according to their size. However, these workers made no attempt to remove size from their analysis, even though overall size in the flies studied (Luciliini) has been shown to be the result of larval competition and larval habitat (Ayala, 1971; Lane, 1975).

In the previous section, standardisation of the data has eliminated any influence of variables measured on different scales (e.g., wing length measured in mm., compared to palp segment lengths measured in microns). Therefore, the first principal coordinate axis reflects a genuine influence of variables which measure overall size of the individual specimens. In the principal coordinate analysis of 53 OTU's based on 72 variables, only some 20% of the total variance is described in the first dimension - rather low compared to other studies. This infers that any potentially useful factors that emphasise differences in shape between species, have been relegated to lower axes. These lower axes are more susceptible to the effects of random elements in the data, and are usually more unreliable in the statistical sense. If size is removed from an analysis, any underlying shape differences between specimens (which are likely to be interspecific differences) will appear on the first axis. Consequently, the relatively small proportion of the trace described by the first few axes will, hopefully, be focused on interspecific differences, rather than size differences between specimens.

2. It has been demonstrated that general size varies seasonally within a species (see Section 7. p.101), and therefore would prove an unreliable character to use in a classification.

3. The last justification for removing size from the analyses concerns the application of the non-specificity hypothesis. The predictions and limitations of this hypothesis were discussed in the previous section. Although it is usually expressed in the context of characters from a restricted part of the body, or basic types of characters (e.g., ecological, morphological), it could equally well be applied to groups of size-related characters. The general failure of the non-specificity hypothesis has led to the suggestion that characters should be spread over the body of an organism. However, many of these characters may be highly correlated, in as much as they are the expression of a limited part of the genome. Size-related characters are an obvious example, and another might be the degree of melanisation of an insect. The general size of an organism is the result of the interaction between its genome and the environment. It is not unlikely that a set of genes controlling the size of one part of the body is likely also to control the size of another part. Hence, the use of many size-related characters from differing parts of an organism does not represent as large a sample of the genome as might first be expected. The influence of a pure size element in characters should therefore be reduced or removed for this example.

Transformation of Data to Remove Effect of Size.

Of the 72 variables listed in Table 21, 27 are linear measurements, showing a high correlation between the character size, and the overall size of the specimen. These size-related variables are listed in Table 23. Sneath & Sokal (1973) have suggested that one method of reducing the influence of size is to express each variable in terms of a general size measure (i.e., a ratio). They suggest that a suitable estimate of overall size might be the cube root of weight, or the square root of an area. Such reference dimensions are inapplicable in the present study, so an alternative was used- the 'mean logarithmic size'. This measure of an individual's size was calculated as follows:

(i) The size dependent characters for each OTU (row) were first converted to natural logarithms.

TABLE 23. Description of 27 'Size Variables'

Code Number	
21	Length of palp segment i+ii
22	" iii
23	" iv
24	" v
25	Width of palp segment iii
26	Head length
27	Proboscis length
28	Cibarium length
29	Pharynx length
32	Wing length
33	Wing width
34	Costa length
48	Fore leg : length of femur
49	tibia
50	tarsus i
51	tarsus ii
52	tarsus iii
53	Mid leg : length of femur
54	tibia
55	tarsus i
56	tarsus ii
57	tarsus iii
58	Hind leg : length of femur
59	tibia
60	tarsus i
61	tarsus ii
62	tarsus iii

(ii) The row mean, $\bar{X}$, of these logarithmically transformed characters, x_i, was calculated. This row mean is the 'mean logarithmic size' and was used as a measure of the general size of an individual.

(iii) Each size-related variable was then expressed in terms of this general size statistic, by subtracting each x_i from its row mean $\bar{X}$.

This transformation was carried out on each of 53 OTU's (details given in appendix) and a principal component analysis was performed on the new matrix.

Results of Principal Component Analysis

A summary of the results is given in Table 24.

The first eigenvector absorbs only 15% of the total variance, and when the first three vectors are combined, they absorb only 37%. This unusually low proportion of the trace suggests that the data are not easily summarised. However, the relative sizes of the eigenvector elements show which variables, if any, contribute most to the variation along any one vector. Histograms showing frequencies of vector loadings in the three largest eigenvectors are given in Fig. 52. The eigenvector is scaled (sum of squared elements = 1), so the average loading is $1/\sqrt{72} = 0.117$. In the first three axes, the loadings are clustered around the mean values ($\bar{+}$ 0.117), indicating that most variables contribute a small, but approximately equal, amount to each axis, making any simple taxonomic interpretation of the axes rather difficult.

On the first vector, the largest loadings are associated with wing pattern (Table 24), but the loadings are only marginally larger than the mean values. The situation is slightly clearer on the second axis where one variable - the presence of cibarial teeth - has a significantly larger loading. Although size no longer dominates the analysis as it did prior to transformation, the results are not easily interpreted taxonomically.

TABLE 24.

Eigenvalues and Eigenvectors of a Principal Component Analysis, Based on a Matrix Transformed to Reduce the Effect of Size.

	1	2	3
Eigenvalue	11.28	8.88	6.48
Cumulative Percentage of Trace	15.66	28.00	37.01
Variable		Eigenvectors	
1	0.045	-0.080	-0.081
2	-0.143	-0.164	0.077
3	-0.032	-0.146	0.015
4	0.008	-0.229	-0.027
5	0.076	-0.178	-0.151
6	0.126	-0.041	-0.113
7	0.118	-0.115	-0.095
8	0.120	-0.184	-0.079
9	0.115	-0.170	-0.184
10	0.176	0.133	-0.077
11	0.122	0.129	0.078
12	-0.010	0.186	-0.050
13	-0.130	0.165	0.057
14	-0.153	0.107	0.116
15	0.151	0.066	-0.203
16	0.136	0.110	0.080
17	0.160	0.059	0.018
18	0.099	-0.011	0.073
19	0.107	0.139	0.103
20	0.099	0.118	0.159
21	0.076	0.024	0.128
22	0.101	0.079	-0.177
23	0.029	0.078	0.227
24	0.075	0.002	0.109
25	0.007	0.214	-0.065
26	-0.058	-0.191	0.049
27	-0.135	-0.138	-0.007
28	-0.005	-0.021	0.050
29	-0.078	0.029	0.063
30	0.159	0.111	-0.071
31	0.158	0.031	-0.001

Variables may be identified from Table 21.

TABLE 24 contin... Variables may be identified from Table 21.

32	-0.073	-0.030	0.130
33	-0.097	0.003	0.232
34	-0.121	0.125	0.086
35	-0.191	0.089	-0.135
36	-0.188	0.077	-0.072
37	-0.208	0.093	-0.019
38	-0.224	0.022	-0.079
39	-0.234	0.030	-0.152
40	-0.080	-0.018	-0.250
41	-0.212	0.056	-0.203
42	-0.175	0.075	-0.136
43	-0.134	-0.054	0.168
44	-0.092	0.048	-0.198
45	-0.153	0.073	-0.089
46	-0.216	-0.127	-0.115
47	-0.186	0.055	-0.104
48	-0.119	-0.103	0.058
49	-0.063	-0.167	0.031
50	0.011	0.017	0.058
51	0.077	-0.006	-0.045
52	0.125	-0.013	-0.096
53	-0.034	-0.131	0.102
54	-0.061	-0.155	0.037
55	0.077	-0.077	0.095
56	0.100	-0.005	-0.229
57	0.134	0.082	-0.057
58	-0.096	-0.062	0.127
59	-0.105	-0.102	0.118
60	0.053	0.092	0.204
61	0.033	-0.001	0.074
62	0.156	-0.006	-0.045
63	-0.044	0.093	0.035
64	0.130	-0.034	-0.263
65	-0.039	0.265	0.056
66	0.048	0.228	-0.153
67	-0.105	-0.240	0.000
68	-0.016	-0.159	-0.152
69	-0.017	-0.194	-0.135
70	-0.050	0.022	0.019
71	0.087	0.126	-0.080
72	-0.081	0.156	-0.003

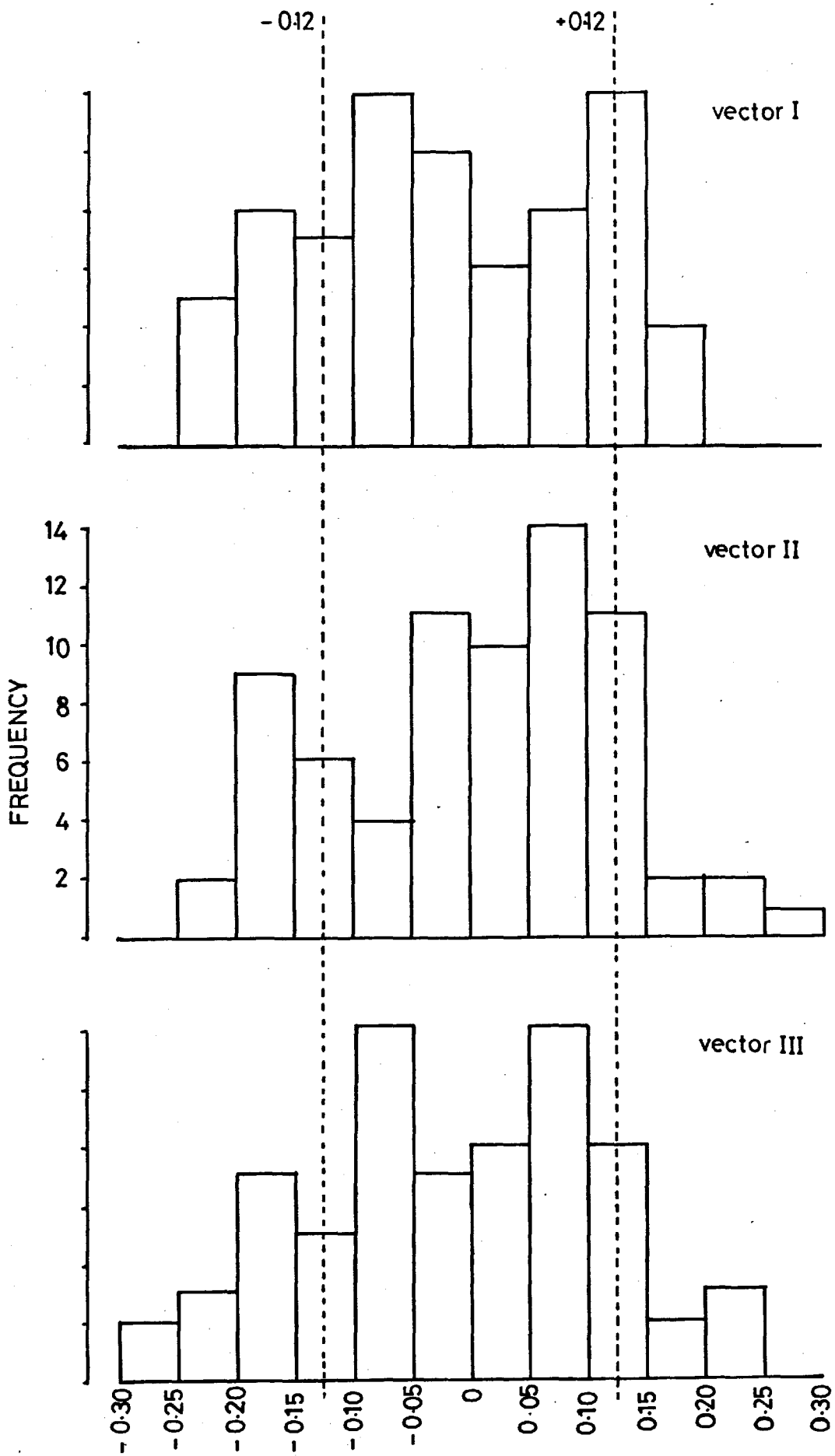


FIG. 52
FREQUENCY HISTOGRAMS OF VECTOR LOADINGS FOR ANALYSIS AFTER ADJUSTMENT FOR OVERALL SIZE

The first two principal components are plotted in Fig. 53. The overall arrangement of OTU's is similar to the analysis based on all 72 variables, the principal difference lying in that the clusters are less distinct. Among the detailed differences between the ordinations, the separation of newsteadi (OTU's 1-4) and impunctatus (OTU's 40-49) is of particular interest. In the plots based on 72 variables (Fig. 53), the two species are mixed, but after attempts to remove the effect of general size, the two taxa become distinct. This infers that, although the species are similar in size (and correlated variables), there are several differences in other characters. Another interesting detail concerns a sample of C. punctatus from Norway (OTU's 26-28). The specimens were placed with the small species, C. newsteadi and impunctatus, in the components plots based on 72 variables. After transformation to reduce size effects, these specimens were distributed throughout the punctatus cluster, thereby confirming the provisional identification. The extensive overlap between the clusters of pulicaris, punctatus and delta, shown in Fig.53 suggests these species have a marked size-independent similarity. Whether basic shape is responsible for this, or a combination of other non-size variables , is difficult to ascertain.

Thus, the results of the component analysis of transformed data, whilst providing some interesting detailed findings, are equivocal and open to two interpretations. Firstly, that the transformation has allowed basic 'shape' differences to be measured, and secondly, that the attempts to remove the effect of gross size have rendered the size variables 'neutral', and the observed classification is therefore based on the 45 non-size variables. The alternative conclusions were tested by repeating the analysis, using only the 27 size variables.

Classification Based on Size Variables Only

A principal coordinate analysis was carried out on the 530TU's (see appendix for identification of specimens), using only the 27 size-related variables. A plot of the first two axes is given in Fig. 54 and the eigenvalues and percentage variance of each axis in Table 25.

As expected, the first axis accounted for a large proportion (45%) of the total variance and the specimens were spread along it according to their size, from the largest to the smallest.

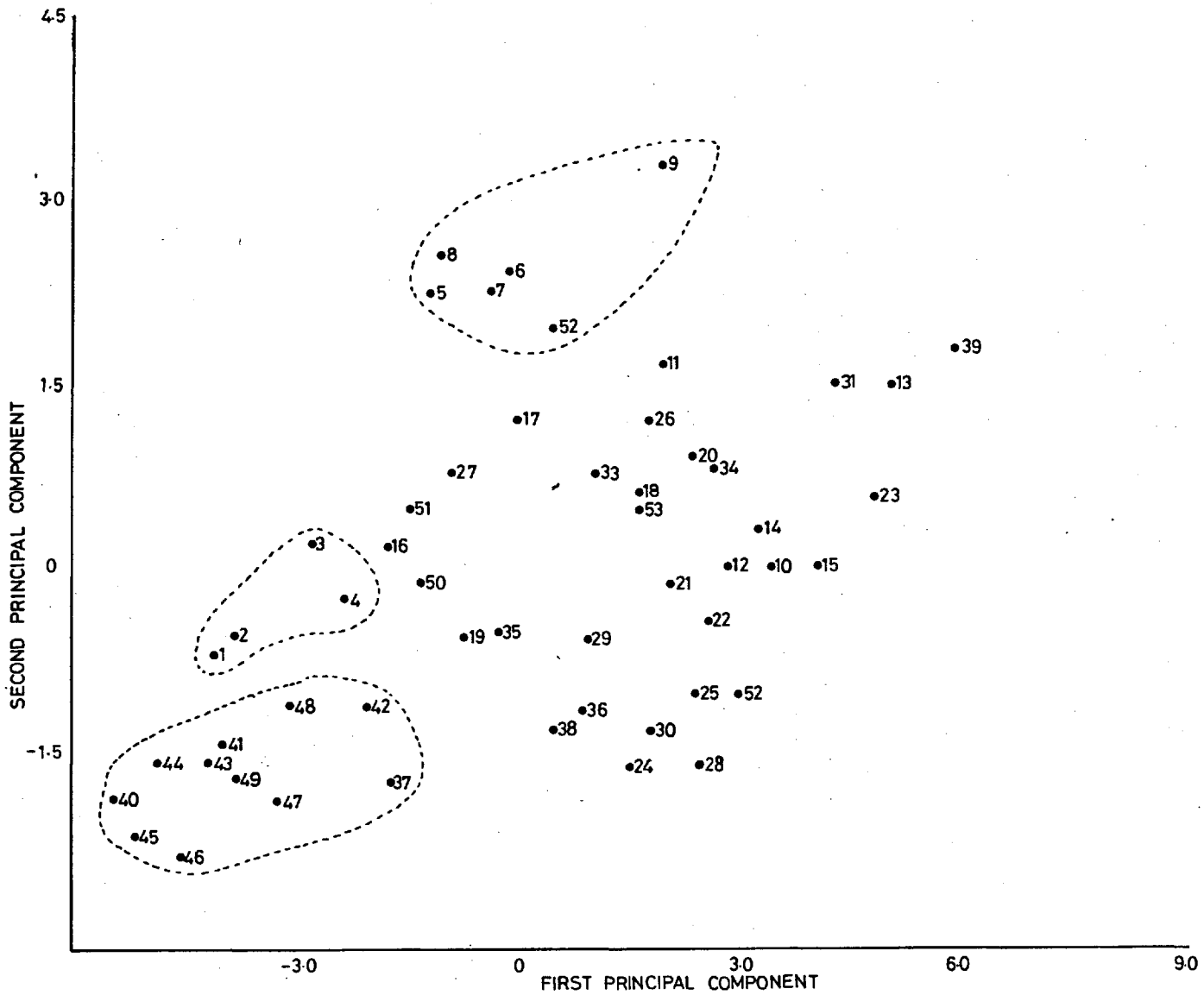


FIG. 53
PRINCIPAL COMPONENTS PLOT: DATA TRANSFORMED TO REDUCE EFFECT OF BODY SIZE

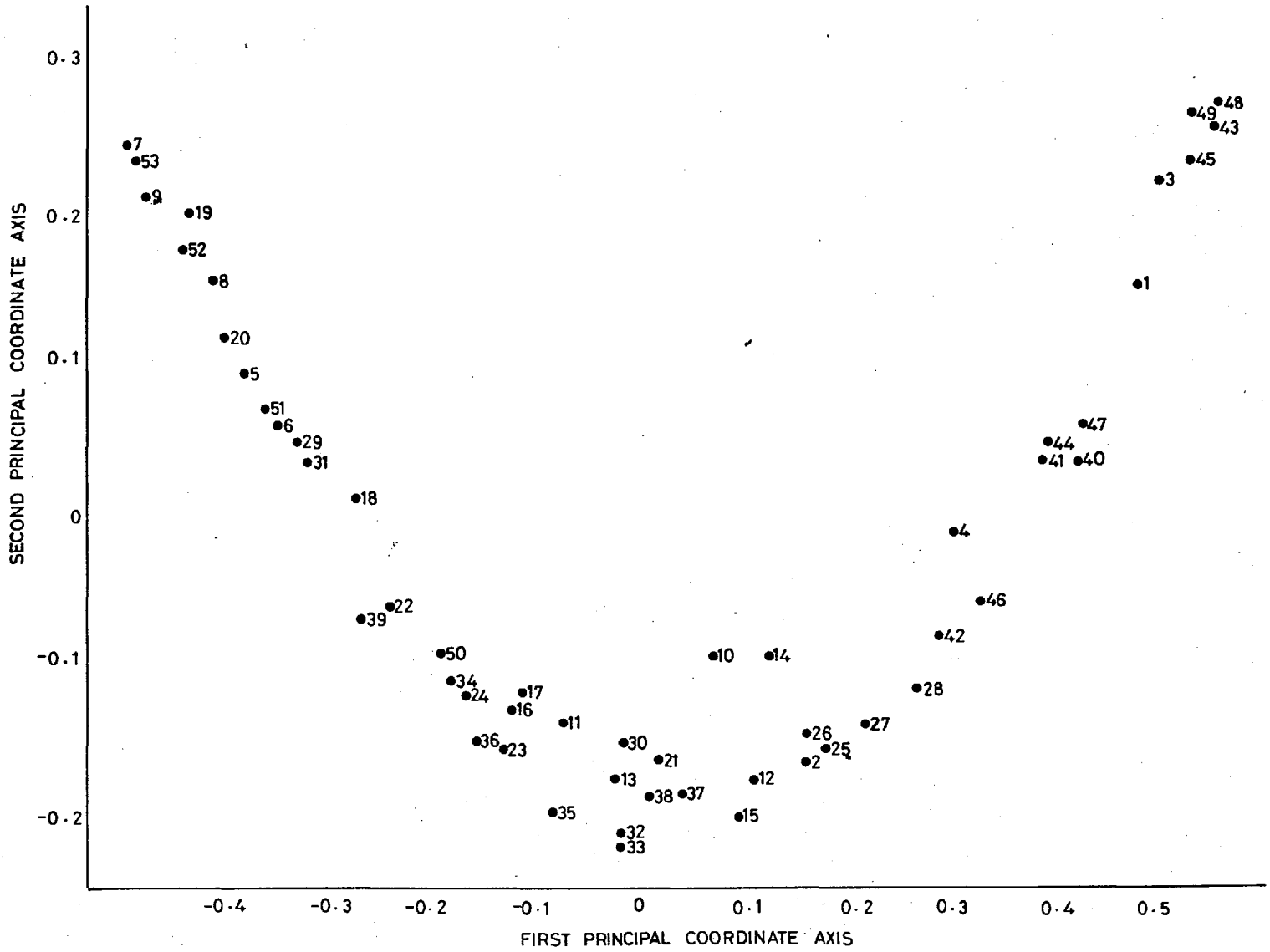


FIG. 54
PRINCIPAL COORDINATE ANALYSIS
BASED ON SIZE RELATED VARIABLES
ONLY.

TABLE 25.

Percentages and Cumulative Percentages of the Total Variance Associated with the Eigenvalues of a Distance Matrix, based on 27 Size variables. (Results of a principal coordinate analysis).

Eigenvalue	Percentage Variance	Cumulative Percentage Variance
1	45.83	45.83
2	11.37	57.20
3	4.45	61.65
4	3.24	64.89
5	2.67	67.56
6	2.03	69.59
7	1.90	71.49
8	1.50	72.99
9	1.48	74.47
10	1.34	75.81

The most obvious feature of the plot in Fig. 54 is that the 53 specimens are arranged in a remarkably symmetrical hyperbola. The first axis is obviously size, but the significance of the second is more elusive. Whatever the factor dominating the second axis, it reaches a maximum, or minimum, at the mean size of the Culicoides studied. Unfortunately, in principal coordinate analysis it is not possible to obtain any indication of which variables are contributing most to this axis (other than by regression analysis), as can be gleaned from an R-mode principal component analysis. Therefore a principal component analysis was performed on the same primary data matrix, in order that the factors influencing this second axis might be found.

The firstthree eigenvalues and eigenvectors of the correlation matrix are given in Table 26 . The first eigenvector absorbs 78% of the total variance and the second eigenvector accounts for a further 4%. The relative sizes of the eigenvalues implies that the distribution of the points representing the OTU's in hyperspace approximates a slender cylinder. The percentage of the total variance described in the first axes of the principal coordinate and the principal component analyses of the data, is surprisingly different. The first axis absorbs 45% of the variance in the former, and 78% in the latter.

All the loadings on the first eigenvector are negative and lie between -0.215 and -0.125, confirming the expectations that this is a general size vector. By definition, the average values of the loadings must be $\frac{1}{\sqrt{27}}$, i.e. 0.192 and, as the frequency histogram for this vector shows (Fig. 55), they are tightly grouped around this value. All the variables thus contribute an approximately equal amount to this vector, with no one group of variables being more important than any other. The large proportion of the total variance absorbed by this vector suggests that variation in size is more prominent than variation in shape.

The range of loadings on the second and subsequent eigenvectors is much greater than the first, with most loadings clustering around zero (Fig.55). The majority of the variables therefore contribute relatively little to these vectors and only those associated with large negative or positive values are of interest.

In the second vector, 3 variables show large negative loadings; width of palp segment iii, length of cibarium and length of pharynx.

TABLE 26.

Eigenvalues and Eigenvectors from a Principal Component Analysis Based on 27 Size Variables

	1	2	3
Eigenvalue	21.19	1.077	0.854
Cumulative Percentage of Trace	78.46	82.45	85.62
Variable	Eigenvectors		
21	-0.191	-0.012	0.010
22	-0.192	0.013	0.037
23	-0.165	0.163	0.336
24	-0.183	0.222	-0.021
25	-0.122	-0.538	-0.506
26	-0.207	-0.065	0.019
27	-0.209	0.047	-0.031
28	-0.120	-0.539	0.521
29	-0.154	-0.348	0.371
32	-0.203	0.024	0.069
33	-0.198	0.178	0.029
34	-0.196	0.219	0.146
48	-0.211	-0.022	-0.076
49	-0.208	-0.055	-0.076
50	-0.202	-0.001	-0.113
51	-0.194	0.023	-0.238
52	-0.195	0.099	-0.167
53	-0.212	-0.045	-0.042
54	-0.207	-0.029	-0.032
55	-0.205	-0.063	-0.077
56	-0.201	0.016	-0.148
57	-0.199	0.005	-0.079
58	-0.185	-0.084	-0.086
59	-0.189	0.200	0.128
60	-0.205	0.073	0.027
61	-0.196	0.010	0.130
62	-0.200	0.098	0.033

Variables may be identified from Table 23.

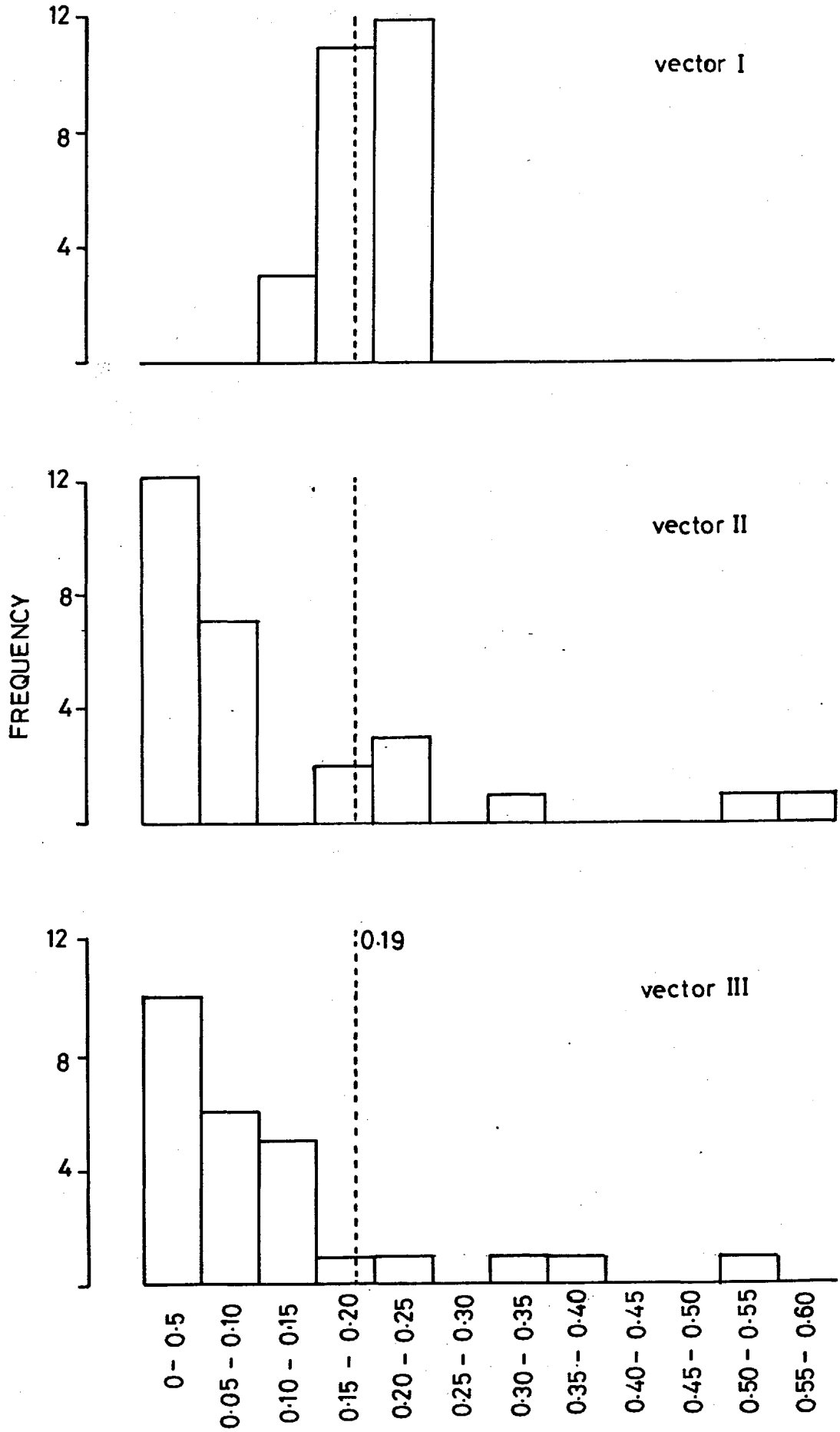


FIG. 55
FREQUENCY HISTOGRAMS OF VECTOR LOADINGS 'SIZE' DATA

The relative width of the palp segment is usually combined with its length (a general measure of size) and expressed as a ratio (palp ratio). It is of considerable value in the taxonomy of Culicoides, and therefore, although the second vector is of minor importance statistically, it is useful taxonomically. For example, it separates two small species, newsteadi and impunctatus, readily from one another.

The third eigenvector also has three variables with large loadings; width of palp segment iii (negative), length of pharynx (positive) and length of palp segment iv (positive). The variables associated with the two largest loadings were important in defining the second vector. Identification of the remaining vectors is far more difficult, and they account for such a small proportion of the total variance that interpretations are more likely to be erroneous as a result of spurious effects.

The plot of the first two principal components (Fig. 56) is quite different from that of the first two principal coordinate axes, based on the same data. This is an unexpected result, as both methods usually produce almost identical results. Although two OTU's are displaced (OTU's 15 and 18), the order of the OTU's along the first axis is not significantly different. The main difference in these ordination diagrams is the position of the OTU's on the second axis. Principal coordinate analysis shows a factor influencing the limits of size variation in a regular manner, whereas there is no obvious regularity in the component analysis. The eigenvalues associated with the second vector also differ between the analyses, accounting for a much smaller proportion of the variation in the component analysis. It is likely that the hyperbolic arrangement of the OTU's in the principal coordinate analysis does not reflect a regular size-dependent growth factor, but a distortion due to the process itself. Further research into the exact cause of this distortion is important in furthering our knowledge of the conditions under which this technique may give misleading results.

In summary, most size-related variables show little interspecific variation in shape. Exceptions are the width of palp segment iii and the length of the cibarium and pharynx. The first of these is already in use in the taxonomy of Culicoides, as the palp ratio. The other two variables may prove to be useful if combined with some measure of overall size, to form a ratio.

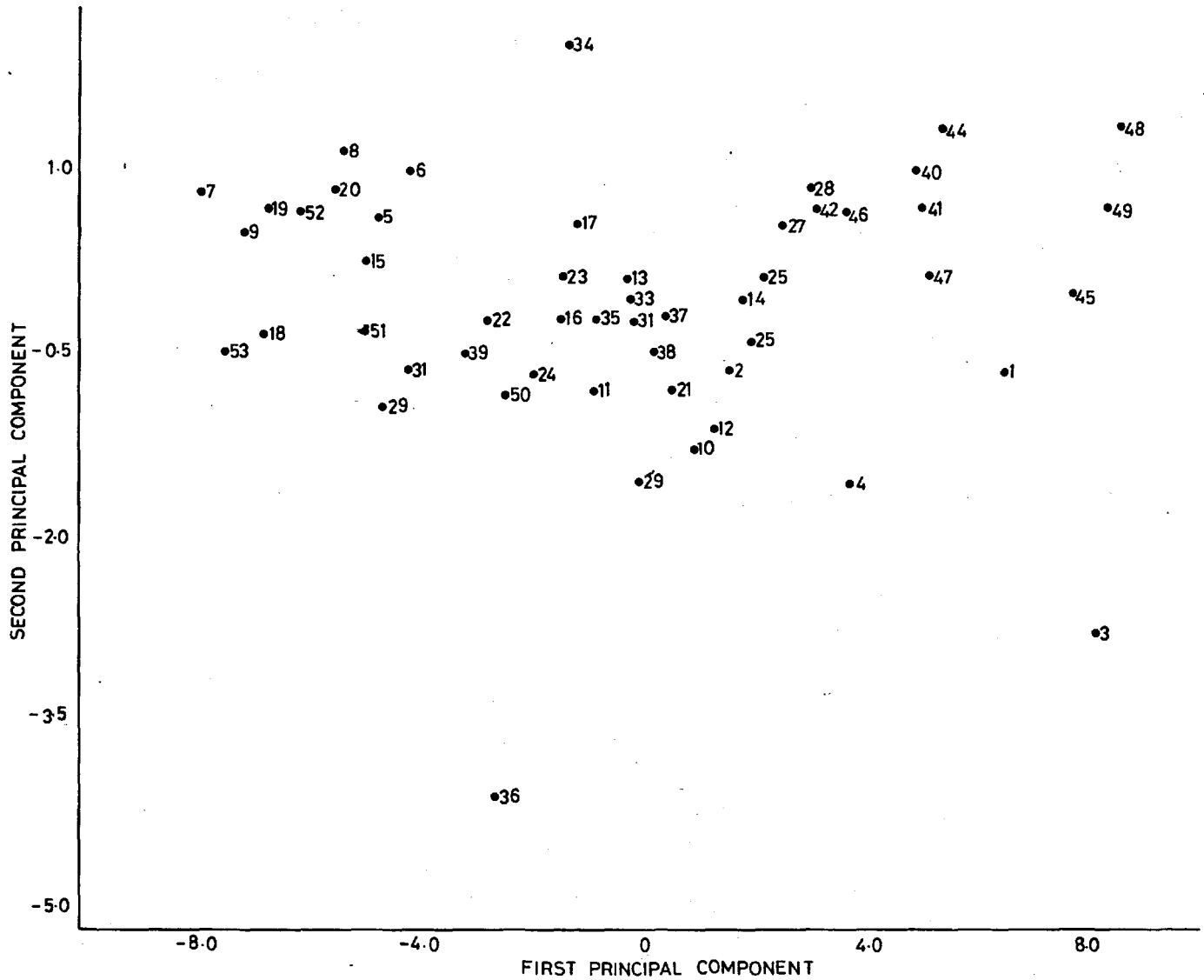


FIG. 56
PRINCIPAL COMPONENT ANALYSIS
BASED ON SIZE RELATED VARIABLES

9.4. REDUCTION OF CHARACTERS BY OBJECTIVE MEANS

There have been a number of 'objective' methods described for discarding a variable in principal component analysis. In this section, two of these methods are tested: cluster analysis of the variables and secondly, the use of the eigenvector elements themselves.

9.4.1. Previous Studies

Although much work has been done with the selection of the best subset of variables in multiple regression analysis, little has been achieved with respect to principal component analysis until Jolliffe (1972, 1973).

Jolliffe (1972) described eight rejection methods, which may be divided into three main groups:

1. Multiple correlation methods.

The first of two methods described in this category retains a set of p variables which "maximises the minimum correlation between the p selected variables" and any of the (K-p) rejected variables. This was the method of Beale et al. (1967) and is very slow. When 30 variables are involved, several hours of computer time are needed. A second and quicker method, was a stepwise procedure which rejects that variable having the maximum multiple correlation with the remaining (K-1) variables, until p variables remain.

2. Four methods use the principal components themselves. The first was that of Beale et al. (1967) in which a principal component analysis was performed on K variables and the eigenvalues inspected. Then if some eigenvalues, p_i, are less than some number λ_o the corresponding eigenvectors are examined, starting with the eigenvector associated with the smallest eigenvalue. The variable with the largest coefficient is then associated with each of these p_i eigenvectors. The p_i variables are then rejected. Another principal component analysis is then done and the process reiterated until all the eigenvalues in the latest analysis are greater than λ_o. This method uses a considerable amount of computer time. The other three methods are much faster.

The next method only does one component analysis and again associates one variable with each of the (K-p) components and

rejects these variables to leave p useful variables. The third method again uses the last (K-p) components, but in this case, the sums of squares of coefficients of all variables in the (K-p) components are calculated, and the (K-p) variables for which this product is the largest, are rejected. The last method in this group is the complement of the second method, in that it concentrates on the first p components, associating a variable with each and then rejecting the (K-p) variables.

3. Cluster analysis.

This third category requires a cluster analysis of the variables themselves, to form groups from each of which a single variable is selected. Jolliffe used single-linkage and average-linkage cluster analyses, and although he found the latter to be better than than the first, he found single-linkage cluster analysis to be a useful technique.

Jolliffe tested five of these methods on artificial data (Jolliffe, 1972) and four (with two further variations on cluster analysis) using real data (Jolliffe, 1973). He found that no one method was significantly better or worse than any other. Therefore, on the criterion of speed of computation, the clustering methods were deemed most successful.

For each rejection method, a suitable criterion was found for empirically deciding how many variables to retain. In the principal component anlysis, the most satisfactory results were obtained when the number of variables rejected(K-p), equals the number of eigenvalues (of a correlation matrix) below 0.70. For single-linkage cluster analysis, Jolliffe found that the appropriate number of variables to retain is the number of clusters present, when the intercluster similarities (= phenon level) falls below 0.55.

9.4.2. Classification Based on the Complete Set of Characters

A classification based on the complete set of characters was produced for the following reasons:

(i) As a standard for comparison with classifications based on subsets of variables.

(ii) To discover the relationships between the taxa, using the maximum information available.

A principal component analysis was performed on a primary data

matrix of order 84 x 72 (given in the appendix). Details of the 84 OTU's are given in Table 27 and for 72 variables in Table 21. The five largest eigenvalues and eigenvectors for this analysis are listed in Table 28, and a plot of the first two components are presented in Fig.57 .

First Principal Component

This accounts for 33% of the total variance - nearly three times as large as the second principal component. The loadings on the associated vector show the influence of two main classes of characters - general size and wing pattern - which have a contrasting effect. The frequency histogram of the loadings (Fig. 58) has a bimodal distribution centred on 0.00 and -0.18. The high negative values are associated with the size variables that dominate this component. The loadings for the wing characters are smaller than for many of the size characters. As expected, the first principal component (Fig. 57) places the specimens according to their size, the small ones to the right of the diagram and the large ones to the left.

Second Principal Component

This axis may be reliably identified as reflecting wing pattern because most of the large loadings are associated with these characters. The only pattern character which does not have a high loading is wing pattern element 9 (variable 43), which has a small negative loading. This pattern element describes the spot in the cubital cell and has been used frequently in the taxonomy of the pulicaris group. Its very small loading suggests that the emphasis traditionally put on this character is not warranted. The frequency histogram for this component (Fig. 58) shows the loadings to be clustered around zero. There is slight evidence of bimodality, resulting from the relatively large loadings for wing pattern characters.

Third Principal Component

The highest loadings on this axis are associated with antennal characters. There is an interesting contrast in the effect of these characters. The segments which form the proximal section of the antenna (segments iii - x, variables 2 - 9) have positive loadings,

TABLE 27.

Collection Data and Code Numbers of Specimens Used in the Numerical Classification of the *C. pulicaris* Complex

OTU Code No.	Locality	Date Colln.	Detailed Locality	Provisional Identificat
1	Wales: Carmarthen	6.vii.69		
2	Wales "	"		
3	Eire: Blasket Isle	vi.	Inishtearght	
4	England: Essex	22.vi	E. Tilbury	
5	N. Ireland: Antrim	v.	Belfast	
6	Eire: Cork	25.ix.		newsteadi
7	Scotland: Perthshire	Bred	Inchture	(=halophilu
8	" "	"	"	
9	Eire: Cork	25.ix.	Skibbereen	
10	" "	"	"	
11	Scotland: Perthshire	8.vii	Powgavie	
12	Scotland: Cromarty	6.viii	Dingwall	
13	Wales: Montgomery	ix.	Staylittle	
14	" "	vii.	"	
15	" "	"	"	
16	England: Durham	viii.	Stockton	
17	Wales: Cardigan	15.vii.	(Aberystwyth)	
18	USSR: Moscow Distr.	25.vii.		
19	" " "	9.viii		grisescens
20	Scotland: Stirlinshire	8.ix.	Plean	
21	" "	"	"	
22	Scotland: Inverness	9.x.	Drinsallie	
23	" "	14.viii.	Glen Affric	
24	Israel: Tel Aviv	18.v.	(Neve Yaar)	
25	England: Surrey	4.viii.	(Pirbright)	
26	England: Hants	vi.	(Bank) p.type	fagineus
27	" "	"	" "	
28	" "	"	" "	
29	England: Kent	8.v.	Beckenham	
30	Scotland: Midlothian	18.vii.	Glentarf	
31	Scotland: Argyl	11.vii.	Lephinmore	delta
32	Scotland: Berwick	21.vi.	Gordon Moss	
33	England: Surrey	19.vi	Woking	

TABLE 27 contin...

OTU Code No.	Locality	Date Colln.	Detailed Locality	Provisional Identificatn.
34	England: Hants	12.v.	(Alice Holt)	
35	England: Surrey	14.v.	(Pirbright)	delta
36	England: Kent	8.v.	Beckenham	
37	" "	"	"	
38	England: Surrey	19.vi.	(Woking)	
39	" "	"	"	
40	England: Hants	30.v.	Fareham	
41	Wales: Pembrokeshire	2.vii.	(Amroth)	
42	England: Hampshire	13.vii.	(Alice Holt)	
43	Norway: Kantokeino	28.vii	69 O'N	
44	" "	"	"	
45	" "	"	"	
46	" "	"	"	
47	" "	"	"	
48	" "	"	"	punctatus
49	" "	"	"	
50	" "	"	"	
51	England: Devon	24.vii	(Lundy Island)	
52	Japan: Okayama	22.vii	(Yoshi Mach)	
53	" "	"	"	
54	" "	"	"	
55	" "	"	"	
56	" "	"	"	
57	" "	"	"	
58	" "	"	"	
59	USSR: Moscow Distr.	13.viii.		
60	" " "	3.ix.		
61	England: Surrey	9.viii.	(Wooton)	pulicaris
62	Scotland: Perthshire	2.ix.	(Powgavie)	
63	Wales: Pembrokeshire	2.vii.	(Amroth)	
64	England: Essex	22.vi.	(E. Tilbury)	
65	England: Hants	12.v.	Alice Holt	
66	England: Surrey	4.x.	Bullswater	
67	" "	"	"	

TABLE 27 contin...

OTU Code No.	Locality	Date Colln.	Detailed Locality	Provisional Identificatn.
68	England: Surrey	4.x.	Bullswater	
69	England: Surrey	10.x.	Trinity	pulicaris
70	England: Hampshire	20.iii.	Bank	
71	Scotland: Inverness	27.viii	Loch Ness	
72	" "	"	"	
73	" "	"	"	
74	Scotland: Buteshire	28.viii.	Kingarth	
75	" "	23.viii.	"	impunctatus
76	Wales: Pembrokeshire	2.vii.	Amroth	
77	Scotland: Buteshire	24.viii.	Rothesay	
78	Scotland: Inverness	26.viii.	(Loch Ness)	
79	" "	"	"	
80	" "	28.viii	"	
81	Scotland: Argyl	24.vii	Lephinmore	
82	England: Surrey	4.viii.	(Pirbright)	
83	" "	13.x.	"	lupicaris
84	England: Surrey	13.vii.	"	

TABLE 28.

Eigenvalues and Eigenvectors from a Principal Component Analysis Using 72 Variables (84 OTU's)

	1	2	3	4	5
Eigenvalue	24.11	9.127	5.735	3.508	3.100
Cumulative Percentage of Trace	33.49	46.17	54.14	59.01	63.31
Variable			Eigenvectors		
1	0.043	-0.155	-0.053	0.017	-0.226
2	0.077	0.032	0.159	0.240	0.074
3	0.054	-0.037	0.182	0.169	0.002
4	0.016	-0.102	0.241	0.091	-0.140
5	-0.005	-0.096	0.295	0.006	-0.208
6	-0.065	-0.057	0.299	-0.042	-0.026
7	-0.023	-0.076	0.289	-0.073	0.121
8	0.125	-0.168	0.205	-0.131	0.147
9	0.004	-0.128	0.211	-0.202	0.177
10	-0.068	-0.112	-0.153	-0.214	-0.097
11	-0.051	-0.079	-0.203	-0.100	0.149
12	-0.024	0.026	-0.214	-0.080	-0.016
13	-0.000	0.109	-0.276	0.024	-0.093
14	0.022	0.140	-0.156	0.145	0.025
15	-0.018	-0.052	-0.083	-0.341	-0.064
16	-0.092	-0.061	-0.052	0.052	0.016
17	-0.081	-0.074	-0.016	-0.079	-0.002
18	-0.065	-0.098	0.125	0.141	-0.053
19	-0.102	-0.048	-0.073	0.149	-0.155
20	-0.087	-0.045	-0.081	0.191	-0.075
21	-0.181	-0.016	0.014	0.077	-0.055
22	-0.179	0.037	0.059	-0.105	-0.057
23	-0.164	0.001	0.041	-0.088	-0.169
24	-0.172	0.043	0.000	-0.031	-0.072
25	-0.088	-0.103	0.038	-0.070	0.361
26	-0.185	-0.083	-0.011	0.037	-0.004
27	-0.196	0.017	0.002	0.007	-0.044
28	-0.123	-0.015	-0.056	0.012	-0.009
29	-0.148	-0.044	-0.027	-0.074	0.001
30	-0.121	-0.041	-0.020	-0.202	-0.045

Variables may be identified from Table 21.

TABLE 28 contin...

Variable	1	2	3	4	5
			Eigenvectors		
31	-0.085	-0.097	-0.045	-0.137	0.090
32	-0.186	0.027	0.004	0.034	0.029
33	-0.178	0.035	0.023	0.074	0.000
34	-0.183	0.083	-0.026	0.066	-0.044
35	0.022	0.224	0.097	0.021	-0.021
36	0.032	0.232	0.035	0.016	0.011
37	0.036	0.227	0.019	0.117	-0.033
38	0.067	0.250	0.063	-0.034	0.125
39	0.018	0.273	0.145	-0.056	-0.025
40	0.057	0.148	0.003	-0.309	0.040
41	0.021	0.272	0.139	-0.101	-0.041
42	-0.003	0.242	0.135	-0.050	-0.013
43	0.034	-0.024	-0.061	0.052	0.346
44	0.069	0.203	0.095	-0.166	-0.120
45	0.056	0.192	-0.078	-0.164	0.186
46	0.016	0.292	0.044	-0.061	0.023
47	0.042	0.235	0.039	-0.100	0.121
48	-0.196	0.025	0.021	0.002	0.036
49	-0.196	0.012	0.026	-0.011	0.045
50	-0.190	0.024	0.021	0.007	0.065
51	-0.181	0.051	0.015	-0.027	0.094
52	-0.095	0.076	-0.005	-0.023	0.097
53	-0.197	0.002	0.019	0.009	0.051
54	-0.193	0.006	0.027	-0.010	0.068
55	-0.193	-0.012	0.025	-0.008	0.041
56	-0.186	0.044	0.006	0.079	0.081
57	-0.176	0.035	-0.033	0.003	0.115
58	-0.179	-0.007	0.003	0.026	0.093
59	-0.182	0.011	0.019	0.027	-0.036
60	-0.194	0.028	0.005	0.018	0.036
61	-0.184	0.042	0.059	-0.027	0.002
62	-0.170	0.083	0.028	0.014	0.093
63	-0.037	0.132	0.004	0.062	-0.000
64	-0.005	-0.087	0.045	-0.392	-0.226
65	-0.039	0.086	-0.363	-0.047	-0.031

TABLE 28 contin...

Variable	1	2	3	4	5
	Eigenvectors				
66	-0.109	0.126	0.025	-0.064	-0.325
67	0.149	-0.118	-0.017	0.048	0.061
68	0.037	0.003	0.106	-0.083	-0.125
69	0.116	-0.094	-0.088	-0.192	0.017
70	0.012	-0.033	-0.017	0.088	-0.031
71	0.054	-0.035	0.026	-0.037	-0.217
72	0.040	0.123	-0.081	0.113	-0.200

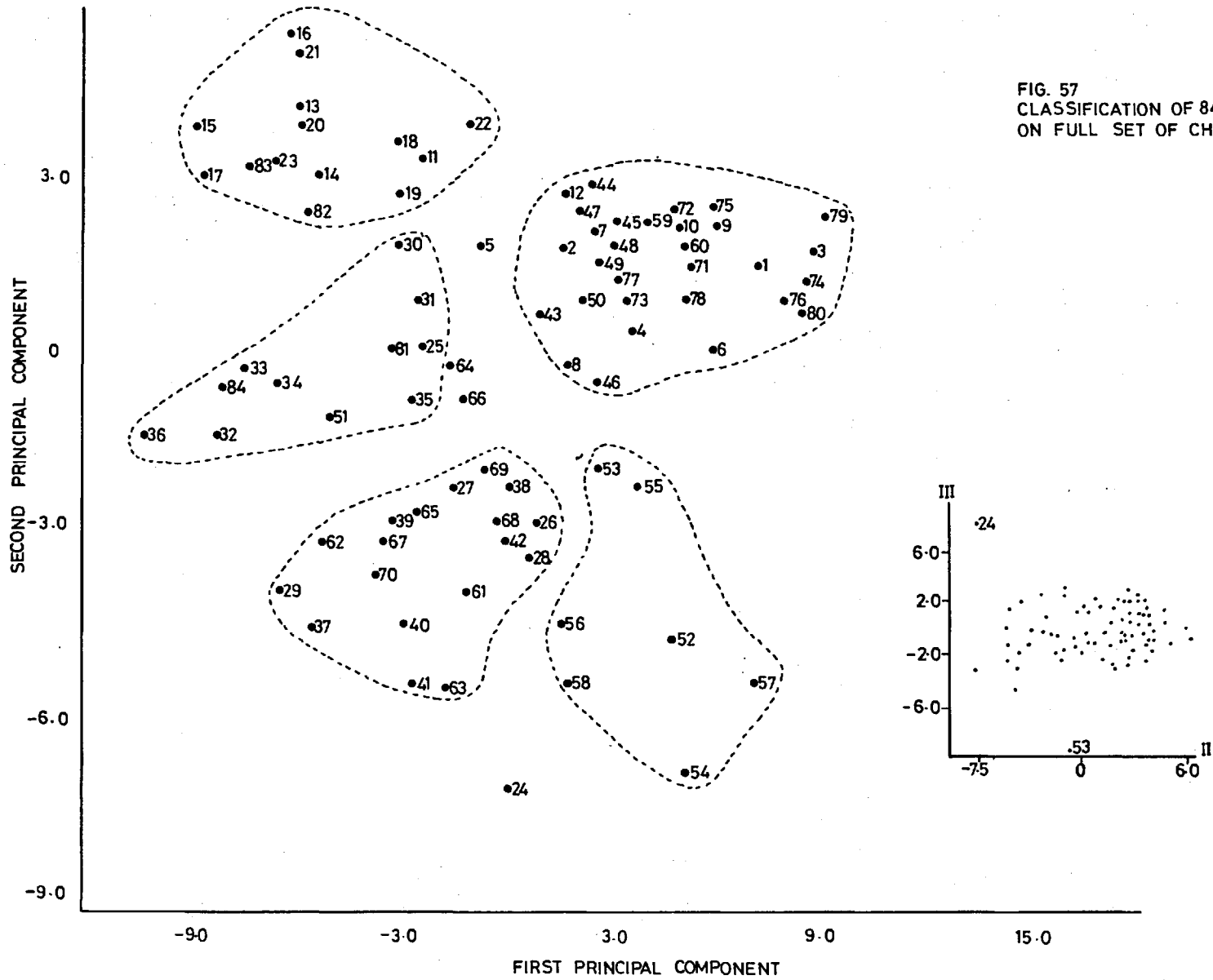


FIG. 57
CLASSIFICATION OF 84 OTU'S BASED ON FULL SET OF CHARACTERS

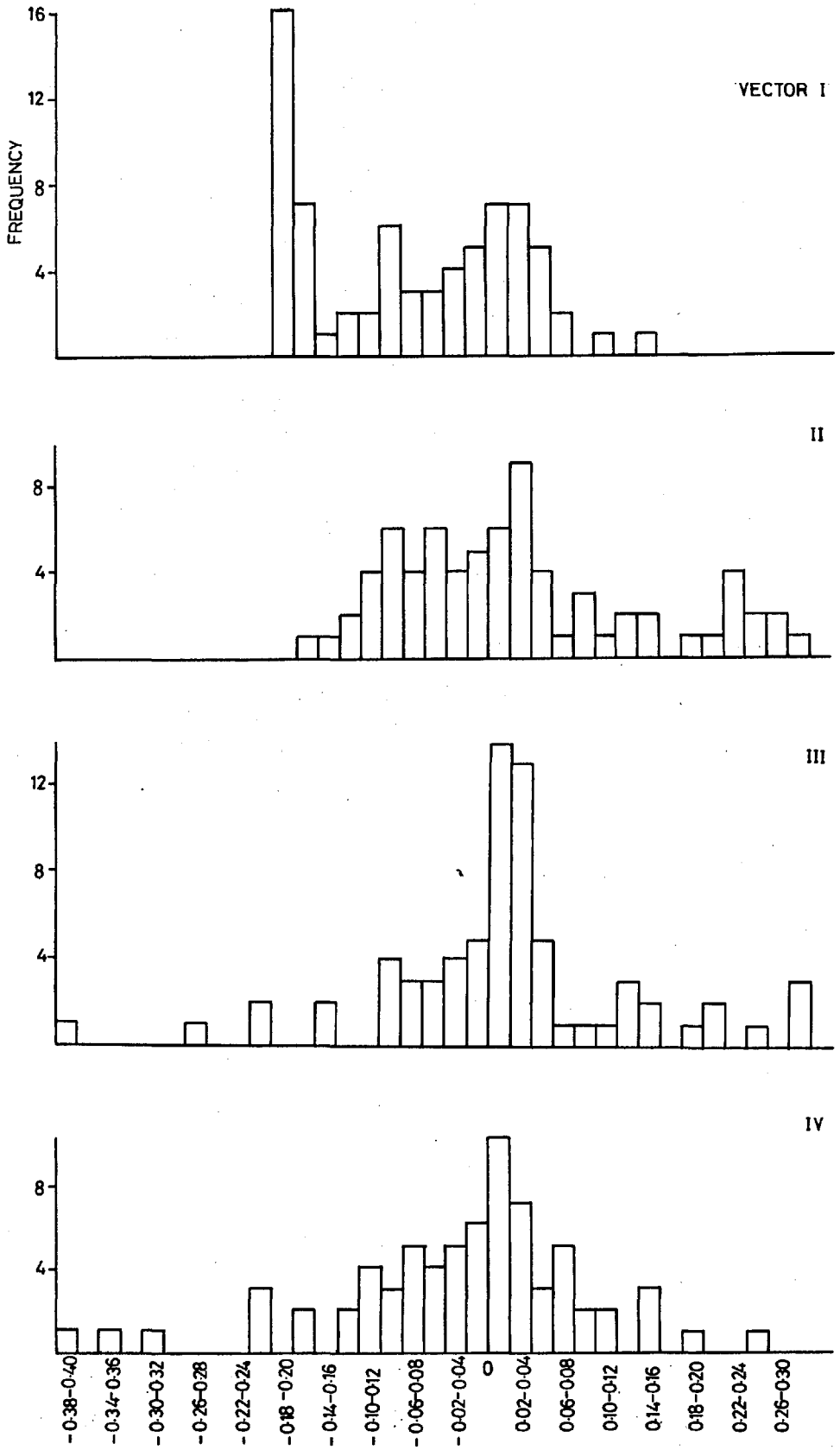


FIG. 58
FREQENCY HISTOGRAMS OF VECTOR LOADINGS FOR ANALYSIS BASED ON ALL VARIABLES

in contrast to the negative loadings for the distal antennal segments (segments xi - xv, variables 10 - 14). Which is positive or negative is immaterial, as it depends on the computational technique. It is the absolute value of the elements and the contrasts that are important. The highest loading is associated with the antennal ratio, which in some respects duplicates this trend. In fact, this ratio measures the 'shape' of the antenna in terms of the proximal section relative to the distal section. The loadings form a unimodal distribution around zero (Fig. 58). Unfortunately, this component is not very useful taxonomically. The inset in Fig. 57 shows that the majority of specimens are placed in a narrow band, with a single OTU placed above (OTU 24) and below (OTU 53).

Taxonomic Discussion

As noted above, most of the useful taxonomic information is concentrated in the first two principal components; the third, which accounts for only 8% of the total variance, has little to recommend it taxonomically. The characters which dominate the first two axes are of very different types. The importance of size for placing specimens along the first axis has already been noted. The effect of wing pattern on the second axis provides a contrast, which summarises the taxonomic variation of the 84 OTU's rather effectively, when the two axes are plotted against one another.

The general grouping of the specimens has been indicated on the plot (Fig. 57) by means of broken lines. The top right group contains specimens provisionally allocated to four species: impunctatus, newsteadi, pulicaris and punctatus. They are united by their small size rather than a basic similarity in qualitative characters. Specimens of grisescens and delta each form their own groups. There's a group composed of specimens of pulicaris, punctatus and fagineus, and a single remaining group, composed of specimens from Japan (OTU's 52 - 58), is also present. There is very little separation of the heterogeneous clusters along the third axis.

9.4.3. Ratios

Some of the characters used in the set of 72 variables are ratios and, in view of the divided opinion concerning their use

in multivariate analysis, some comment seems appropriate.

Many authors, including Jeffers (1967) and Blackith & Reyment (1971) have argued against the use of ratios. Among their objections are:- the common occurrence of allometry distorts the ratio; ratios use only two variablesto measure a shape, which may be much more complex; ratios imply only one contrast in the form of structures, and finally, ratios may duplicate other measurements made. Despite these objections, ratios are frequently used in multivariate studies and used extensively in general biological studies. They were incorporated into the present account for the following reasons:

(i) They have been used extensively in Culicoides taxonomy, and the results of this study have to be related to current taxonomic procedure. A potential loss of rigor in the analysis had to be accepted so that its taxonomic significance could be evaluated.

(ii) The contribution of ratios to the classification of Culicoides was to be tested empirically.

(iii) It was necessary to test empirically whether the few ratios compounded from other variables in the analysis had the same contribution to the results as their component parts.

Although principal component analysis is usually used for quantitative data (continuous variables), the technique is robust enough to be usable on data which are not of this type, e.g., compound variables or ordered multistate, etc. (Clifford & Stephenson, 1975; Roback & Moss, 1978).

In the analysis based on 72 variables described above, ratios contributed relatively little to the classification, in terms of character loadings on the first five principal components, which accounted for 63% of the total variance. The only exceptions were the antennal ratio - an important character on axis three - and the palp ratio, on axis five. This implies that ratios have a minimal contribution to the overall variation within the taxa studied.

In a numerical study of the tanypodine Chironomidae (a sister family of the Ceratopogonidae), Roback & Moss (1978) also found that ratios had little effect on the results. Ratios constituted 14% of their variables and were retained "because of their traditional use in midge [Chironomidae] classification". It is

interesting to note that the antennal ratio, an important character in the Chironomidae, was the only ratio with a high loading in their analysis, as found in the present study.

One important result from this empirical use of ratios, is that the concern expressed by Atchley et al. (1976) may not be justified. The use of ratios may be unsatisfactory for mathematical reasons, but their reputed disadvantages are diluted in a study that employs many characters. This is especially true in a study such as this which, not only uses a large number of characters, but also uses many different types, e.g., lengths of structures, wing pattern, meristic characters. The use of diverse types of characters has a limited effect in a statistical sense, because correlation matrices were used so that the data were, in effect, standardised in respect of the units of measurement. The main advantage is that the results of analyses using them are often taxonomically more informative.

Some studies have not used ratios directly in the multivariate analysis, but have used the results of these analyses to suggest useful taxonomic ratios. Sands (1972) used logarithmic transformation of the raw data, to test the idea that the pattern of variation between different species of termites might be at least as well expressed by ratios, as by linear combinations of the characters. Having established that loadings on the variables, using logarithmically transformed data, were almost identical to those based on the raw data, he was able to use the loadings to to suggest which characters would make useful taxonomic ratios. Characters with a positive weighting were interpreted to be useful when multiplied, and negative loadings useful when divided.

Having already shown that ratios contribute little to the classification, in terms of character loadings, their use was further tested by repeating the analysis, but omitting eight of the ratios. Comparing the results produced by the smaller set of 64 variables with those based on the complete set, would show the effect of the eight ratios. The ratios deleted were variables 65 - 72. The results of this second analysis are presented as a plot of the first two principal components (Fig. 59). The arrangement of the specimens is almost identical to that based on the complete set of characters, except that the plots are mirror images. The 'mirror image' effect is merely an artefact of computation, and has no significance, numerically or taxonomically. The percentage

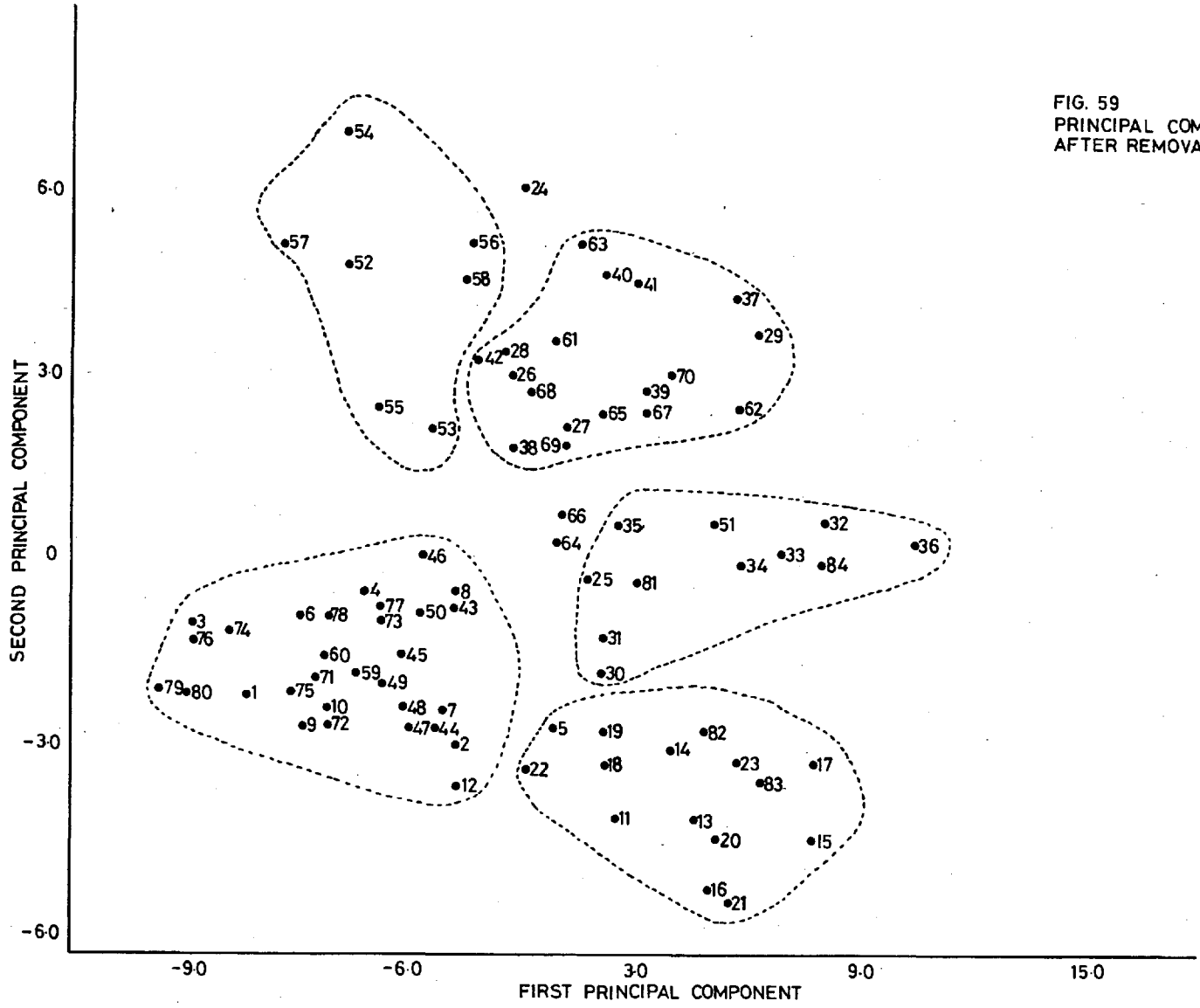


FIG. 59
PRINCIPAL COMPONENTS PLOT OF ANALYSIS
AFTER REMOVAL OF RATIOS

of the trace absorbed by the first five dimensions are 35%, 14%, 7%, 5%, and 4% respectively. The slightly higher proportion of the total variance described by these axes is because fewer variables were used. The loadings for the characters common to both axes are approximately the same but of opposite sign. The size variables have negative loadings in the analysis of the complete set of variables, but positive loadings in the analysis based on 64 variables. This change of sign is responsible for the mirror imaging of the plots. The results of this experiment confirm the conclusion made earlier, that the inclusion of ratios has no obvious deleterious effect on the arrangement of OTU's.

9.4.4. Reduction of Variables Using Between-Character Correlations

The notion of correlation between characters is generally used by biologists to describe a variety of situations. The word is often used to mean concordance, rather than statistical correlation.

Jardine & Sibson (1971) recognised five basic types of correlation between characters, of which three are relevant here: statistical, taxonomic, and functional.

Statistical correlation is the association of the characters within a population and may vary between otherwise similar populations. The concept of taxonomic correlation between characters has often been studied and is, (according to Jardine & Sibson), unrelated to the statistical correlation of characters. If two characters discriminate (or classify) OTU's in a similar manner, they are said to be taxonomically correlated. Hence, the concept of taxonomic correlation is the basis of testing the concordance of classifications, based on different characters, or sets of characters. Functional correlation of characters is more obvious than the previous two. Two or more characters which are jointly involved in the performance of the same function, are said to be functionally correlated, e.g., the various parts of the mouthparts in Culicoides are functionally correlated, as they are jointly involved in piercing and withdrawing blood from a particular host.

Jardine & Sibson (1971, p. 172) suggest that they cannot conceive any general procedure for eliminating redundant characters, although they emphasise that the study of statistical and taxonomic correlation may play a part in the selection of variables for

numerical taxonomy. Redundancy cannot be determined by the sole use of statistical dependence within populations, as this confuses redundancy in describing a given OTU with redundancy relevant to the classification of a set of OTU's.

In the present study, both taxonomic and statistical correlation (sensu Jardine & Sibson) are investigated, by using cluster analysis and principal coordinate analysis of an association matrix between characters (a modification of Jolliffe's C1 method). The starting point of both analyses is a character-character correlation matrix. From each group of characters shown by these analyses, one representative is selected. In cluster analysis, a predetermined similarity level may be used as a criterion of a cluster. The groupings shown by principal coordinate analysis are not usually so easily defined and selection must therefore remain a more subjective decision.

The Between-Character Correlation Matrix

Correlations between characters were calculated using a Pearson product-moment correlation coefficient over all 53 OTU's, so that each coefficient in the matrix is based on 53 observations.

The matrix is not given, as its inclusion would take up too much space. However, its important features are summarised below. Table 29 contains details of all the significant off-diagonal correlation coefficients in each row of the matrix, divided into positive and negative coefficients. A 5% level of probability was chosen, making values greater than 0.272 significant for 50 degrees of freedom. The proportions of significant positive and negative correlations are expressed as percentages of all row elements and not just of the significant correlations.

By inspection of the row totals of significant coefficients, it is possible to see which characters have a high overall correlation with other characters. As expected, many of the size measurements, e.g., leg lengths and wing lengths, show a large number of positive correlations with each other. A high proportion of positive correlations are also shown by the number of sensilla on the antennal flagellum (characters 15 - 20). Most of these correlations are with size variables. This emphasises an interspecific difference in the total number of antennal sensilla. A small species

TABLE 29.

Table to show Number and Percentage of Significant Off-Diagonal Correlation Coefficients in Each Row of the Character-Character Matrix.

Character (Row)	Number Positive	Percent	Number Negative	Percent	Total Percent
1	3	4.2	12	16.6	20.8
2	5	6.9	39	54.1	61.0
3	3	4.2	29	40.3	44.5
4	6	8.3	30	41.7	50.0
5	7	9.7	7	9.7	19.4
6	24	33.3	5	6.9	40.2
7	5	6.9	6	8.3	15.2
8	7	9.7	10	13.8	23.5
9	9	12.5	9	12.5	25.0
10	37	51.4	8	11.1	62.5
11	34	47.2	10	13.8	61.0
12	8	11.1	7	9.7	20.8
13	7	9.7	9	12.5	22.2
14	8	11.1	10	13.8	24.9
15	19	26.4	3	4.2	30.6
16	33	45.8	6	8.3	54.1
17	32	44.4	7	9.7	54.1
18	23	31.9	3	4.2	36.1
19	30	42.3	4	5.5	47.8
20	30	42.3	5	6.9	49.2
21	37	51.4	8	11.1	62.5
22	40	55.6	8	11.1	66.7
23	38	52.8	6	8.3	61.1
24	37	51.4	6	8.3	59.7
25	36	50.0	20	27.8	77.8
26	39	54.2	12	16.6	70.8
27	41	56.9	11	15.3	72.2
28	33	45.8	5	6.9	52.7
29	37	51.4	9	12.5	63.9
30	34	47.2	4	5.5	52.7
31	34	47.2	8	11.1	58.3
32	38	52.8	10	13.8	66.6
33	39	54.2	8	11.1	65.3
34	38	52.8	7	9.7	62.5
35	13	18.1	5	6.9	25.0

TABLE 29 contin...

Character (Row)	Number Positive	Percent	Number Negative	Percent	Total Percent
36	11	15.3	4	5.5	20.8
37	10	13.8	8	11.1	24.9
38	13	18.1	27	37.5	55.6
39	12	16.6	28	38.9	55.5
40	10	13.8	25	34.7	48.6
41	11	15.3	33	45.8	61.1
42	12	16.6	1	1.4	18.0
43	2	2.8	5	6.9	9.6
44	13	18.1	0	0.0	18.1
45	12	16.6	3	4.2	20.8
46	14	19.4	7	9.7	29.1
47	11	15.3	4	5.5	20.8
48	40	55.6	11	15.3	70.9
49	38	52.8	11	15.3	68.1
50	39	54.2	10	13.8	68.0
51	37	51.4	7	9.7	61.1
52	38	52.8	10	13.8	66.6
53	38	52.8	11	15.3	68.1
54	38	52.8	10	13.8	66.6
55	38	52.8	10	13.8	66.6
56	38	52.8	12	16.6	69.4
57	37	51.4	10	13.8	65.2
58	38	51.4	9	12.5	63.9
59	39	54.3	12	16.6	70.8
60	37	51.4	8	11.1	62.5
61	39	54.2	8	11.1	65.3
62	39	54.2	8	11.1	65.3
63	0	0.0	0	0.0	0.0
64	9	12.5	4	5.5	18.0
65	18	25.0	10	13.8	28.8
66	34	47.2	6	8.3	55.5
67	6	8.3	39	54.2	62.5
68	4	5.5	26	36.1	41.6
69	8	11.1	27	37.5	48.6
70	2	2.8	4	5.5	8.3
71	29	40.3	3	4.2	44.5
72	9	12.5	5	6.9	19.4

such as newsteadi has a low total number of sensilla (mean = 8.5 sensilla/antenna; s.d.=1.08) compared to the large number of sensilla on the bigger species such as orisescens (mean = 12.0 sensilla/antenna; s.d.=1.09). The variances in the total number of sensilla per antenna does not differ significantly between these two species.

Ratios involving the proboscis (variables 67,68,69) have a large number of negative correlations (36% - 54%), again mostly with size. Most of the other ratios: antennal ratio (65), cibarium/pharynx ratio (70), mandible/maxilla ratio (71) and the costal ratio (72) show little affinity with other characters. In contrast, the palp ratio (66) shows a large number of positive correlations with size variables. These correlations reflect the interspecific variation in shapes of the third palp segment (see Figs 11-14 in Section 3). In C. newsteadi (a small species), the third palp segment is swollen and hence the palp ratio is low (mean = 2.28; s.d.=0.26). In larger species such as C. orisescens or C. delta, the third palp segment is slender with a corresponding large palp ratio (means = 3.82; s.d.=0.64 and 3.10; s.d.=0.32 respectively).

The majority of characters describing wing pattern (35 - 47) show little correlation with other types of characters, but a few (38 - 41) do show some negative correlations with size. These last four characters describe the pigmentation around the medial veins of the wing, emphasising a taxonomic distinction between small species, such as newsteadi and impunctatus, and large species such as pulicaris and delta. The differences are summarised in Table 30.

TABLE 30.
Table of mean values for four wing pattern characters in two large and two small species

Character	38	39	40	41
impunctatus	3.60	4.30	1.80	2.87
newsteadi	4.25	4.50	2.66	3.00
pulicaris	2.70	2.90	1.60	1.50
delta	3.00	3.00	1.63	2.00

A few characters are typified by a very low number of significant correlation coefficients. A typical example is character 64

(presence/absence of cibarial teeth), which is only present in C. fagineus and for which only 18% of correlations are significant.

Results of Cluster Analysis of Variables

The results of single linkage cluster analysis of the between-character correlation matrix are presented as a dendrogram (Fig. 60).

Basically, at a phenon level of 0.45, four distinct clusters are evident. Reading from the top of the dendrogram, the first group (48 - 18) is the largest and is composed mainly of size related characters with a few antennal characters (10,11,16,17,18). The straggly nature of this cluster is a typical feature of the technique of single linkage cluster analysis. The second cluster (69 - 9, reading downwards) is composed of the proportional lengths of antennal segments and the three proboscis ratios (67,68,69). Cluster three, (65 - 14, reading downwards) is composed of four antennal characters, and the last cluster (41 - 44, reading downwards) is composed entirely of wing pattern characters.

One of the first problems encountered in interpreting cluster analysis is the recognition of clusters. The relationship between the number of clusters present at each level of similarity for these data is shown in Fig. 61. This approximates to a shallow logistic curve.

The level of similarity chosen, and consequently the number of clusters present, is a subjective decision, based on the final number of characters required and the exact structure of the dendrogram. The general structure of the dendrogram has already been discussed. At a phenon level of 0.45, four distinct clusters are present, together with five single-character clusters (outliers): 1,3,43,63,70. At a similarity level greater than 0.45, the proportion of single-character clusters increases considerably. With the inclusion of outliers, a total of nine variables would be selected at a similarity level of 0.45.

Once the problem of specifying a cluster is resolved, the next problem is the selection of one character from each cluster. Among the possible ways of selecting a variable are:

1. Choose the last variable to join a cluster (outer clustering).
2. Choose one of the innermost members of a cluster (inner clustering).
3. Choose one of the variables at random.

Analysis of real data by Jolliffe (1973) showed inner clustering

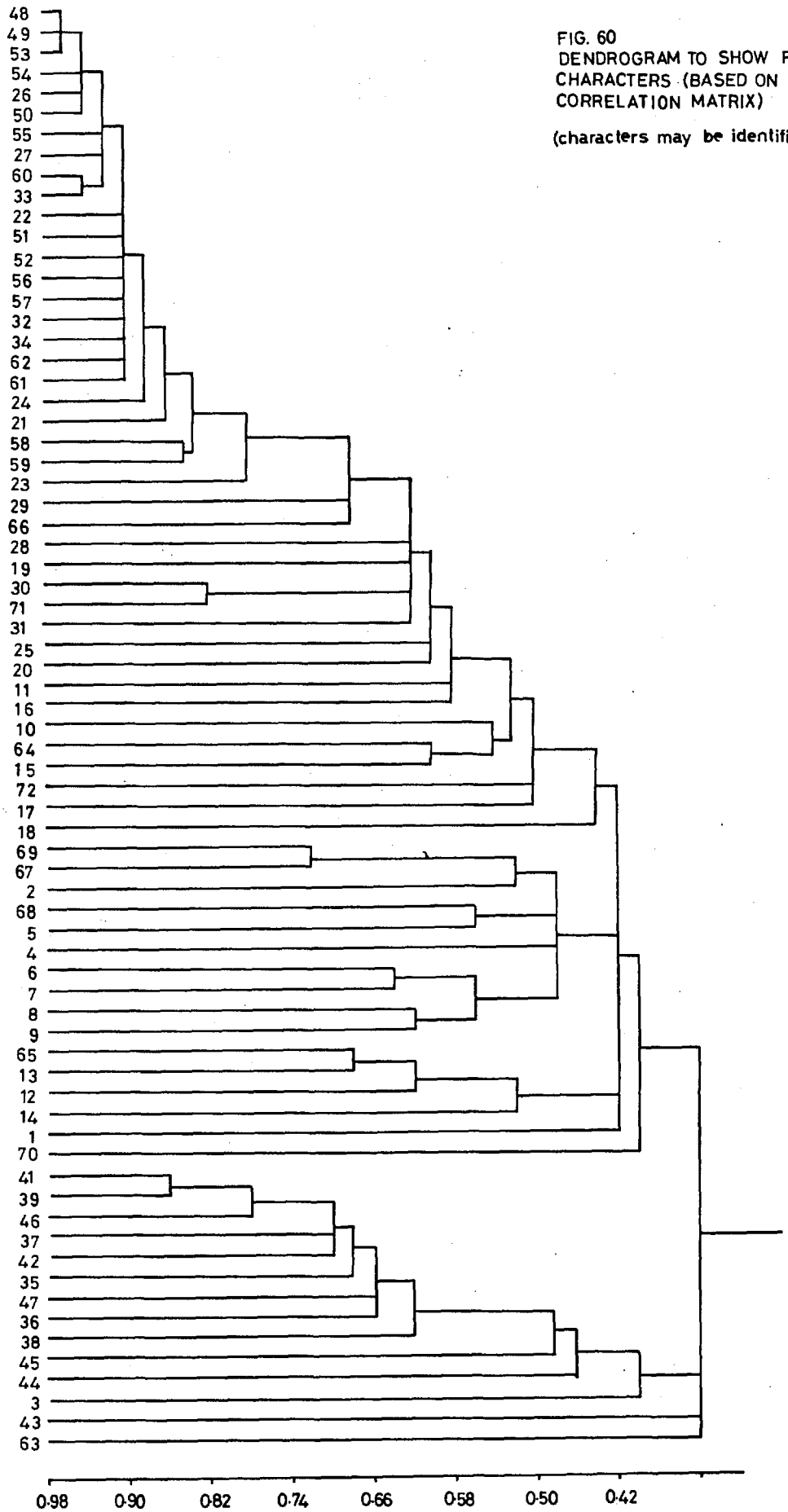


FIG. 60
DENDROGRAM TO SHOW RELATIONSHIPS OF CHARACTERS (BASED ON BETWEEN-CHARACTER CORRELATION MATRIX)

(characters may be identified from table 21)

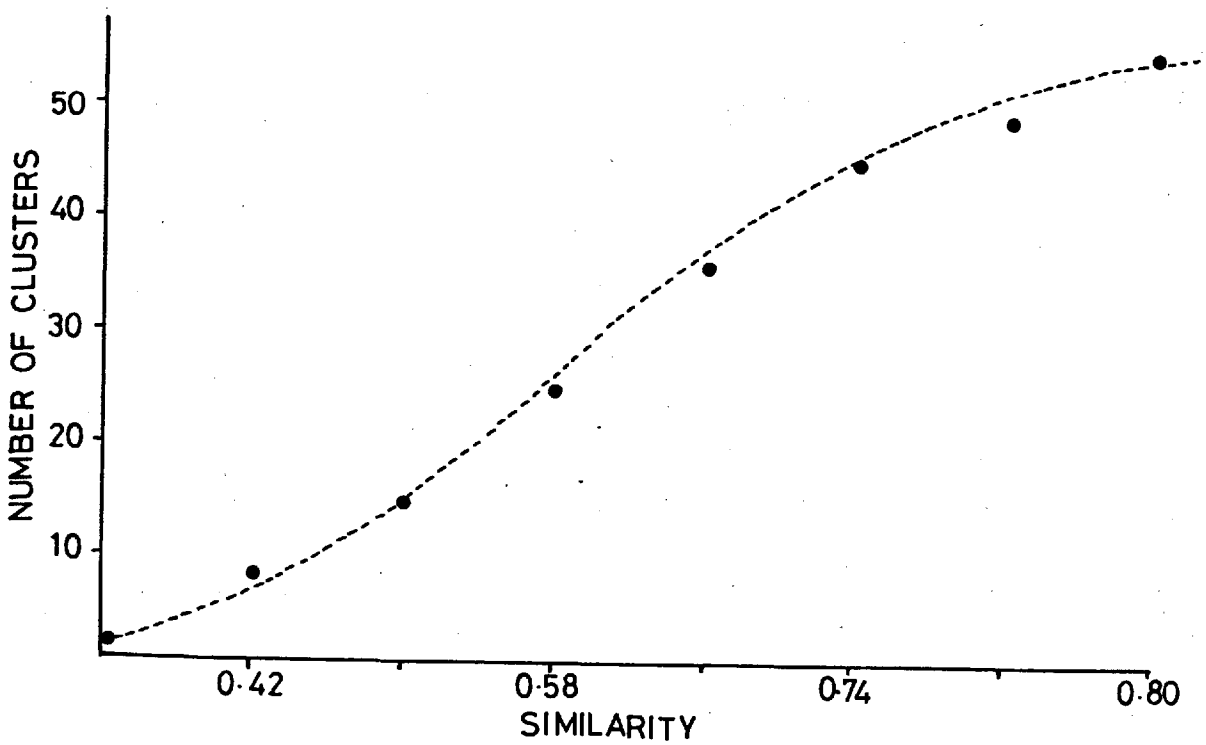


FIG. 61
NUMBER OF CLUSTERS AT EACH LEVEL OF SIMILARITY

to be the most effective and was the method used here. From the innermost pair of characters, one was selected at random to give the following set:

Character

53 from the general size cluster.

67 from the cluster of proximal antennal and proboscis characters.

65 from the cluster of distal antennal characters.

41 from the wing pattern cluster.

The total number of characters, including outliers therefore is:

Character

53 Length of mid femur

67 Head length/proboscis length

65 Antennal ratio

41 Wing pattern element 7

1 Eye contiguity

70 Cibarium/ pharynx ratio

3 Relative length of antennal segment iv

43 Wing pattern element 9

63 Setae in hind tibial comb

Before this selection is employed to produce a classification of the 84 specimens, the results of the principal coordinate analysis should be considered.

Results of Principal Coordinate Analysis of Between-Character Distance Matrix

In a principal coordinate analysis of 72 variables, the first three eigenvalues account for 32%, 14%, and 7% of the total variance respectively. The three principal coordinate axes associated with the eigenvalues are plotted in Fig. 62. This diagram shows a smaller number of groups than the cluster analysis, but more outliers. However, examination of the first and second axes shows the general grouping of the variables to be very similar to the cluster analysis, with respect to size measurements and wing pattern, but the antennal characters are more widespread. Visual inspection of these two axes (which account for approximately 50% of the trace) shows the characters to be grouped as follows:

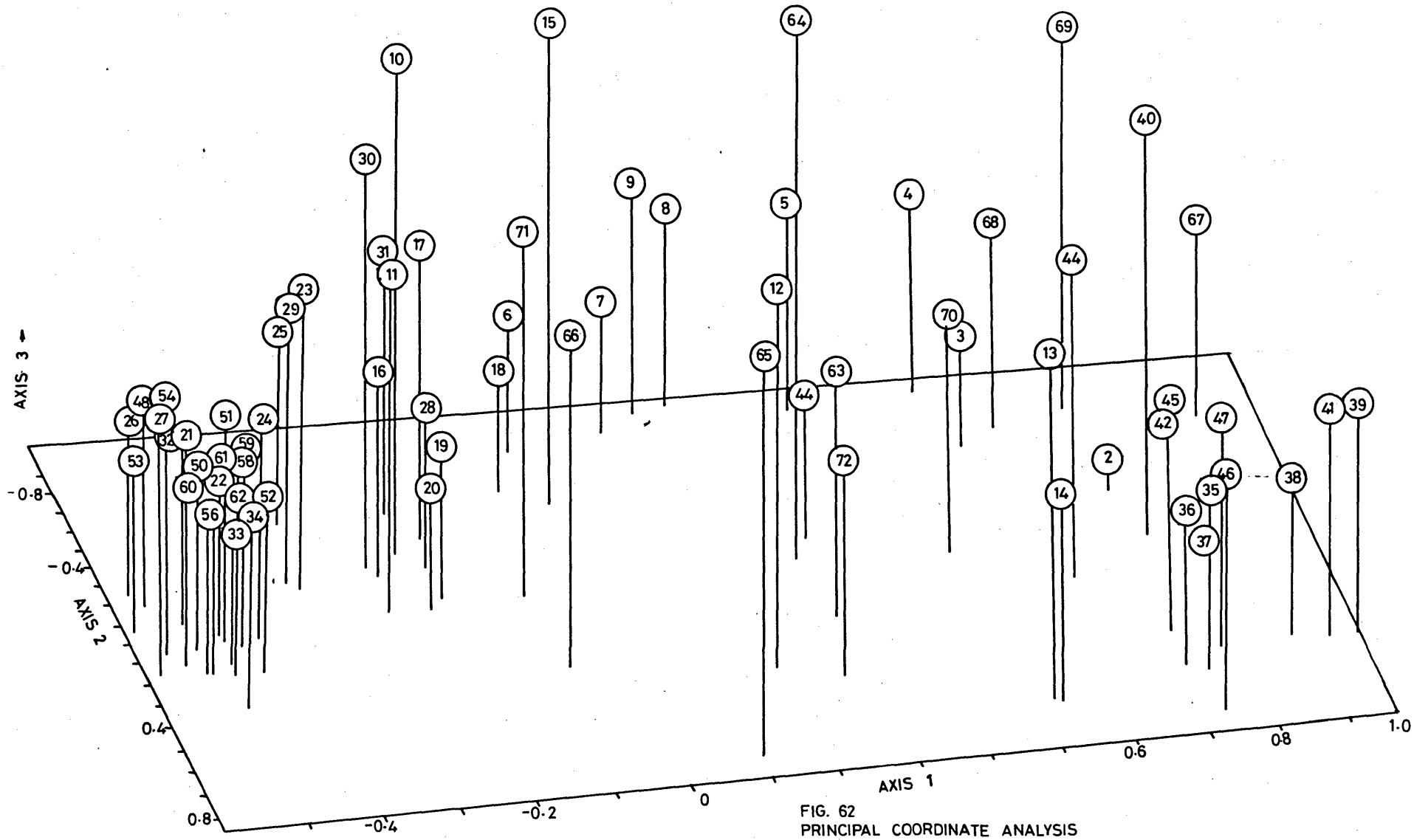


FIG. 62
PRINCIPAL COORDINATE ANALYSIS
PLOT SHOWING RELATIONSHIP OF CHARACTERS
(BASED ON BETWEEN-CHARACTER DISTANCE MATRIX)

Group identification	Variables
Wing pattern	35,36,37,38,39,41,42,45,46,47.
Size	10,11,16,17,19,20,28,30,71; 23,25,29; 21,22,24,26,27,31,32,33,34,48,49,50,51, 52,53,54,55,56,57,58,59,60,61,62.
Proboscis ratios	67,68,69.
Proximal antennal segments	5,6,7,8,18.
Distal antennal segments (part)	13,14.
Sensilla ant. seg. iii + cibarium	15,16.
Ungrouped	2,3,4,65,66,12,40,43,44,63,70,72.

The two characters 15 and 64 form a group which defines C. fagineus and is the only one with such a clear taxonomic interpretation.

This grouping suggests that a total of 19 characters summarise the 72 characters, if all the outliers are included. The most central character from each group was selected as a representative in an analogous manner to inner clustering(used for selecting variables in the cluster analysis). The 19 characters suggested are: 1,2,3,4,7,11,12,14,40,43,44,45,60,63,65,66,69,70,72.
This is considerably more characters than the nine suggested by the cluster analysis approach, in which an arbitrary phenon level of 0.42 was provisionally selected. In retrospect, it was probably too optimistic to expect only nine characters to classify these difficult species. The results of the cluster analysis may now be examined again, to see which characters are suggested, if a total of 19 variables are to be used. Recourse to Fig. 61 shows that for 19 clusters to be recognised (1 variable from each cluster), a phenon level of 0.52 should be used. This level is remarkably similar to the figure of 0.55 which Jolliffe (1972) found by empirical means, to give the most reasonable selection of variables. Davies & Boratynski (1979) approached this rather differently. They first decided that about 25 out of 101 characters would be a suitably sized subset, and then looked for a phenon level that would yield that number.

The 19 characters suggested by cluster analysis and principal

TABLE 31.

Subset of Characters Suggested by Cluster Analysis and Ordination of Between-Character Distance Matrix

Ordination	Cluster Analysis
OTU 1	OTU 1
2	2
3	3
4	4
7	7
11	-
12	12
14	-
40	40
43	43
44	44
45	45
60	-
63	63
65	-
66	-
69	69
70	70
72	72

Variables may be identified from Table 21.

coordinate analysis of the between-character/distance matrix are summarised in Table 31. On the whole, there is a good correspondence between the selections, with only two exceptions. Firstly, ordination shows three distinct characters (11,60,66), which are then grouped together (general size cluster) by cluster analysis. The second difference presents a similar case in which principal coordinate analysis shows variables 12,14, and 65 to be well separated, but cluster analysis groups them together. The two differences may be due to properties of the techniques themselves. Cluster analysis represents close associations clearly, whereas ordination summarises distant associations more accurately. The outliers are therefore more reliably identified by ordination. As the differences between the variables selected by each method is in the outliers, the 19 variables suggested by principal coordinates is accepted.

To test the effectiveness of these 19 variables in classifying members of the C. pulicaris complex, they were used in a principal component analysis, composed of a sample of 84 specimens.

Results of Classification Based on a Subset of 19 Variables, Selected by Character Correlations

A summary of the results of this analysis is given in Table 32 and a plot of the first three components is given in Fig. 63 . A discussion on the identifications will be given later in this section.

The first principal component describes some 19% of the total variance, rising to 32% in two dimensions and 44% in three. This is a relatively low proportion of the total variance to be described in three dimensions, certainly much smaller than is generally found in the literature. The relative sizes of the eigenvalues suggests that the data form an elongate spheroid in hyperspace.

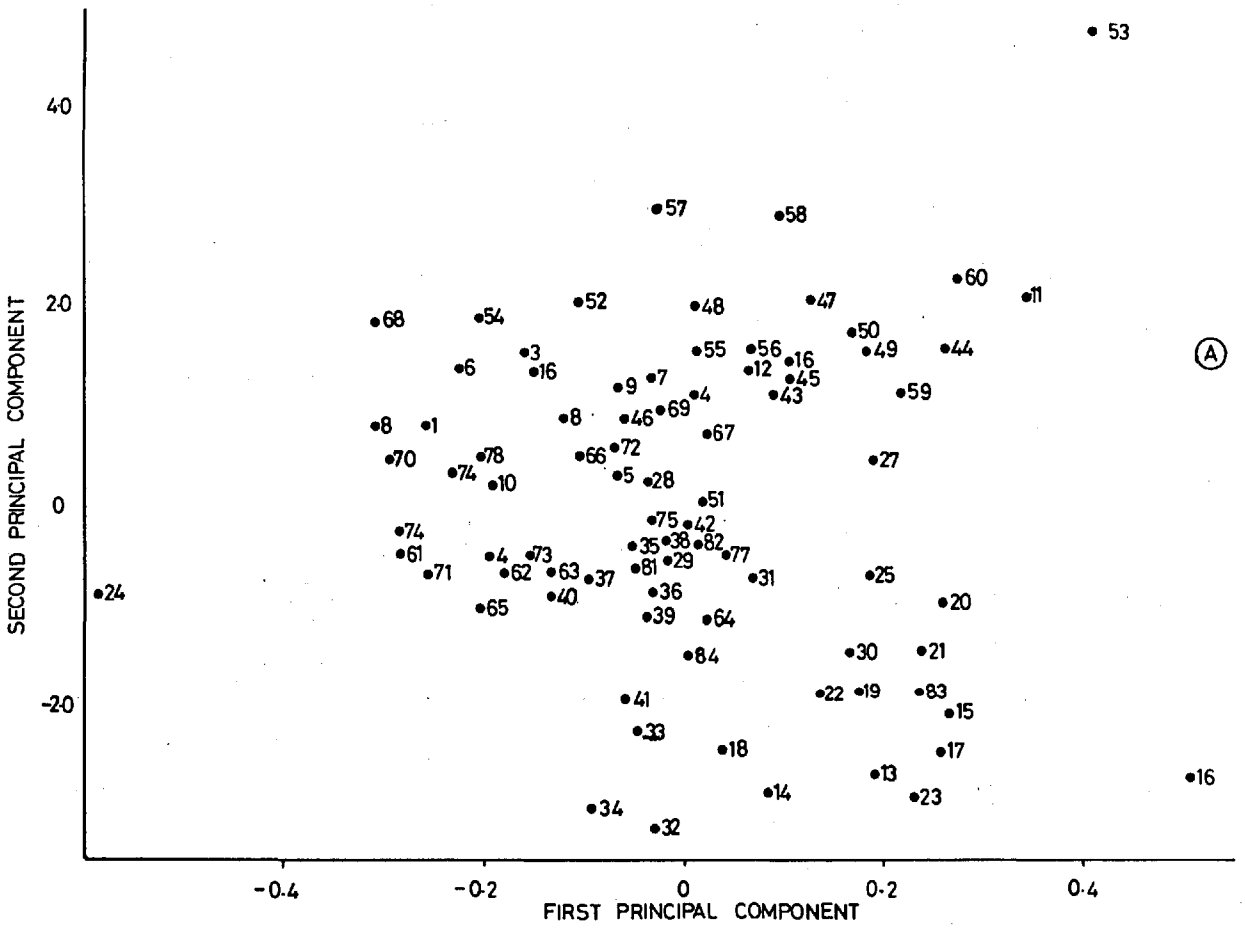
The plot of the first three components (Fig. 63a) shows that most OTU's form one large heterogeneous clump, with little distinction between taxonomic groups. Within this large group, which is similar in shape in both plots, some of the species form monospecific clusters, for example, grisescens (13 - 23) and species 'a' (52 - 58). However, none of these clusters are well separated. It is clear that the 19 variables suggested by correlation of

TABLE 32.

Summary of a Principal Component Analysis, Using a Subset of 19 Variables Selected by Between-Character Correlations

	1	2	3	4	5
Eigenvalues	3.616	2.587	2.290	1.571	1.393
Cumulative Percentage of Trace	19.03	32.65	44.70	52.96	60.29
		Eigenvectors			
Variables (Code numbers)					
1	-0.121	0.205	-0.231	0.475	-0.202
2	-0.279	0.004	0.246	-0.054	-0.242
3	-0.331	-0.136	0.112	-0.063	-0.256
4	-0.354	-0.173	0.051	0.080	-0.139
7	-0.267	-0.176	0.052	-0.223	0.502
11	0.183	0.125	-0.347	-0.111	0.151
12	0.287	0.182	-0.176	0.108	0.008
14	0.233	-0.002	0.208	-0.289	-0.396
40	0.130	0.185	0.415	0.250	0.233
43	-0.022	0.259	0.087	-0.491	0.059
44	0.127	-0.127	0.433	0.260	0.131
45	0.202	0.205	0.419	-0.029	0.198
60	0.183	-0.409	-0.231	-0.089	0.208
63	0.150	-0.232	0.214	-0.150	-0.122
65	0.462	0.129	-0.129	-0.624	-0.168
66	0.217	-0.438	0.028	0.268	0.009
69	-0.063	0.481	-0.313	0.268	0.077
70	0.011	-0.002	0.093	0.062	0.321
72	0.187	-0.162	0.045	0.305	-0.271

Variables may be identified from Table 21.



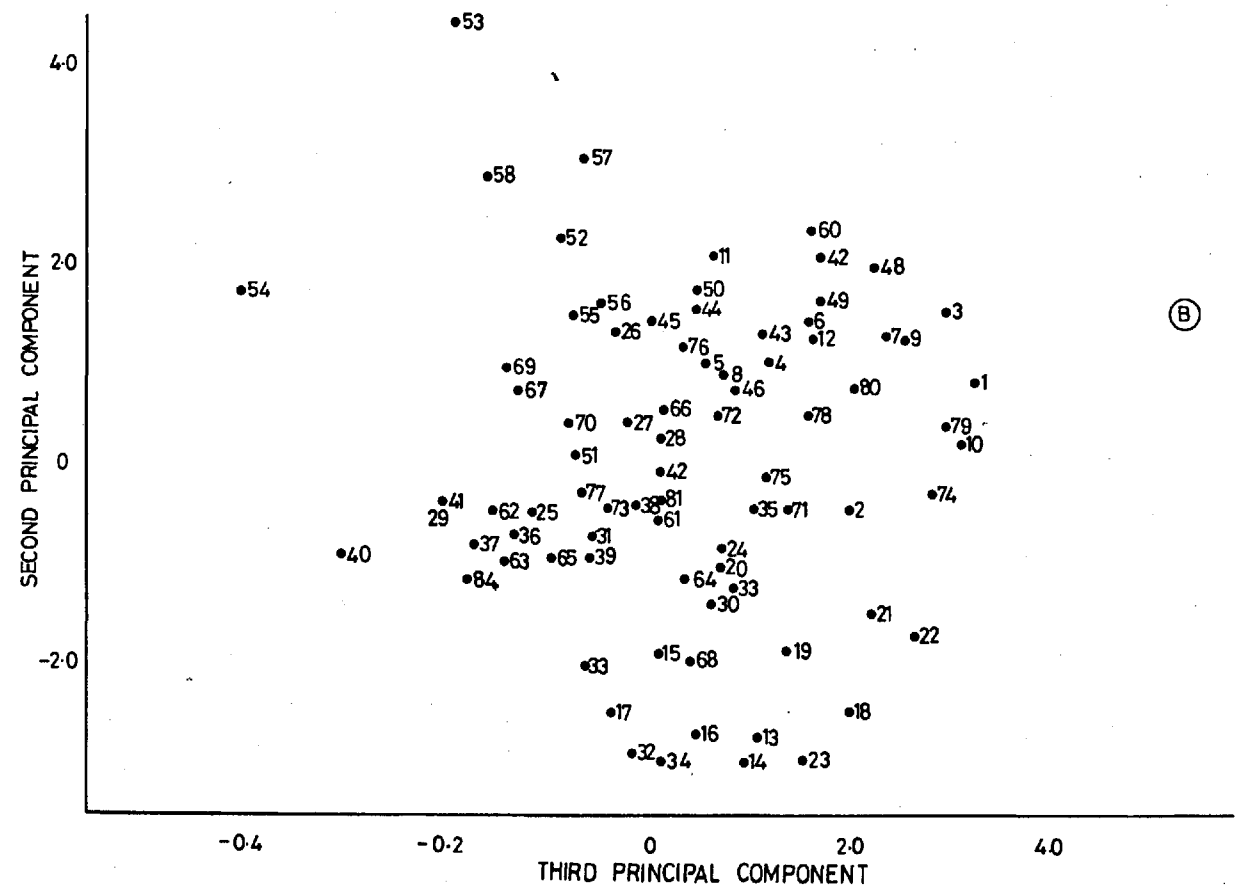


FIG. 63
CLASSIFICATION BASED ON 19 VARIABLES SELECTED BY CHARACTER-CORRELATION

characters does not lead to a clear classification of the species in the pulicaris group (if indeed there is such an arrangement).

It is worth noting that the first principal component does not separate specimens according to their size, as is usually the case. The loadings associated with each of the 19 variables on the first eigenvector are given in Table 32. The five largest loadings are associated with antennal characters. In decreasing size, they are: antennal ratio, proportional lengths of antennal segments v, iv, xiii, and iii. Frequency histograms of the loadings appropriate to each variable on the first five eigenvectors were constructed (Fig. 64). By definition, the average value of the loading must be $\frac{1}{\sqrt{19}}$ i.e., 0.229. For the first eigenvector, the loadings are clustered around the class containing the expected average value (0.20 - 0.25). Fig. 64 also emphasises the importance of the antennal ratio on this vector, by showing its loading, + 0.46, which is considerably larger than others.

The second eigenvector is dominated by a number of ratios. The five largest loadings in descending size are associated with the following variables: antennal length/proboscis length; palp ratio; length of hind tarsus i; proportional length of antennal segment iii, and finally, setae on hind tibial comb. Most of the loadings are absolutely smaller than the expected average 0.23 (fig. 64). This axis exhibits some separation of specimens according to their size, presumably the influence of the hind tarsus length. However, this ordering by size is minimal, and may be coincidental. The first three loadings are considerably larger than the rest (Fig. 64).

The third principal component separates the specimens mainly on their wing pattern. The three highest loadings are associated with wing pattern elements 10, 11, and 6, and here again, are significantly larger than any other loadings (Fig. 64). However, as found in the previous section, wing pattern does not provide the best set of characters for a clear classification of members of the pulicaris group.

The third antennal segment (proportional length) is one of the variables with the largest loadings on all the first three eigenvectors, suggesting that it is of considerable taxonomic importance.

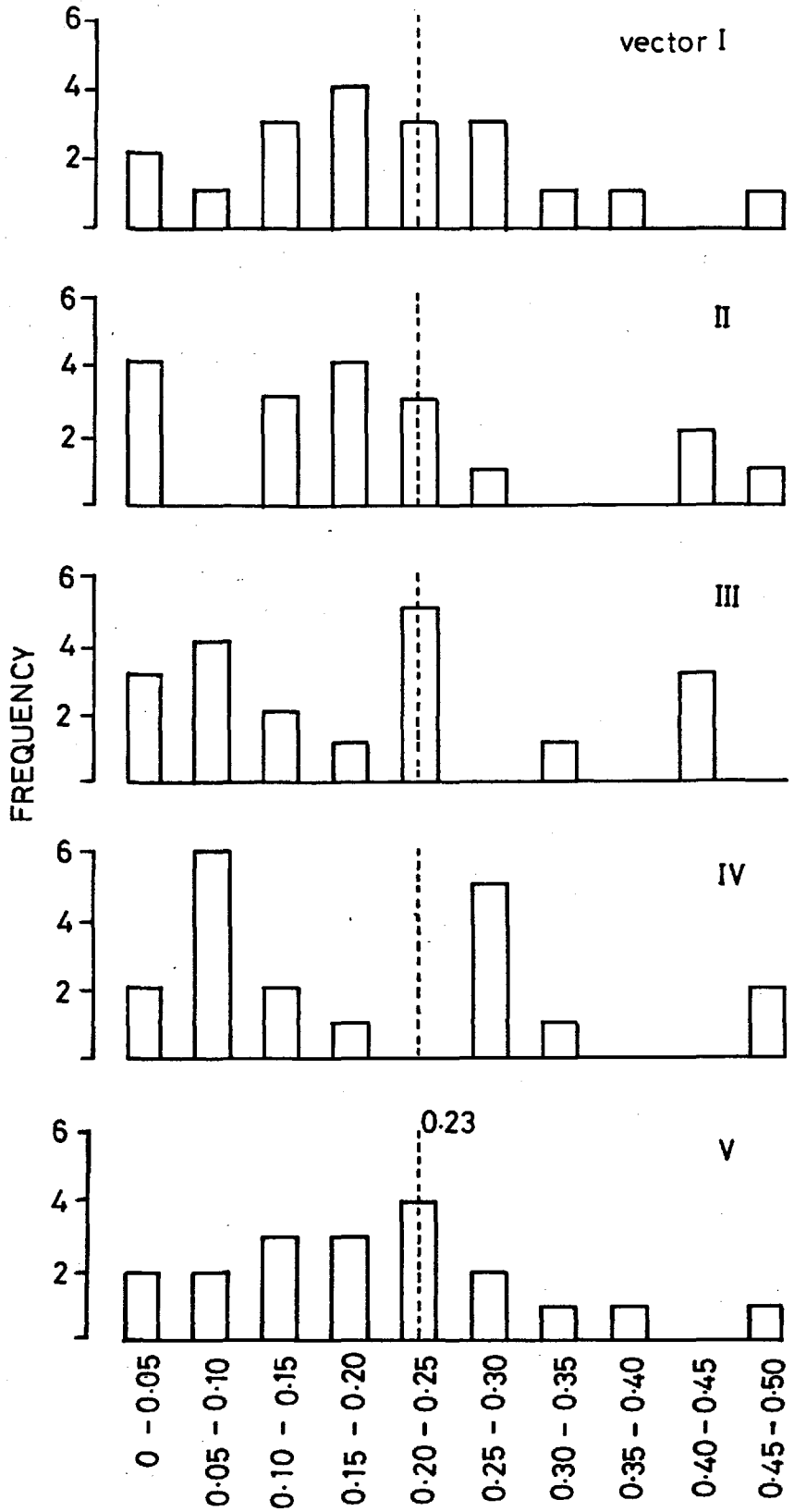


FIG. 64
FREQENCY HISTOGRAMS OF VECTOR LOADINGS
(FROM ANALYSIS BASED ON 19 CHARACTERS)

9.4.5. Discarding Variables Using Principal Component Loadings

The second numerical method employed here, to choose a subset of variables, uses the principal components themselves. The technique used here is a variation on Jolliffe's method B4, in which he associated one variable with each of the p largest principal components and then rejected the (K-p) variables.

In the principal component analysis based on the complete set of 72 variables (described on p. 215), the first five components accounted for 63% of the total variance. The next five components only added a further 12%. Axes four to ten absorb between 2% and 4% each - a rather low proportion. To associate a single variable with each of these small components, as suggested by Jolliffe (1972), may give rather misleading results, because the likelihood of random effects influencing them is rather high. This is a particularly important aspect when a subset of 20 variables is required. Therefore it was decided to concentrate on the first few components, by selecting those variables associated with the five largest loadings on each axis, until a total of 20 variables were obtained. The five variables with the highest loadings on each of the first four axes are given in Table 33 . It is of considerable interest, and convenience, that these axes are associated with different sets of variables, and no one variable receives a high loading on more than one eigenvector. This set of 20 variables is the same as that selected from the analysis based on 64 variables (i.e., 72 variables less 8 ratios), with the exception of one variable - the antennal ratio. This character has the highest loading on the third vector in the analysis based on 72 variables, but obviously was absent from the analysis based on 64 variables. In both analyses, the third component was indentified as a vector describing the relative proportions of the distal and proximal sections of the antenna. Because the antennal ratio measures this contrast, there would be no loss of information by omitting it and using the individual segments of the antenna instead.

A principal component analysis was therefore performed on 84 OTU's, using the 20 variables described in Table 33.

Results

A plot of the first two principal components is given in Fig. 65.

TABLE 33.

Variables with the 5 Largest Loadings on the First Four Eigenvectors

Eigenvector	% Variance of Eigenvector	Variable Code No.	Loading	Description
I	33.5%	53	-0.198	Length of mid leg femur
		26	-0.196	Length of head
		48	-0.196	Length of fore-leg femur
		49	-0.196	Length of fore-leg tibia
		60	-0.196	Length of hind metatarsus
II	13%	46	0.292	Wing pattern element 12
		39	0.273	" 5
		41	0.272	" 7
		38	0.250	" 4
		42	0.243	" 8
III	12%	6	0.300	Proportional length of antennal seg. vii
		5	0.295	" vi
		7	0.289	" viii
		13	-0.277	" xiv
		4	0.241	" v
IV	5%	64	0.392	Cibarial teeth
		15	-0.341	No. sensilla on antennal seg. iii
		40	-0.309	Wing pattern element 6
		2	0.241	Length of antennal segment iii
		10	-0.214	" xi

Variables may be identified from Table 21.

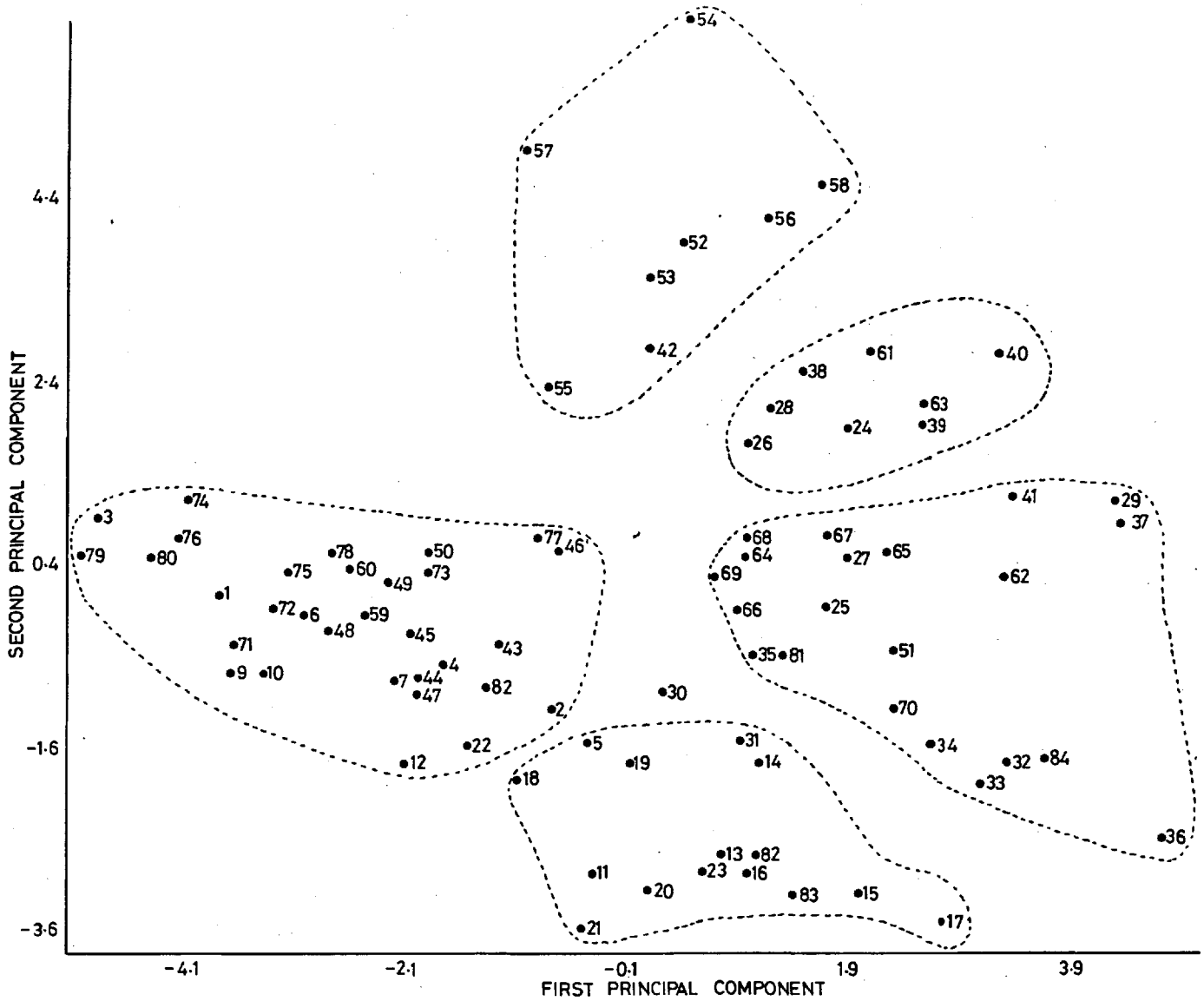


FIG. 65

CLASSIFICATION BASED ON 20 VARIABLES SELECTED BY VECTOR LOADINGS METHOD

and the five largest eigenvalues and their associated eigenvectors in Table 34 .

Overall, the arrangement of OTU's is rather similar to both the classification based on the complete set of variables (Fig. 57), and that based on 64 variables (Fig. 59). Because of this similarity, little need be said on the taxonomic aspects, other than that most taxa are less distinct. This detracts little from the overall similarity between the results, however.

The first vector is influenced by two contrasting classes of characters, size and wing pattern. Size variables have positive loadings and wing pattern characters have negative loadings. The relative importance given to these groups of characters is very similar to the analysis based on the full character set. The second vector is also dominated by these characters but, in contrast to the first axis, all these characters have negative loadings. Wing pattern characters have the largest loadings on this eigenvector, as found in the reference analysis. Antennal characters clearly dominate the third component by their large loadings, also found in the analysis based on a larger number of variables. It is most interesting to note that when a small set of variables was used, the same characters were found to influence the same vectors as an analysis of a much larger number of variables.

The first component of this anlysis absorbs 28% of the total variance, only slightly less than the 33% in the analysis based on 72 characters. The difference between these values is surprisingly small, considering that one analysis used only a quarter of the variables of the other. Two variables, 14 and 64, have very small loadings on the first three components, which account for 64% of the trace, suggesting that the number of variables could be further reduced by removing these two.

In conclusion, the classification produced by the set of variables with high loadings in a principal component analysis is far superior to that produced by variables selected by cluster analysis analysis of characters. The classification is superior both in terms of minimal distortion in the arrangement of OTU's, and in its taxonomic application.

In a survey of the techniques for reducing the number of variables in principal component analysis, Davies & Boratynski (1979) also found the vector method to be very effective.

TABLE 34.

Summary of Principal Component Analysis Using 20 Variables Selected be Vector Loading Method

	1	2	3	4	5
Eigenvalue	5.767	3.969	3.106	2.130	1.046
Cumulative percentage of trace	28.84	48.68	64.21	74.86	80.09
Variable	Eigenvectors				
2	-0.182	-0.027	-0.255	-0.339	0.031
4	0.023	0.085	-0.397	-0.022	-0.565
5	0.064	0.028	-0.449	0.112	-0.247
6	0.147	-0.113	-0.390	0.135	0.224
7	0.085	-0.039	-0.391	0.142	0.495
10	0.210	0.089	0.192	0.313	-0.115
13	-0.058	-0.055	0.416	-0.105	-0.232
14	0.064	0.069	0.152	0.492	0.234
26	0.381	-0.127	0.016	-0.057	-0.036
38	-0.288	-0.237	0.033	-0.011	0.260
39	-0.219	-0.369	-0.067	0.088	-0.095
40	-0.215	-0.158	0.080	0.345	0.034
41	-0.219	-0.370	-0.060	0.129	-0.142
42	-0.160	-0.384	-0.070	0.056	-0.132
46	-0.219	-0.352	0.067	0.026	-0.040
48	0.324	-0.291	0.035	-0.065	-0.045
49	0.331	-0.281	0.029	-0.044	-0.012
53	0.343	-0.263	0.026	-0.071	0.002
60	0.312	-0.281	0.056	-0.085	0.019
64	0.069	0.076	-0.049	0.558	-0.279

Variables may be identified from Table 21.

In addition, they found one of the clustering methods to be similarly effective (single linkage, based on simple matching coefficients for multistate data), whereas the other clustering method (single linkage of correlation matrix) gave very poor results. The conclusions of the present study concur well with those of Davies & Boratynski.

9.5. SELECTION OF VARIABLES BY SUBJECTIVE CRITERIA

In contrast to the objective methods used above, this section attempts to find an effective subset of variables by subjective or intuitive means.

The selection of the subsets of characters was made, using the results of previous studies on character variation (Section 7), together with a knowledge of those characters thought to be important in the classification of this complex.

Between 15 and 20 variables were used so that the classifications produced from intuitively selected characters could be compared to the classifications based on objectively chosen characters. In this approach, trial and error is an important component and inevitably leads to many trial classifications being tested. Of the many trials undertaken, three have been selected and are described here, to represent the range of results obtained.

The majority of trials were first carried out on 53 specimens representing the 8 taxa. Details of these specimens are given in the appendix. However, it was thought that this number of specimens was insufficient to determine whether clear boundaries existed between species, or if the observed discontinuities were an artefact from using small samples. The number of specimens was therefore increased to 84, making a primary data matrix of order 84 x 72 and incorporating over 6,000 measurements.

In each experiment, the relationships between the OTU's were summarised by principal component analysis of a correlation matrix. The eigenvectors were scaled such that the sum of the squared loadings was equal to unity. This enabled the relative sizes of loadings to be compared for different axes. As in the previous part, the specimens were provisionally identified so that the taxonomic effectiveness of a combination of variables may be readily ascertained. Indications of possible misidentifications were noted and they are discussed later in this section.

9.5.1. Selection 1

Seventeen variables were used in this experiment, many of which were size variables. They were selected because most were associated with large loadings in one or more of the principal

component analyses in the previous section. They are:

Variable	
1	Contiguity of eyes
2	Proportional length of antennal segment iii
3	" iv
4	" v
5	" x
6	" xi
17	Number of sensilla on antennal segment xii
18	" xiii
31	Number of mandibular teeth
34	Wing pattern element 1
35	Length of costa
51	Length of fore tarsus ii
53	Length of mid femur
54	Length of mid tibia
59	Length of hind tibia
61	Length of hind tarsus ii
63	Number of setae on hind tibial comb

The selection of variables was first applied to a sample of 53 OTU's representing eight taxa. Details of the specimens are given in the appendix, where OTU's are indicated by an asterisk*.

Although the first axis accounted for only 24% of the total variance and the second a further 12%, these two axes summarised the relationships between the species quite well (Fig. 66). A total of 37% of the trace absorbed by the first two dimensions is lower than usually recorded in the literature. The relative sizes of the eigenvalues (Table 35) show that most of the variation is in the first two dimensions, indicating that the OTU's are distributed approximately as a plane in the hyperspace.

A plot of the first two axes is given in Fig. 66 . The main factor affecting the first axis appears to be size and, as expected, the two small species, impunctatus (40 - 49) and newsteadi (1 - 4) are well separated from the other larger species. Within the larger species, most of the taxonomic variation is along the second axis, where the specimens fall into three main groups. Those of grisescens (5 - 9) form a coherent cluster at the top of the diagram with two specimens of lupicaris (52, 53) and a specimen from Lundy Island (29) provisionally identified as punctatus (the status of this

TABLE 35.

Percentages and Cumulative Percentages of the Total Variance Associated with the Eigenvalues in a Principal Coordinate Analysis. Analysis of 53 OTU's Based on the 17 Variables in Subjective Selection 1.

Eigenvector	Percentage Variance	Cumulative Percentage Variance
1	24.70	24.70
2	12.69	37.39
3	6.60	43.99
4	5.97	49.96
5	4.50	54.46
6	4.28	58.74
7	3.42	62.16
8	3.12	65.28
9	2.86	68.14
10	2.72	70.86

Variables may be identified from Table 21.

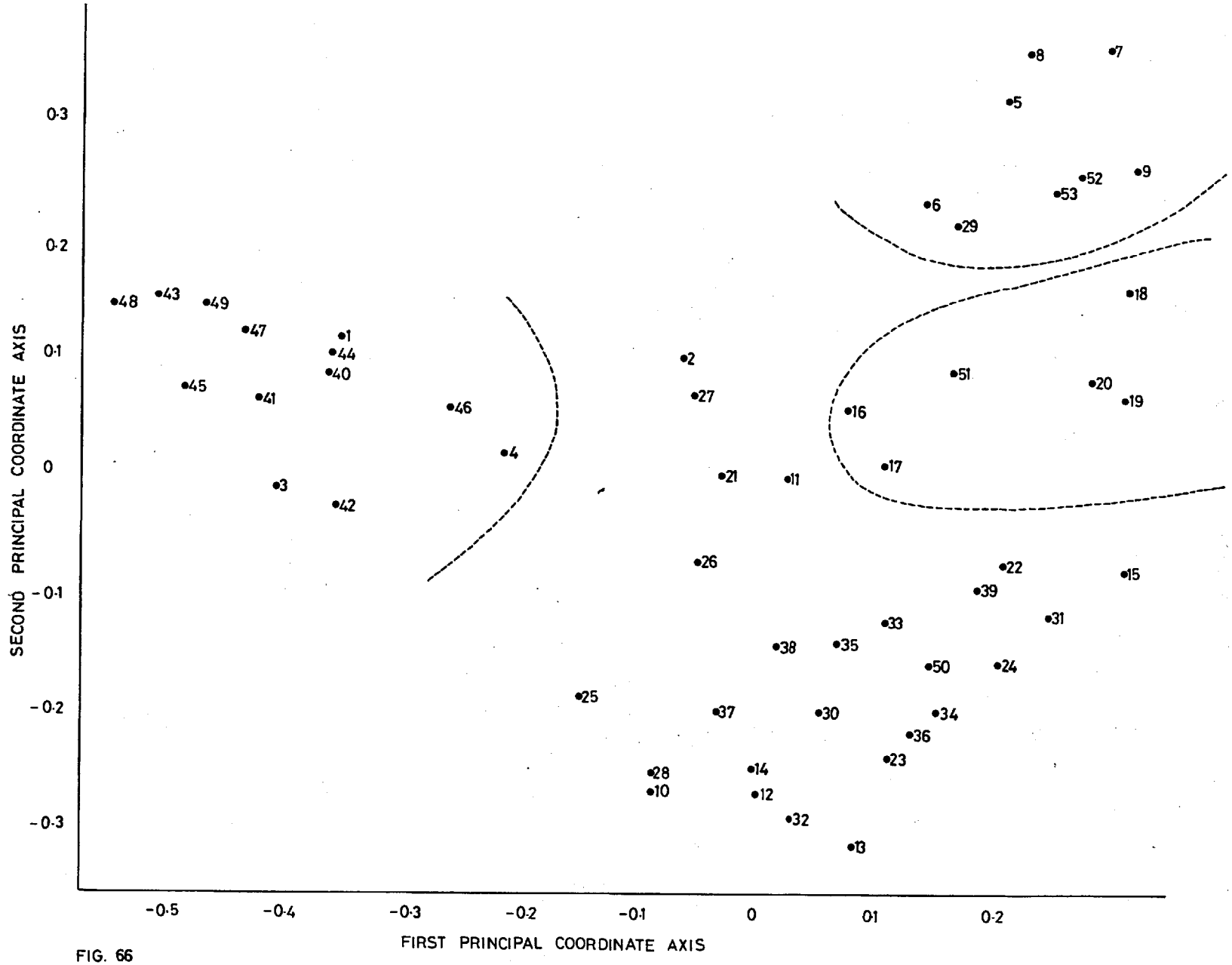


FIG. 66
PRINCIPAL COORDINATES PLOT OF 53 OTU'S
SUBJECTIVE SELECTION 1

specimen is discussed below). The middle group consists mainly of delta (16 - 20) and of a specimen of lupicaris (51). The two specimens of punctatus (21, 27) and a specimen of fagineus (11), placed close to delta, are well separated in the third dimension. The remaining specimens form a large heterogeneous cluster containing pulicaris (30 - 39), punctatus (21 - 29) and fagineus (10, 12 - 14). The specimens of fagineus (readily identified by the presence of cibarial teeth) lie at the periphery of this large group, and form a reasonably discrete group in the third dimension.

The taxonomic relationships of both the species and their constituents are relatively clearly demonstrated, considering the nature of the problem of the pulicaris species group. However, this situation changed when the number of specimens was increased from 53 to 84. Details of the 84 specimens may be obtained from Table 27 . The percentage of the total variance described by the first three eigenvectors increased to 36%, 13%, and 11% respectively, for this larger data set (Table 36), giving a total of 60%. Although the first two principal components are more significant in a statistical sense than the analysis based on 53 specimens, the grouping of the OTU's on these axes is much less obvious. This result typifies the problems encountered in the pulicaris complex, outlined in the introductory section : if only a few examples of each species are considered, then the complex falls into more or less distinguishable groups, but as the number of specimens is increased (including intermediate forms) the complex becomes broad and diffuse and shows few clear boundaries between species.

The first two principal components of the analysis, based on 84 OTU's are plotted in Fig. 67, and again, size is the predominant factor affecting the first component. The largest loadings on the first eigenvector (Table 36) are associated with characters which are simple size measurements. The second component is not so readily interpreted. The characters with the five largest loadings represent very different types of character - eye contiguity, proportional length of antennal segment xi, number of mandibular teeth, wing pattern element 1, spines on the hind tibial comb - in fact, most character types other than size.

The relative positions of the taxa do not differ significantly in this plot from that based on 53 OTU's, only the extent of overlap between them which, as already noted, is very much greater.

TABLE 36.

Eigenvalues and Eigenvectors from a Principal Component Analysis Based on the Variables in Subjective Selection 1

	1	2	3	4	5
Eigenvalue	6.159	2.266	1.905	1.098	1.003
Cumulative Percentage of Trace	36.23	49.56	60.76	67.22	73.12
Variable			Eigenvectors		
1	-0.099	-0.357	0.042	-0.527	-0.429
2	-0.184	0.235	0.300	0.140	-0.257
3	-0.132	0.095	0.459	-0.035	-0.201
4	-0.074	0.093	0.577	0.093	-0.032
9	-0.007	-0.295	0.281	0.397	0.263
10	0.152	-0.449	-0.226	-0.120	-0.045
17	0.161	-0.229	0.044	0.212	0.268
18	0.116	-0.046	0.423	-0.459	0.219
31	0.170	-0.302	-0.001	0.379	-0.376
34	0.364	0.144	-0.006	-0.147	-0.160
35	-0.042	0.428	-0.107	0.185	-0.405
51	0.368	0.112	0.014	0.099	0.014
53	0.386	0.025	0.098	0.046	-0.079
54	0.384	0.033	0.096	0.051	-0.073
59	0.363	0.048	0.030	-0.109	-0.124
61	0.364	0.065	0.139	-0.013	-0.027
63	0.089	0.371	-0.014	-0.188	0.395

Variables may be identified from Table 21.

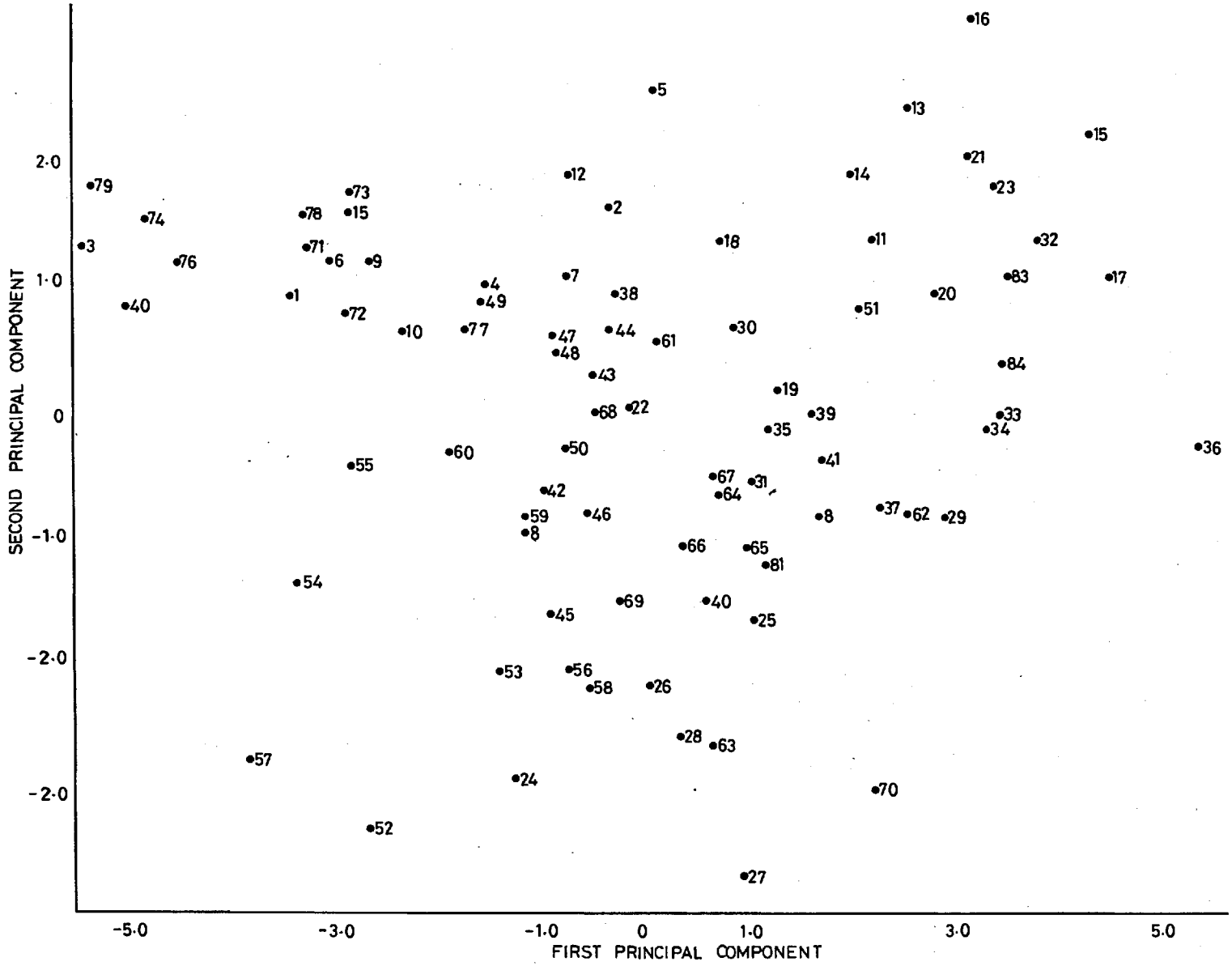


FIG. 67
CLASSIFICATION OF 84 OTU'S USING SUBJECTIVE SELECTION 1

The overlap is more apparent in the plots of subsequent axes, making them of little taxonomic use for assessing the homogeneity of taxa. A series of specimens of 'punctatus' from Japan (52 - 58) form a reasonably coherent group to the left of the main pulicaris-punctatus group.

9.5.2. Selection 2

Seventeen variables were also used in this trial, but were selected with more emphasis on traditional taxonomic characters of the wing pattern, and ratios. They are:

Variable	
1	Contiguity of eyes
2	Proportional length of antennal segment iii
9	" x
15	Number of sensilla on antennal segment iii
17	" xii
18	" xiv
40	Wing pattern element 6
43	" 9
44	" 10
45	" 11
60	Length of hind metatarsus
64	Cibarial teeth
65	Antennal ratio
66	Palp ratio
69	Ratio of antennal length to proboscis length
70	Ratio of cibarium length to pharynx length
72	Costal ratio

These characters embody many of the traditional characters (in a quantified form) and therefore might be expected to produce a useful classification of the pulicaris group. Surprisingly, this was not the case, for as Fig. 68 shows, many of the specimens are not grouped by taxa. Furthermore, the first three dimensions absorb only 16%, 15%, and 14% of the total variance, in an analysis based on 84 OTU's. This result does not differ much from an analysis based on 53 OTU's (using the same characters), in which the three largest eigenvectors absorbed 18%, 17%, and 14% respectively. The relatively low proportion of the total variance described in the first few dimensions suggests that the relationships (taxonomic or otherwise)

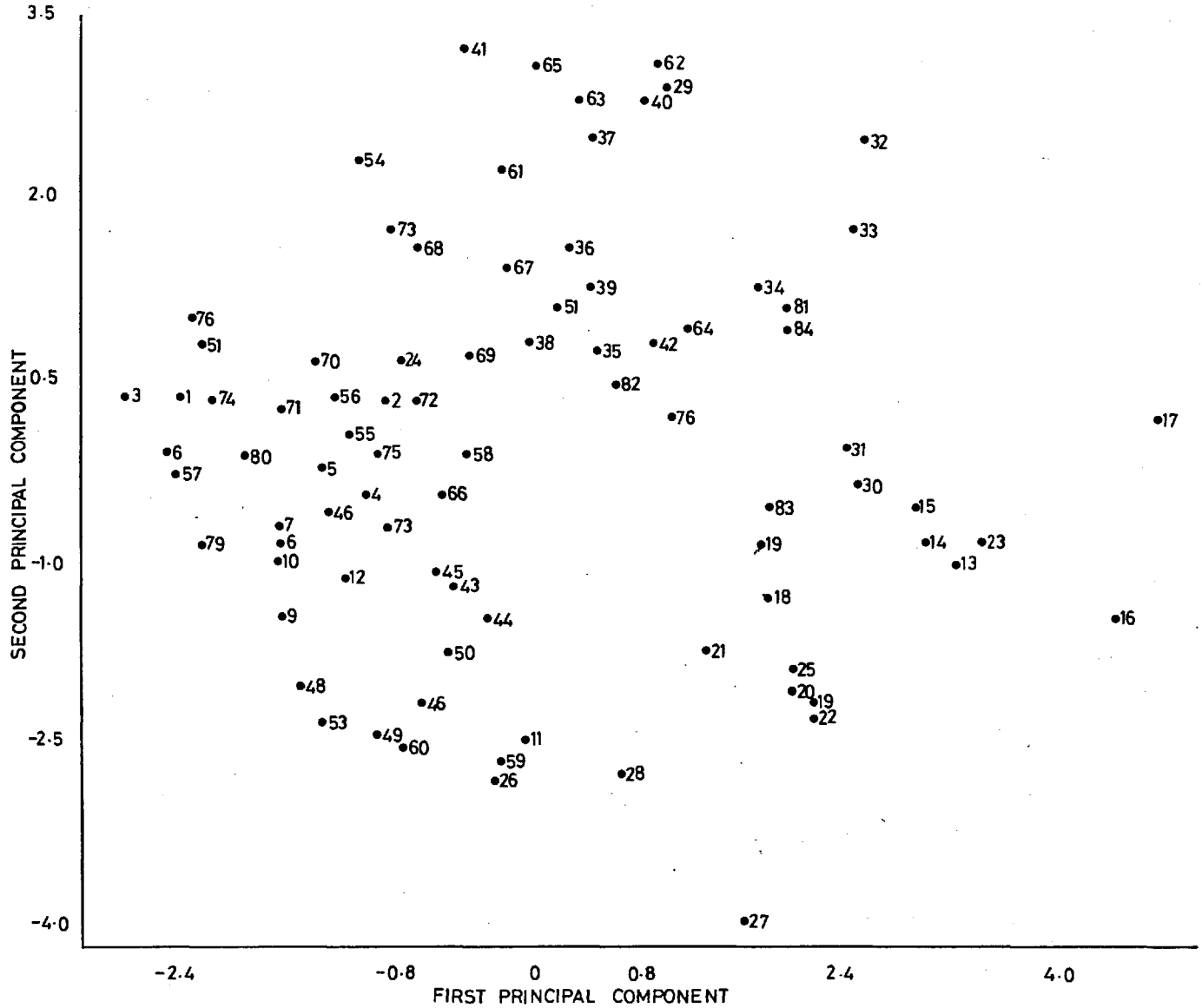


FIG. 68
PRINCIPAL COMPONENTS PLOT OF 84 OTU'S USING VARIABLES IN SUBJECTIVE SELECTION 2

between the specimens cannot be effectively summarised by this analysis. In traditional taxonomic studies of this complex, fewer characters (than 17) have usually been used. These results show that even when many of the traditional characters are considered simultaneously (as they are in principal component analysis) a concise description of the taxonomic relationships is not forthcoming.

Fig. 68 shows that specimens fall into three main groups: a large heterogeneous group to the left, and equally heterogeneous group in the centre and a small and relatively homogeneous group to the right. The grouping of the specimens is briefly as follows:

- newsteadi (1 - 12): all the specimens are in the large, left hand cluster, and are generally placed close together. One specimen, OTU 11, with a high antennal ratio, is placed well apart from other specimens of this species.
- grisescens (13 - 23): this is the only species in which the specimens are grouped together and apart from those of other species. In Fig. 68 the small cluster of points to the right is mainly composed of specimens of grisescens, but also in this group are two specimens of punctatus (30, 31), one lupicaris (83) and some specimens of fagineus. However, the fagineus are quite distinct in the third dimension, with component values of +4.0 - +6.6, compared with values of -0.8 - -0.9 for grisescens.
- impunctatus (71 - 80): although specimens of this species are not generally spread over the diagram, they do not form a distinct group apart from the other taxa. They are placed in a large group to the left with newsteadi and punctatus which remains heterogeneous in three dimensions.
- lupicaris (81 - 84): these are generally placed close together, in the centre of the diagram.

Specimens representing three species - pulicaris, punctatus and delta - are distributed throughout the three main clusters, especially punctatus.

The characters which are most important on the first eigenvector are the antennal ratio, length of hind metatarsus and the ratio of antennal length to proboscis length (Table 37). Considerable emphasis is put on the first two of these characters (wing length is normally used instead of the hind metatarsus) in diagnostic keys to the pulicaris complex, and therefore the loadings given to

TABLE 37.

Eigenvalues and Eigenvectors from a Principal Component Analysis Based on the Variables in Subjective Selection 2

	1	2	3	4	5
Eigenvalue	2.848	2.543	2.283	1.676	1.580
Cumulative Percentage of Trace	16.75	31.71	45.14	55.00	64.29
Variable		Eigenvectors			
1	-0.153	0.072	0.162	0.519	0.012
2	-0.284	0.181	-0.289	-0.068	0.361
9	-0.193	0.115	0.355	-0.289	0.244
15	0.102	-0.268	0.428	-0.051	-0.091
17	0.254	0.078	0.278	-0.298	-0.185
18	0.153	0.334	0.133	0.025	0.141
40	-0.129	-0.488	-0.003	-0.143	0.180
43	-0.247	0.030	-0.126	-0.349	-0.365
44	0.110	-0.359	-0.080	-0.046	0.295
45	-0.118	-0.431	-0.189	-0.300	0.017
60	0.456	0.137	-0.005	-0.157	-0.028
64	0.051	-0.209	0.495	0.165	0.169
65	0.218	-0.278	-0.128	0.148	-0.541
66	0.467	-0.135	-0.041	0.037	0.254
69	-0.373	-0.179	0.198	0.306	-0.166
70	0.014	-0.039	-0.083	-0.038	0.007
72	0.197	-0.103	-0.343	0.372	0.098

Variables may be identified from Table 21.

them in this study confirm their importance. However, it must also be noted that the first vector accounts for only a small proportion of the total variance (16%).

The second component may be identified as a wing pattern vector because three of the four wing pattern characters receive large loadings on the corresponding eigenvector. The separation of C. fagineus on the third principal component noted above, is attributable to the high loadings given to three characters important in defining this species: cibarial teeth, large number of sensilla and relative length of antennal segment iii.

9.5.3. Selection 3

As in the previous experiment, this selection is also based on traditional characters, although some groups, e.g., wing pattern, were represented by individual characters. The experiment is based on 15 characters, they are:

Variable	
1	Contiguity of eyes
2	Proportional length of antennal segment iii
9	" x
15	Number of sensilla on antennal segment iii
17	" xii
18	" xiii
37	Wing pattern element 3
38	" 4
43	" 9
46	" 12
60	Length of hind tarsus i
64	Cibarial teeth
65	Antennal ratio
66	Palp ratio
72	Costal ratio

Details of the eigenvalues and eigenvectors are given in Table 38 . The first axis absorbs 22% of the total variance, rising to a total of 50% by the third. This is higher than in the previous experiment, but not as high as the first, although it should be remembered that only 15 variables were used in this trial, compared

TABLE 38.

Eigenvalues and Eigenvectors from a Principal Component Analysis Based on the Variables in Subjective Selection 3

	1	2	3	4	5
Eigenvalue	3.323	2.624	1.623	1.540	1.490
Cumulative Percentage of Trace	22.15	39.65	50.47	60.73	70.66
Variable			Eigenvectors		
1	0.205	-0.128	0.332	0.277	-0.402
2	-0.219	-0.379	0.115	-0.246	-0.164
9	0.219	-0.256	-0.353	-0.284	-0.096
15	0.229	0.214	-0.437	0.181	-0.090
17	0.259	0.213	-0.104	-0.250	0.320
18	0.193	0.009	0.285	-0.417	-0.030
37	-0.421	0.072	-0.111	-0.072	-0.023
38	-0.434	-0.060	-0.302	-0.019	-0.059
43	-0.028	-0.265	-0.015	0.149	0.564
46	-0.419	0.150	-0.268	-0.085	-0.062
60	0.076	0.378	0.152	-0.412	0.180
64	0.300	0.129	-0.342	0.165	-0.353
65	-0.064	0.369	0.120	0.465	0.301
66	-0.072	0.474	-0.023	-0.237	-0.212
72	-0.235	0.240	0.367	0.084	-0.262

Variables may be identified from Table 21.

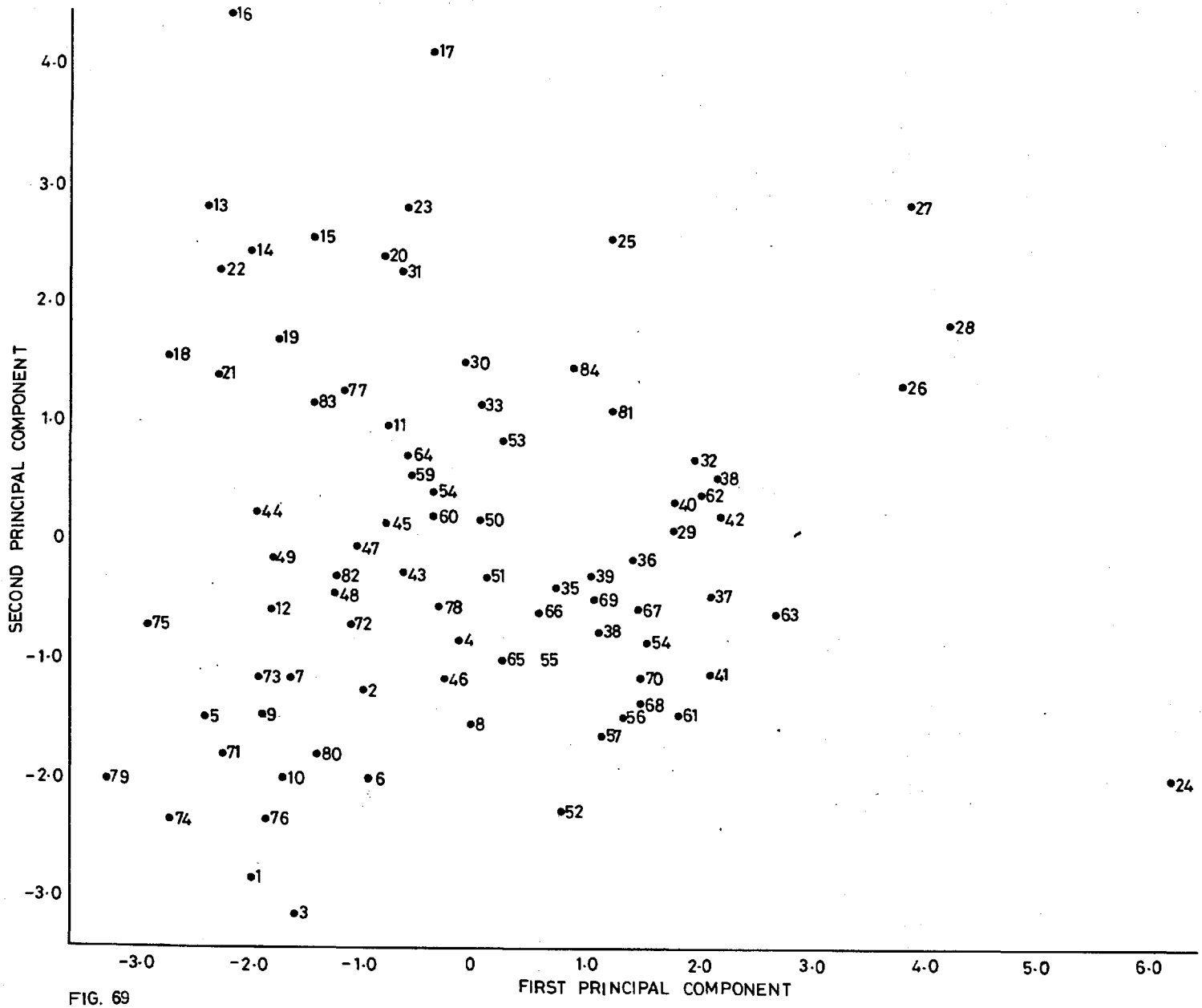


FIG. 69
PLOT OF FIRST TWO PRINCIPAL COMPONENTS, FROM ANALYSIS USING VARIABLES IN SUBJECTIVE SELECTION 3

to 17 in the previous two. The plot of the first two principal components (Fig. 69) shows that generally, the specimens are grouped according to their taxa. As found previously, the species do not form well separated groups, but overlap to a considerable extent. Four species - grisescens (13 - 23), fagineus (24 - 27), impunctatus (71 - 80) and newsteadi (1 - 12) form 'satellite' clusters around a large, central group of pulicaris, punctatus, delta and lupicaris. The plot of the second and third components (Fig. 70) shows that the shape of the central conglomerate does not alter significantly when a further dimension is considered. Two small species, impunctatus and newsteadi, are superimposed when only two dimensions are considered, but are well separated in the third. The position of each taxon may be summarised as follows:

- newsteadi (1 - 12): specimens are all well grouped, except OTU 11, which is nearer grisescens cluster in both plots (Figs 69-70).
- impunctatus (71 - 80): again, a fairly well defined species, with one outlying OTU - 77 - well within the delta cluster.
- grisescens (13 - 23): a distinct species, especially in the plot of the second and third axes.
- delta (29 - 37): specimens identified as this species are spread throughout a central heterogeneous cluster, together with pulicaris and punctatus. One specimen, OTU 31, is placed well within the grisescens group.
- fagineus (24 - 28): no apparent grouping of specimens, as they are generally distributed in the central multispecies cluster.
- punctatus (38 - 51); pulicaris (59 - 70); lupicaris (81 - 84): these three species are the basis of a conglomerate, in the centre of the diagrams in both plots (Figs 69-70).

Three of the specimens are possible misidentifications and will be discussed in the section concerning the homogeneity of the taxa.

The large loadings associated with variables 37, 38 and 46 on the first axis allows it to be identified as a wing pattern vector. These loadings are considerably larger than any others, except that associated with cibarial teeth. This last variable is important in identifying specimens of fagineus, which are placed to the right on the first axis. On the second vector, three ratios - palp ratio, proportional length of antennal segment iii and the antennal ratio, together with the length of the hind metatarsus, have large loadings and subsequently dominate this axis. The length of the latter

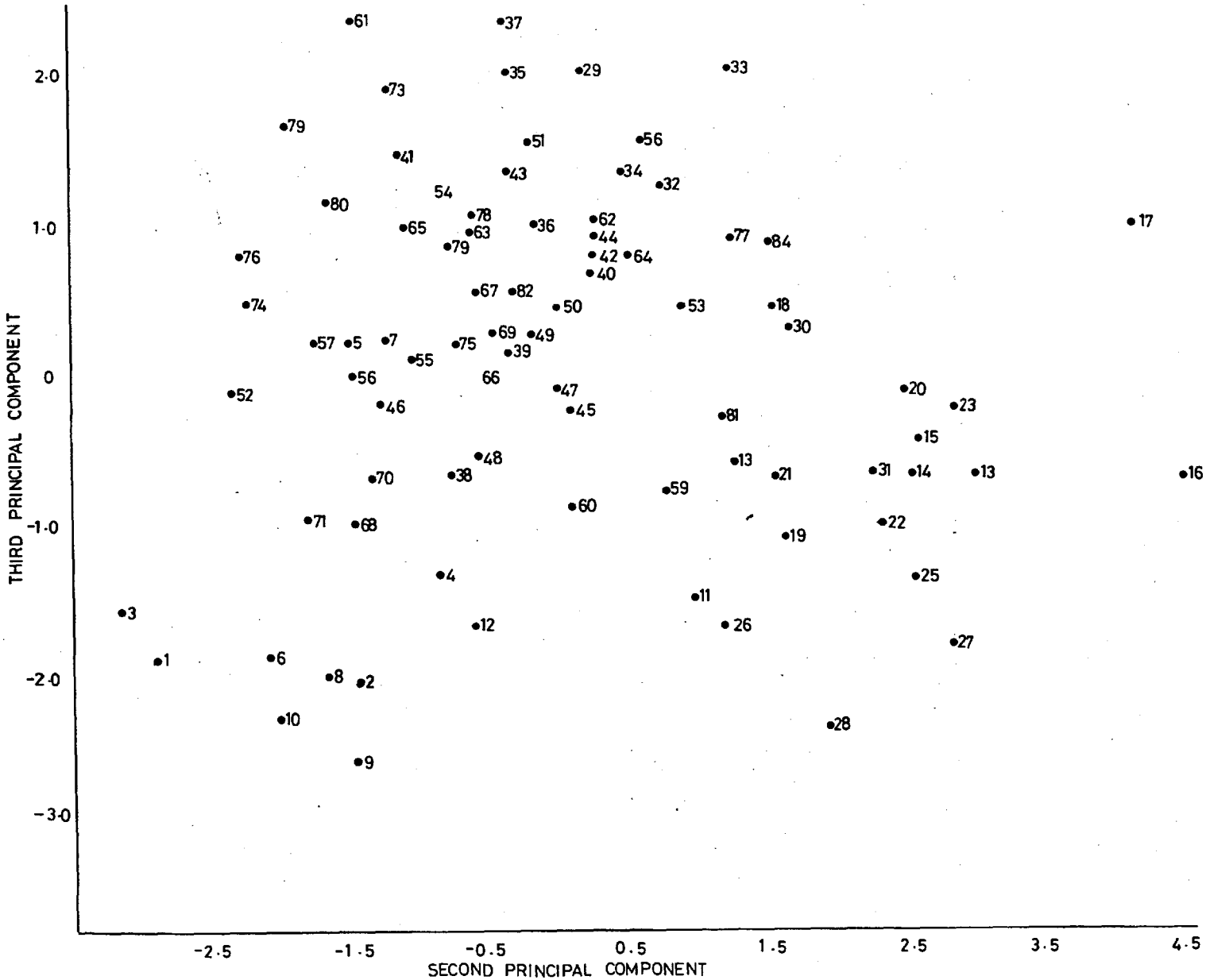


FIG. 70
PRINCIPAL COMPONENTS PLOT, FROM ANALYSIS USING VARIABLES IN SUBJECTIVE SELECTION 3

is the only overt measure of size in this selection of variables. The association of wing pattern with vector one, and ratios and size with the second, is the reverse of that found in previous experiments, including that based on the complete set of variables. The third vector is not easily interpreted because there is little difference between the three or four largest loadings and the remainder, as found in the first two axes.

9.6. SUMMARY

9.6.1. Comparison of Techniques for Reducing the Number of Variables

This section comprises the effectiveness of those methods used here to eliminate redundant variables for principal component analysis. Three main approaches were used in selection of a representative subset of variables:

(i) the use of variables from only one body region
(ii) objective numerical methods
(iii) subjective selections

The success of any given method is determined by comparing visually, the arrangment of a sample of OTU's based on the subset, with the arrangment of the same OTU's, based on an analysis of the complete set of variables. In this study, the congruence between classification was evaluated by eye and therefore is rather subjective. Numerical methods for comparing classifications have been discussed above (p.186). By far the mostsuccessful method, according to the criteria above, used the loadings from a principal component analysis. The other numerical method employed cluster analysis (using a correlation matrix) to produce several groups, and then a single representative was selected from each. The classification based on the 19 variables selected by this method revealed little similarity to the component analysis using all 72 variables. Not only did this technique fail to reproduce the arrangement of the OTU's, but bunched them so tightly as to render the results of little taxonomic use.

A main difference between the two numerical methods is that using the loadings on a principal component analysis makes it possible to observe directly the effect of each character on the classification and therefore select those characters which exert the

most influence. One of the theoretical reasons for the use of cluster analysis,is that redundant variables are being removed by selecting only one variable from each group of correlated variables. In a biological context, this could be interpreted as an attempt to sample a wider range of the genome and to minimise bias due to influence of any one set of characters/genes. However, one of the disadvantages of this method is the inability to distinguish between variables which vary within a group, from those which vary between groups. For example, brown wings may be highly correlated with long legs and long wings, when all the species are pooled. This may mean that one species is large with brown wings and all the other species are small with green wings. Furthermore, long legs and long wings are both measures of overall size and therefore a cluster would be formed of three characters, two of which represent interspecific differences (brown wings and size), and two are correlated within a species (long legs and long wings). From this trio of characters, only one would be selected, most likely long legs or long wings and an important interspecific difference would be rejected. This technique is quite different from removing all characters which are highly correlated within a species. By the method used here, it would be possible to obtain a set of variables which vary little, or not at all, between species and are unlikely to give any reasonable clustering of OTU's.

Of the three selections of variables made on subjective grounds, the first most resembled the classification produced from the total number of variables, as it contained many size characters. The remaining two showed little similarity with the reference classification, undoubtedly because the variables were chosen for their taxonomic importance, rather than for their ability to duplicate the classification.

Perhaps the worst method for selecting a subset of variables used characters from only one part of the body, in this case, the wing. This is probably because general size, and other correlated variables, play an important rôle in the classification, when 72 variables are used. It was found however, that wing pattern was the second most important group of characters influencing the first principal component, and dominated the second in the classification based on many variables. Similar results were obtained when only size-variables were used (Section 9.3). Treated

separately, these suites of characters have little taxonomic use, but together they have a "synergistic" effect. From these results it may be concluded that the prediction made by the non-specificity hypothesis - that similar classifications will be produced from characters taken from different parts of an organism - does not apply to these data.

One point of a more general nature, suggested by these results, is that the recommendation by Sneath & Sokal (1973), to use as many variables as possible in numerical taxonomy, may not always be appropriate. The classification produced from 20 characters gave very similar results to that using 72 characters, implying that 52 of these are in some way redundant. It is possible that because these 20 variables contain representatives of all the character types used in this study (e.g., wing pattern, meristic characters, etc.), they incorporate most of the taxonomically useful information. The addition of characters does not seem to extend the range of character types, but merely duplicates those already present in the subset. In conclusion, the most effective method for reducing the number of variables, to preserve an arrangement of taxa, uses the loadings from a principal component analysis.

9.6.2. Summary of Taxonomic Results of Multivariate Analysis

The main objective of Section 9 is to produce a classification of the taxa in the C. pulicaris species complex. Several have been produced in association with experiments to eliminate redundant variables. This summary compares the taxonomic merits of these ordinations.

In the introduction to this section, it was stated that the general methodology was to treat all specimens as OTU's, group them by numerical techniques, and then inspect these groups to determine which taxa are homogeneous. Prior to multivariate studies, the specimens were provisionally identified to ensure that, where possible, the variations of all the taxa were represented. This also facilitated taxonomic evaluation of the classifications produced from different subsets of variables. These provisional identifications were based on experience gained through working on the complex over a considerable period of time. Often, criteria were used which were not always objective, a practice frequently encountered in the identification of taxonomically difficult species. The present study was undertaken in an attempt to make this rather intuitive and vague

approach more objective, or explicit.

The multivariate analyses have shown that species do not form discrete and well separated groups, but a somewhat large heterogeneous cluster, with peripheral subgroups. This results means that establishing homogeneous taxa is not simply a case of recording the specimens in each cluster (taxon) and noting any overlap between them. Therefore the provisional identifications were used as hypotheses which could be tested by attempting to refute them. For example, if n specimens were provisionally assigned to taxon Z, the hypothesis is that Z is a homogeneous group, defined by the n specimens. If, during the production of several alternative classifications, two specimens are placed apart from the other n-2 specimens, they may be considered as misidentifications and then allocated to another group. A new hypothesis is then set up for Z and it is redefined. In principal component analysis, the allocation of specimens to taxa is notional, in as much as it does not effect the ordination. In discrimination analysis, specimens have to be allocated to taxa, which then serve to define within-group and between-group covariance matrices. Therefore the a priori assumptions made concerning the homogeneity of taxa will influence the results through the group means and dispersion matrices. It is, of course, possible to use multiple discriminant analysis iteratively, by redefining groups after each cycle of computation, in the hope that re-allocated individuals will improve the clarity of discrimination.

This blend of subjective and objective criteria was employed here to assess the homogeneity of taxa. A comprehensive study of the interplay between the subjective and objective methods used by practising taxonomists, would provide an interesting line of future research and help greatly in defining those areas of traditional taxonomy which might benefit most from numerical techniques.

Several methods are available for the objective comparison of classifications and some, such a matrix correlation and rotational methods, have been discussed above (p.176). However, these techniques only allow estimates to be made of the similarity between classifications and do not decide which are the "best" or most taxonomically acceptable. There are very many possible classifications of a set of objects and such terms as "best" or "most acceptable"

are rather enigmatic. The significance of a classification depends on the aims of a study, how the classification is to be used, or the type of data available. As most are made for practical use, the selection of one classification rather than another is a value-judgement, made by the user. Only in some cases will the aims be presented in a form which allows the results to be tested by objective means. Often this is not possible and (like beauty) the quality of a classification is in the eye of the beholder!

The objective of the present study is to group specimens into morphological taxa and a 'good' classification should be able to summarise the relationships of the taxa concisely. In terms of the analytical techniques used here, such a classification would ideally use relatively few characters to group specimens of the same taxon together, and have the minimum of overlap between the taxa. Clearly, the results obtained so far suggest that the specimens do not form well defined groups. With these data, it is not possible to use a technique which simultaneously groups specimens into taxa objectively, and discriminates between them. This must be undertaken in two stages - classification and then discrimination.

The general inability of the classification produced so far, to group specimens into convenient taxa, suggests that the relationships between the taxa cannot be summarised easily in morphological terms. Associated with this conclusion is the relatively low proportion of the total variance which the first few dimensions of a principal component analysis can describe. This in turn is due to the relatively low correlations between characters. Table 39 summarises the percentage and cumulative percentage of the trace absorbed by the first five eigenvectors (largest), of the classifications produced in this section. All analyses used 84 OTU's, but those based on size or wing pattern alone (p.186and 196 respectively) are not included for reasons given in the discussions of these classifications.

The analysis based on 20 variables selected by the vector loadings method, described the highest percentage of the variance within the data in five dimensions. This is considerably larger than the 64% for the analysis using all 72 variables, and confirms the conclusions made earlier on the efficiency of this reduction

TABLE 39.

Comparison of Techniques for Reducing the Number of Variables. Percentages and Cumulative Percentages of the Total Variance, Associated with the Five Largest Eigenvalues in Each Analysis.

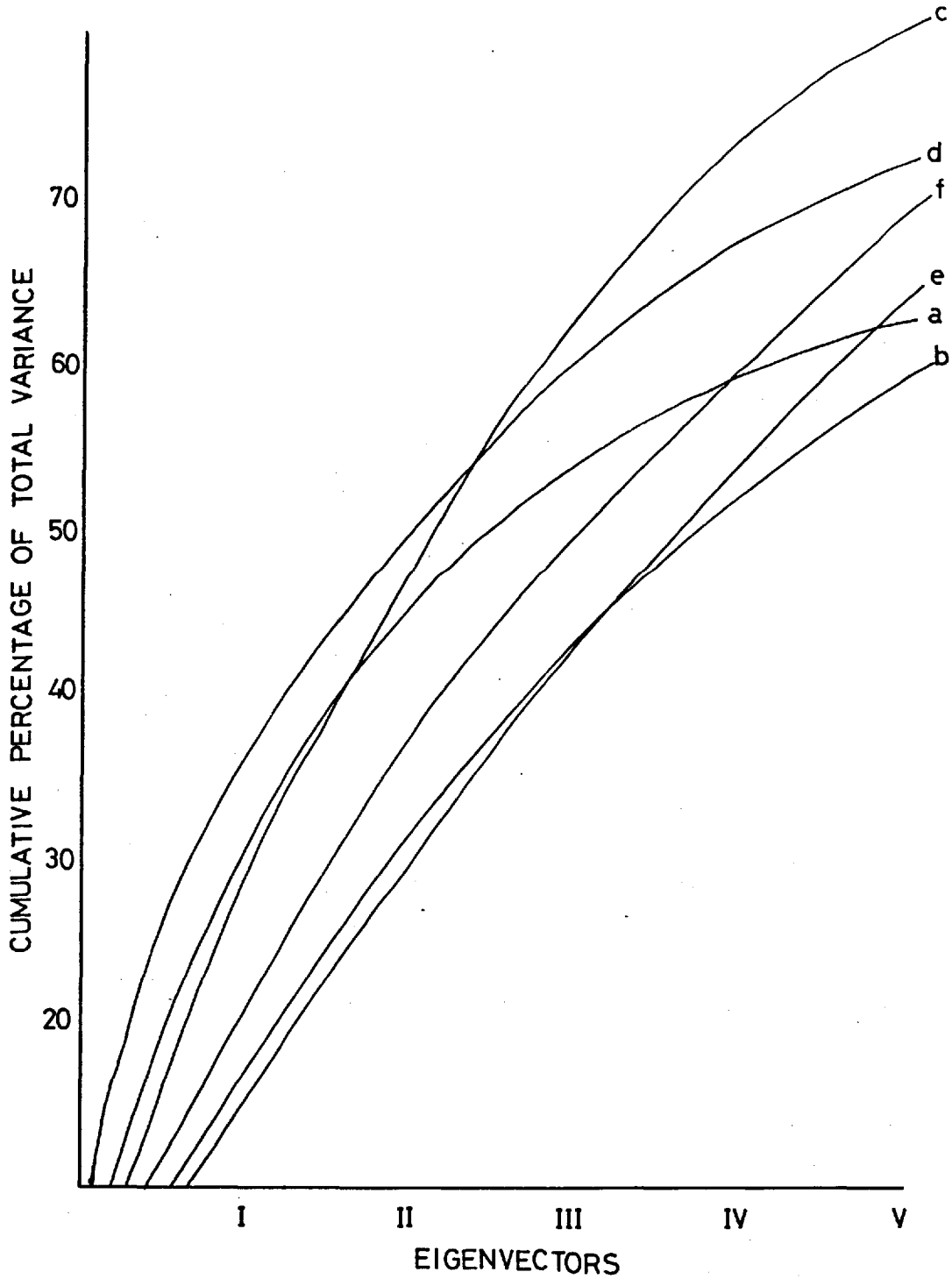
	Number of Variables	Vector I	II	III	IV	V
Complete set of Variables	72	33.4	46.1	54.1	59.1	63.3
		33.4	12.7	8.0	4.9	4.3
Variables selected by Cluster Analysis	19	19.0	32.6	44.7	52.9	60.2
		19.0	13.6	12.1	8.2	7.3
Variables selected by Vector Loadings	20	28.8	48.7	64.2	74.8	80.1
		28.8	19.9	15.5	10.6	5.3
Subjective Selection 1	17	36.3	49.5	60.7	67.2	73.1
		36.3	13.2	11.2	6.5	5.9
Subjective Selection 2	17	16.7	31.7	45.1	55.0	64.3
		16.7	15.0	13.4	10.1	9.3
Subjective Selection 3	15	22.1	39.6	50.5	60.7	70.6
		22.1	17.5	10.9	10.2	9.9

method. The poorest summary of the data in five dimensions was based on the 19 variables suggested by cluster analysis of a between-character correlation matrix. The first of three analyses (complete set of variables, variables selected by vector loadings and subjective selection 1) was dominated by size. In these analyses, the first axis described a much larger percentage of the total variance than in the other three classifications.

To facilitate easier comparison between relative importance of the vectors in each analysis, the cumulative percentage of the trace was plotted (Fig. 71). In general, the relationship between the vectors and the variance they describe, is in the form of a shallow curve. The slopes of these curves are similar for all six analyses. Two classifications, a and d, in which size variables predominated, the curves were attenuated, indicating that the fourth and fifth axes contribute relatively less than they do in other analyses. Although the OTU's are arranged in a character hyperspace, whose dimensions differ according to the subset used, the general resemblence in the shape and slope of the curves implies that for each classification, the overall geometric arrangement of the specimens in hyperspace is very similar.

This method of comparing classifications reveals which summarise the data most efficiently, and which are most reliable when only a few vectors are inspected. However, it does not help decide which is most acceptable in taxonomic terms. The taxonomic merits of each classification will therefore be judged subjectively, by visual inspection of the principal component plots. The combination of variables which produce a classification misplacing the least number of specimens from the provisionally identified taxa, and showing the limits of the taxa most clearly, will be considered the most useful. This approach is analogous to the concept of parsimony used in numerical phyletic studies.

By these criteria, the most acceptable classification is produced by the 15 variables in subjective selection 3. It shows the taxa most clearly in three dimensions. When only two dimensions are considered, the analysis using 72 variables gives satisfactory results. It is possible that artificially coherent clusters are obtained when a small subset is used, because characters emphasising the within-group variation are omitted. This factor could distort the visual and subjective assessment of the taxonomic usefulness of a particular selection.



a = complete set of variables
b = variables selected by between character correlation
c = variables selected by vector loadings method
d = variables from subjective selection 1
e = variables from subjective selection 2
f = variables from subjective selection 3

FIG. 71

PROFILES OF VARIANCE ABSORBED BY FIRST FIVE AXES IN ANALYSES BASED ON DIFFERENT SUBSETS OF VARIABLES

From all the experiments carried out, two important conclusions were reached. Firstly, that as the number of specimens in the study was increased, the boundaries between the taxa became less distinct and the degree of overlap between them increased. The second conclusion was that even when such powerful techniques as principal component analysis were used, a concise summary of the data was not possible.

Subdivisions of the C. pulicaris Species Group

The subdivision of the pulicaris group has been proposed on two previous occasions, by Wirth & Blanton (1969) for the North American species and by Kremer (1965) for the western Palaearctic species. An outline of these proposals has been given above (p. 59). This section discusses how the results of the multivariate analysis concur with these suggestions. The divisions of Wirth & Blanton were designed for North American species and proved most unsatisfactory when applied to the Palaearctic taxa, which are the subject of the present study. Comment will therefore be confined to the proposals of Kremer (1965).

In general, the multivariate studies lend little support to the idea of subdividing the group. It was found that the species intergraded to such an extent that any rational division would seem inappropriate. In addition to this conclusion, a few more detailed comments might be pertinent. Kremer placed grisescens and fagineus in the Grisescens sub-group. Although the present analysis confirms the distinction of the two species, it does not place them together. On the contrary, fagineus and grisescens are usually placed on opposite sides of an ordination diagram, confirming the previous comments that this subgroup is composed of species which share only a dissimilarity to other species in the complex rather than similarity to each other. This division, therefore, has little taxonomic value. Kremer proposed two other groups, the first containing pulicaris, punctatus, newsteadi (as halophilus) and lupicaris, and the second containing delta and impunctatus. While these subdivisions draw attention to general trends in the group, they are of little practical use, because the boundaries between them are so vague. For example, there is as much evidence that delta is closely related to impunctatus as there is of its affinity with pulicaris and punctatus. A more appropriate

division suggested by the multivariate analysis, if it is felt that the group requires dividing, would be a large core group containing pulicaris, punctatus, newsteadi and delta, with three peripheral groups each containing one species, viz., grisescens, fagineus and impunctatus. This grouping of the species, in common with the other divisions, has little to offer and is therefore not recommended.

Possible Misidentifications Shown by Multivariate Analysis

One of the benefits of using individual specimens as OTU's in a principal component analysis is that misidentifications can be recognised. During the course of the analysis, doubt was thrown on the provisional identification of several specimens, especially the following:

OTU 11

This was provisionally identified as C. newsteadi, but was consistantly placed apart from the other specimens of this taxon in analyses based on either size characters or ratios. The main features which placed the specimen in the C. grisescens cluster was its high antennal ratio and its size.

OTU 24

This specimen from Israel was an outlier in most analyses, but was located with the small specimens in experiment two. It was positively identified as C. fagineus by the presence of the cibarial teeth. Unfortunately, this is the only representative of a series recorded from Israel (Braverman et al., 1976) which was seen, the others apparently having been lost. This lack of supporting material made it impossible to determine whether this small, pale specimen was evidence of geographical variation in fagineus, or a representative of a new taxon.

OTU's 30 and 31

Campbell & Pelham-Clinton identified these two specimens as C. delta during the preparation of their 1960 paper. The specimens were collected at different localities in Scotland. In most of the ordinations, they were placed on the margin of the grisescens

cluster. Those classifications which concentrated on variables other than size (experiments 2 and 3) placed them well within the C. orisescens cluster, strongly suggesting that they should be identified as this species.

OTU 51

This specimen of punctatus from Lundy Island has been commented on above, in the study of seasonal variation (p. 102) and in this section on p. 251. It is a large specimen and therefore in the analysis based on 72 variables, it was separated from the other punctatus and within the cluster of C. delta. In experiments 2 and 3, in which size is of little importance, this specimen is shown to be a 'typical' member of the punctatus group. Features of the wing pattern confirm this.

OTU 77

Originally identified as C. impunctatus, this specimen is often put into the delta group, especially in those analyses which use few size variable, e.g., experiment 3, particularly in the plot of the second and third components. Although this specimen is much smaller than most C. delta, its high antennal ratio is typical of this species and distinguishes it from impunctatus.

Recognised Taxa in the C. pulicaris Complex

Using the combined results of the multivariate analyses, the following taxa are recognised:

C. pulicaris Specimens of pulicaris were usually grouped together with specimens of C. punctatus in the centre of most classifications. Edwards (1939) fixed the name pulicaris to this taxon. The multivariate analysis shows this to be a most sound judgement, in view of its central position. Specimens 59 - 70 are included in this taxon.

C. punctatus Specimens of this species fall into two groups. The first is composed of OTU's 38 - 42, collected in a range of British localities, and the second is composed of OTU's 43 - 50, a series of small specimens from Norway. The second group was placed with specimens of newsteadi in the analysisbased on 72 variables.

However, they were grouped with the remaining punctatus in classifications such as experiment 3. This analysis was more or less independent of size. The taxonomic position of OTU 51 from Lundy Island has already been discussed in the section on misidentifications. This species consists of OTU's 38 - 51.

C. impunctatus Specimens provisionally allocated to this species were placed with those of newsteadi in the classifications dominated by size. They were distinct however from newsteadi in experiment 3. OTU's 71 - 76 and 78 - 80 represent this species.

C. newsteadi Owing to its small size, this species was often grouped with C. impunctatus and the Norwegian specimens of punctatus. As noted under these species, all three taxa were shown to be distinct in experiment 3. One specimen (OTU 11) provisionally identified as newsteadi, was later shown to be C. grisescens. This species comprises OTU's 1 - 10 and 12.

C. delta The similarity of this species to C. impunctatus and C. pulicaris was demonstrated repeatedly by the extent to which the three species overlapped in the multivariate analysis. This species is most distinct in the classification based on the complete set of variables. Two specimens (OTU's 30 and 31) originally thought to be delta, were shown to be C. grisescens. OTU 77 was transferred to this species from the closely related C. impunctatus. Specimens allocated to this species are 29, 32 - 37 and 77.

C. fagineus In the analysis based on all 72 variables, the five specimens of this species are placed ambiguously in the centre of the diagram. In classifications such as experiment 3, they form a distinct and well defined group. The species is easily recognised by the presence of cibarial teeth. This character was discovered when accurate measurements were being made of the cibarium. Because this species is readily identified, it was not included in the discriminant analysis.

C. grisescens This species formed a distinct cluster in all of the analyses, proving it to be the most homogeneous of all the taxa. In addition to OTU's 13 - 23, this species is represented by OTU's 11, 30 and 31.

C. lupicaris Unfortunately, only four specimens of this rare species were available for study. In all the analyses, the specimens were scattered throughout other taxa, especially delta, grisescens and pulicaris. Many authors have rejected this species and either

synonymised it with pulicaris or delta (see p. 73 for details). The results of the multivariate analysis, although based on a few specimens, substantiate these doubts. This taxon was not recognised for the purpose of discriminant analysis, on the grounds that the species was not homogeneous and the sample available was very small.

Sp. A A series of specimens from Japan could not be identified with any confidence and were allocated to C. punctatus originally. In the multivariate analysis, they form a distinct group, including that based on all 72 variables. The seven specimens of this taxon were recognised as a separate taxon for the purposes of the discriminant analysis.

Section 10. DISCRIMINATION AND IDENTIFICATION

10.1. INTRODUCTION

The principal objective of the previous section was to characterise the species in the C. pulicaris complex. Having achieved this, a system may now be developed for discriminating between them and for the identification of additional specimens.

It is most important to distinguish between the processess of classification and discrimination. The former is concerned with forming classes, recognising clusters of specimens and constructing taxa. Discrimination, on the other hand, is aimed at selecting variables, and if necessary, weighting those which emphasise the differences between groups. It is a precursor to identification in which unidentified individuals are allocated (as far as possible) to known groups. It is important to note that statisticians often speak of 'classifying' when, to the taxonomist, they are involved in the process of identification. This difference in terminology has led to some confusion in the past. The main objective of identification is to associate a specimen with similar specimens, with ease and certainty. All other considerations (such as phylogenetic relationships) are secondary. Sneath & Sokal (1973) state that in recent years, considerably more effort has been put into the perfection of classification rather than discrimination techniques. They suggest that the topic of discrimination and identification will expand rapidly in the next few years, bringing substantial advances.

Identification techniques may be divided into two broad groups: sequential methods and simultaneous methods. The former group contains the most familiar and commonly used technique - the dichotomous key. There have been several advances on the traditional diagnostic key, summarised in a symposium on this subject (Pankhurst, 1975). These innovations include multiple entry keys, polythetic polyclaves and on-line computer systems. Simultaneous methods rely on some level of agreement over all characters so that identification is made in one step. The discriminant function is typical of this group and, like other methods, is probabilistic. This means that identification of a specimen is associated with an estimate of the likelihood of its

membership of a group.

The application of any method depends to a large extent on the type of data available. Sneath & Sokal suggest that for large studies on well separated taxa, sequentialmethods are more appropriate. Discriminant analysis is most valuable in problems involving closely related and overlapping groups. Quantitative variables are more conveniently incorporated into discriminant methods.

For the pulicaris complex, discriminant analysis is most appropriate. However, it is possible that the 'identification space' methods devised by Gyllenberg & Niemela (1975) for binary data may be adapted for this problem.

Discriminant function analysis was first developed by R.A. Fisher to separate two groups. The method was later generalised for use with several groups, and termed multiple discriminant function analysis. Frequently, increasing the number of taxa decreases the ease of interpretation and subsequent identification. Therefore the method should be applied to as few taxa as necessary when considerable overlap exists between them. Multiple discriminant and canonical variate analysis are equivalent, except for minor details. Unknown specimens can be readily identified by using the linear function which defines each canonical variate, and so places the specimen on the canonical axes. The specimen is named according to its proximity to a group centroid.

In multiple discriminant analysis, the basic matrix has been termed the identification matrix. It consists of a number of submatrices, one for each taxon. The submatrix contains measurements of several variables made on a sample of specimens. The identification matrix of this stduy is based on the same data as the primary data matrix used in the previous section on classification. The matrices differ however, in that the specimens are grouped into taxa for discriminant analysis, whereas those in the classification study were considered individually.

There are several methods for computing canonical variates (see Davies, 1971), that used here having been outlined in Section 6. Usually, the eigenvalues and eigenvectors (the stage prior to the calculation of discriminant scores) are extracted directly from the product matrix $W^{-1}B$, where W is the within groups sums of squares and products matrix and B is the between groups sums of squares and products matrix. Under these

circumstances, the elements of the eigenvectors, when suitably scaled, indicate the contribution each variable makes to the corresponding canonical variate. In the method used here, to use the terminology of Fig.24 (see Section 6), the eigenvectors are expressed in terms of the w axes, and not in terms of the original variables. Therefore, to assess the contribution of each variable to an axis, the correlation between the variable and the canonical variate is calculated. The correlation of the i-th canonical variate with the j-th variable in x is:

$$\underline{w}'_j \underline{b}_i \,/\sqrt{w_{jj}}$$ where $\underline{w}'_j$ = j-th row of $\underline{W}$; w_{jj} = jj-th diagonal of $\underline{W}$; $\underline{W}$ is the within groups sums of squares and cross products matrix; $\underline{b}_i$ = vector of weights for canonical variate i.
This correlation can also be used to determine whether the linear combination of all ten variables may be replaced by only one. This is only feasible when there is a high correlation with only one, or perhaps two, variables. In this study, none of the correlations were sufficiently large to make this simplification practical. The canonical variates were computed by using a series of programs developed by Dr. M. Hills, Biometrics Section, B.M.(N.H.), based on the programs published periodically in the journal Applied Statistics.

Canonical variate analysis assumes equality of the within groups covariance matrices, which should therefore be tested for homogeneity before attempting the analysis. In general, few studies involving multivariate analysis mention the subject of significance tests, but as multiple discriminant techniques rely so much on homogeneity of dispersion matrices, it is important to establish whether the data conform to this assumption. Tests were carried out using the method outlined in Cooley & Lohnes (1962), by a program written by R.G. Davies of Imperial College.

The test is inapplicable however, if either a matrix is singular (i.e., has no inverse) or if the determinant is negative. In this analysis, some of the within group covariance matrices were singular and had negative determinants, and so the test could not be applied. Therefore, the assumption had to be made that the within group dispersion matrices were homogeneous (see general discussion for further comment). The reason for the singularity of some matrices is not clear, but one possible explanation is given by Brown (1969) which seems applicable here.

Briefly, the explanation is that a correlation or covariance matrix based on n replicates of k variables must, when n is less than k, as in some instances occurs in these data, be of rank not greater than n. This means that it has at least (k-n) zero eigenvalues, and is therefore singular. At present, there do not appear to be any tests available to test the homogeneity of dispersion matrices, when either the matrices are singular or have negative determinants.

Since one of the major objectives of this section is the identification of unknown specimens, practical considerations are of prime importance. For a specimen to be identified, each variable will have to be measured in order to calculate the position of the specimen on a diagram. Clearly 72 variables is excessive and the number used should be reasonably small if the method is to be practical. Selecting variables for discriminant analysis is, for computational reasons, considerably more difficult than for other multivariate methods, such as principal component analysis. Sneath & Sokal (1973) suggest that a nearly optimal set of variables will be selected by inspecting the data and choosing those characters with means that are well separated in relation to the variances, and that are not highly correlated with other characters.

Although size variables differ significantly between species, they show seasonal and geographical variation, which renders them taxonomically unreliable (see section 7). If the analysis was dominated by size variables, there is a likelihood of specimens from localities not included in this analysis, falling outside the ranges of the taxa. For example, if the centroid of punctatus was calculated without the sample from Norway, then because of their small size, these specimens might not have been identified as punctatus. For this reason, only one size variable, length of hind metatarsus, was included in the discriminant analysis.

By inspecting the matrices, 14 variables were selected to separate the seven taxa in the C. pulicaris complex (newsteadi, grisescens, delta, pulicaris, punctatus, impunctatus and sp. A). C. fagineus was not included in the discrimination analysis because it is easily recognised by the presence of cibarial teeth. When a canonical variate analysis was carried out with the 14 variables, four of the characters had very low correlations with

any of the first three canonical variates. This implies that these characters contribute relatively little to discrimination between the species and their removal would improve the ease with which specimens could be subsequently identified. The ten remaining characters were:

Character	
1	Contiguity of eyes
2	Proportional length of antennal segment iii
17	Number of sensilla on antennal segment xii
18	Number of sensilla on antennal segment xiii
37	Wing pattern element 3
43	Wing pattern element 9
46	Wing pattern element 12
65	Antennal ratio
66	Palp ratio
72	Costal ratio

These characters represent several different types of character taken from different parts of the midge, and as such are not subject to the limitations of the non-specificity hypothesis.

10.2. DISCRIMINATION OF SPECIES

Canonical Variate Analysis of Specimens Classified in Section 9.

In Section 9, 84 specimens were used to determine which taxa of the complex were homogeneous and to recognise which of the provisional identifications were incorrect. Originally, eight taxa were thought to exist in the complex, but the principal component analysis proved one of them, C. lupicaris, too heterogeneous to be allowed simple specific status. The analysis also showed that a sample from Japan, referred to as sp. A, was sufficiently distinctive to warrant recognition as a separate species. An analysis was therefore carried out on the seven taxa and ten characters listed in Section 10.1.

A summary of the data and dispersion matrices for each taxon is given in the appendix. The zero values in some of the dispersion matrices are due to the invariance of a character within the taxon.

The pooled variance-covariance matrix did not show any zero values (Table 40). A summary of the analysis is given in Table 41.

Figs 72 and 73 show the position of the specimens on plots of the first three axes, which account for 90% of the total discrimination. This is an encouraging result and suggests that the relationships of the groups are a good summary of the original seven dimensions. In view of the undesirable use of absolute size to separate species discussed earlier (p.), it is of interest that none of the axes separate the species by size. This result, therefore, is a substantial improvement on previous taxonomic studies of the group (e.g., Campbell & Pelham- Clinton, 1960). Two variables have relatively high correlations with the first canonical variate: contiguity of eyes and wing pattern element 12, and are therefore important in separating the taxa on this axis. The second axis is dominated by antennal variables: the antennal ratio and proportional length of antennal segment iii. Wing pattern element 9 is the most important variable on the third axis. In the plot of the first and second canonical variates (Fig. 72), two species are distinct - newsteadi and grisescens- and to a lesser extent, sp. A. One of the most notable aspects of this diagram is the degree of overlap between pulicaris and punctatus. A group of seven punctatus specimens are intermediate between newsteadi and the main group of punctatus and pulicaris. These are from Norway and their intermediate nature has been commented on previously (p. 108). Although they fall within the range of values of punctatus, they contribute disproportionately to the difference in the group means of punctatus and pulicaris. In this projection, C. impunctatus appears to be intermediate between newsteadi and punctatus, but when the third canonical variate is inspected (Fig. 73), impunctatus becomes quite distinct. A clearer understanding of the relative positions of the taxa is gained from the three dimensional diagram in Fig. 74 , where only the group means have been used. C. grisescens, impunctatus and newsteadi, in the foreground, appear well separated from the remaining four species.

To resolve the differences between pulicaris and punctatus, an additional 46 specimens were included in the study. The specimens of punctatus and pulicaris were taken from a wide geographical range of localities, and allocated to the appropriate taxon according to wing markings. It has already been noted (p. 65) that

TABLE 40

Group Means and Pooled Correlation Matrix for Canonical Variate Analysis of 7 Taxa.

GROUP MEANS FOR THE TEN CHARACTERS

4.8182	.0744	.9091	1.0000	3.2727	1.3636	4.0000	1.0528	2.2822	.5707
5.5000	.0701	1.2143	1.1429	3.6429	.0714	4.4286	1.1378	3.8242	.6376
8.3750	.0701	1.0000	1.7500	2.2500	.5000	2.1250	1.0896	3.1026	.6241
7.8250	.0732	1.0000	1.1000	2.4125	1.1125	2.0625	1.1082	2.7771	.5974
9.4286	.0716	1.0000	1.0000	2.4286	.8571	1.1429	1.1397	2.1935	.6041
7.9394	.0719	1.1818	1.3333	2.3788	1.1364	1.8182	1.0568	2.7103	.6038
9.6957	.0806	.2609	1.0435	3.6957	.0435	2.4348	1.0211	2.8146	.6233

OVERALL MEAN

7.8015	.0737	.9338	1.1765	2.8088	.8015	2.4228	1.0801	2.8241	.6072

DEGREES OF FREEDOM OF POOLED COVARIANCE MATRIX = 129

POOLED STANDARD DEVIATIONS

1.2176	.0054	.3588	.3848	.6957	.3595	1.1806	.0588	.3858	.0309

POOLED MATRIX

1.00									
.24	1.00								
.06	.07	1.00							
-.06	.14	.13	1.00						
-.12	-.22	-.16	-.20	1.00					
-.14	-.20	-.06	-.01	.21	1.00				
-.01	-.19	-.10	-.18	.39	.35	1.00			
-.27	-.59	-.05	-.15	.19	.26	.29	1.00		
-.14	-.07	.02	-.03	.11	-.07	-.01	.12	1.00	
.01	-.03	-.17	-.02	.08	.10	.31	.09	.12	1.00

****

TABLE 41.

Summary of Canonical Variate Analysis of Seven taxa of the C.pulicaris Complex

Cumulative Percentage of Trace	I 39.71		II 68.80		III 90.18	
Variables (code numbers)	Coeffic.	r	Coeffic.	r	Coeffic.	r
1	0.55	0.50	-0.25	-0.18	0.41	0.55
2	-8.88	-0.00	182.3	0.37	93.8	0.37
17	-0.67	-0.03	-0.41	-0.20	-1.35	-0.39
18	-0.37	0.02	-0.88	-0.27	-0.07	-0.03
37	-0.28	-0.32	0.45	0.16	0.26	0.27
43	1.31	0.26	0.29	0.30	-1.55	-0.61
46	-0.81	-0.50	0.18	0.15	0.08	0.13
65	0.66	0.01	0.52	-0.14	1.33	-0.13
65	-0.42	-0.37	-1.79	-0.51	-0.07	0.09
72	5.36	-0.05	-10.8	-0.26	-0.11	0.24

r = correlation of canonical variate with original variable

Variables may be identified from p. 284

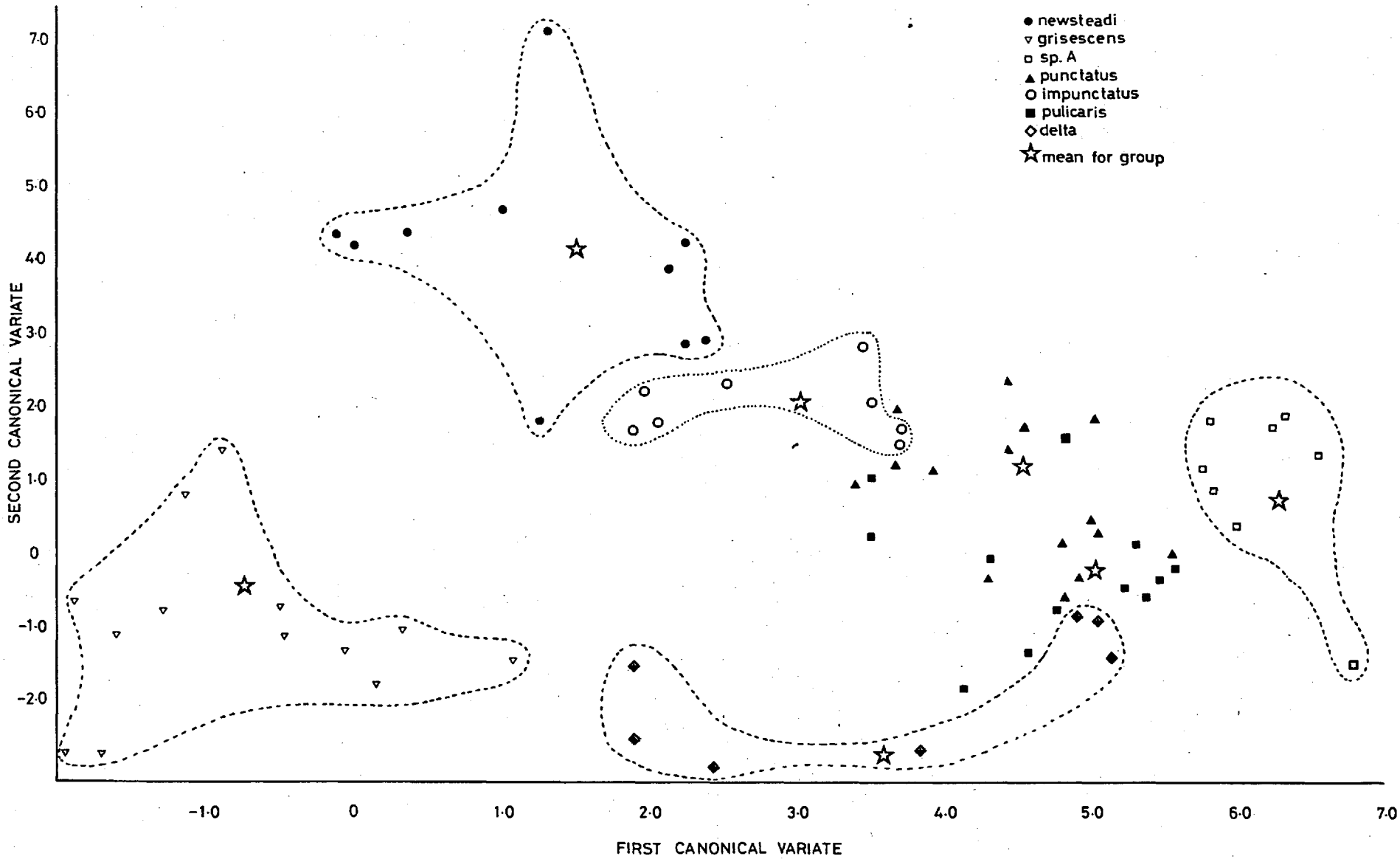


FIG. 72
CANONICAL VARIATES PLOT OF SEVEN TAXA - FIRST TWO AXES

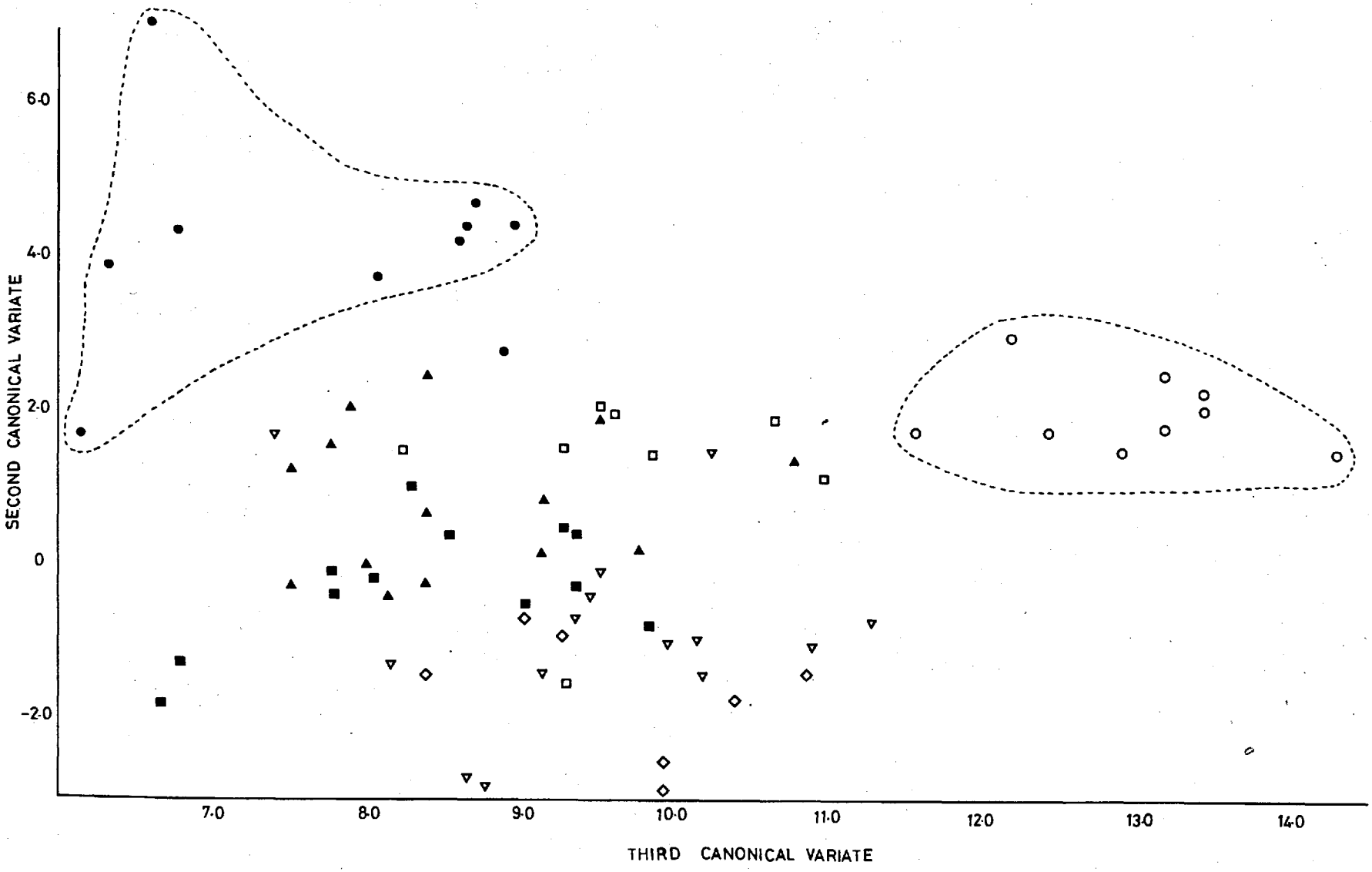


FIG. 73
PLOT OF SEVEN TAXA ON SECOND AND THIRD CANONICAL VARIATES

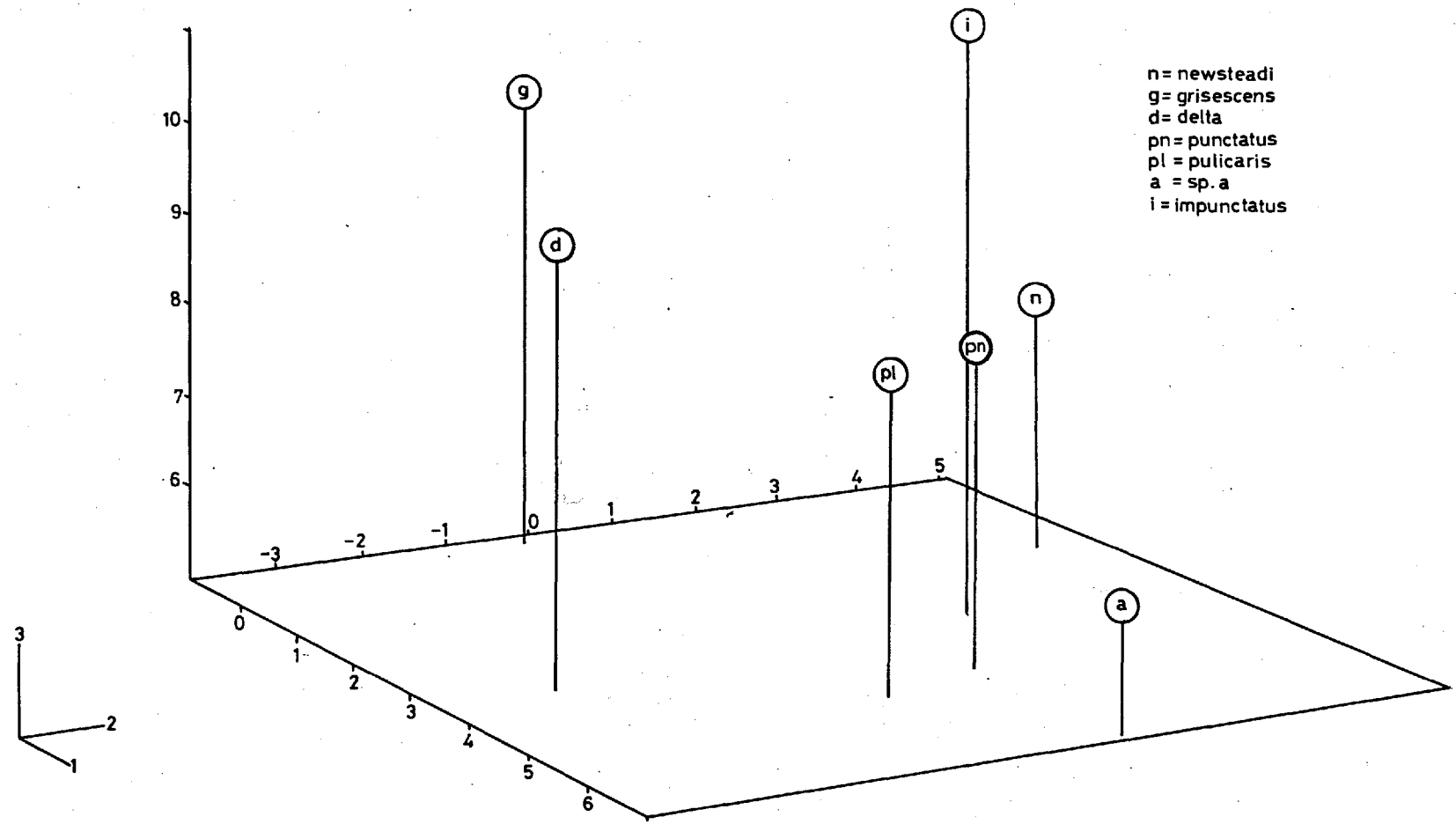


FIG. 74
THE MEANS OF SEVEN TAXA PLOTTED ON THE FIRST THREE CANONICAL VARIATES

wing features are not completely reliable, but are the only morphological differences on which this pair of taxa have been separated. Increasing the number of specimens should help to resolve the problem of whether the vague differences in wing pattern are associated with differences in any other characters. A summary of the changes in group means, as a consequence of increased sample size, is given in Table 42.

Overall, increase in sample size has led to relatively minor changes in the group means. Additional specimens of another common species, C. impunctatus, were also included. These additional specimens were provisionally identified by traditional means and then confirmed by large values on the third canonical variate of Fig. 73. Because the sample was increased in size from 9 to 24 specimens, there were some associated changes in the mean values for a few characters (Table 42). The differences are fairly small considering the sample size was increased nearly threefold. In this example, the small sample was fairly representative of the species. This may not be the case for other species and therefore a close check should be kept on the sample statistics whenever the sample size is increased. Another potential problem is the risk that new material brought into the analysis, after the main calculations have been done, may differ from earlier material in some attribute and for which earlier analysis has not allowed, i.e., the new taxa may differ in variables which were not originally measured. This risk has to be balanced against the benefits of increasing the sample size and improving the statistical validity of the results.

With the larger samples available for the common species, a new set of canonical variates was calculated. A summary of the analysis is given in Table 43 , and the results used to construct a three dimensional model (Fig.75). This model shows the 90% percentiles (= centours of Cooley & Lohnes, 1962) about the means of each species as flat discs. The calculation of these confidence intervals was described in Section 6, based on theoretical distributions about the means. The photographs in Fig.75 show two views of this model. The general shape resembles the results obtained from the previous canonical variates analysis (Fig. 74) with both impunctatus and grisescens well separated from other taxa. C. newsteadi, delta and sp. A overlap with other taxa to varying degrees. The last two species, pulicaris and punctatus,

TABLE 42.

Effect of Increase in Sample Size on the Means of Ten Characters

	punctatus		pulicaris		impunctatus	
Sample Size	14	40	13	33	9	24
Variable						
1	7.714	7.852	8.154	7.939	9.444	9.666
2	0.072	0.072	0.071	0.072	0.078	0.079
17	1.000	1.000	1.076	1.181	0.333	0.291
18	1.071	1.100	1.384	1.333	1.111	1.091
37	2.500	2.412	2.612	2.378	3.555	3.708
43	1.285	1.112	1.077	1.136	0.111	0.042
46	2.357	2.062	2.000	1.818	2.666	2.500
65	1.123	1.108	1.064	1.056	1.004	1.025
66	2.753	2.777	2.818	2.710	2.748	2.833
72	0.606	0.597	0.617	0.604	0.634	0.623

Variables may be identified from p.284.

TABLE 43.

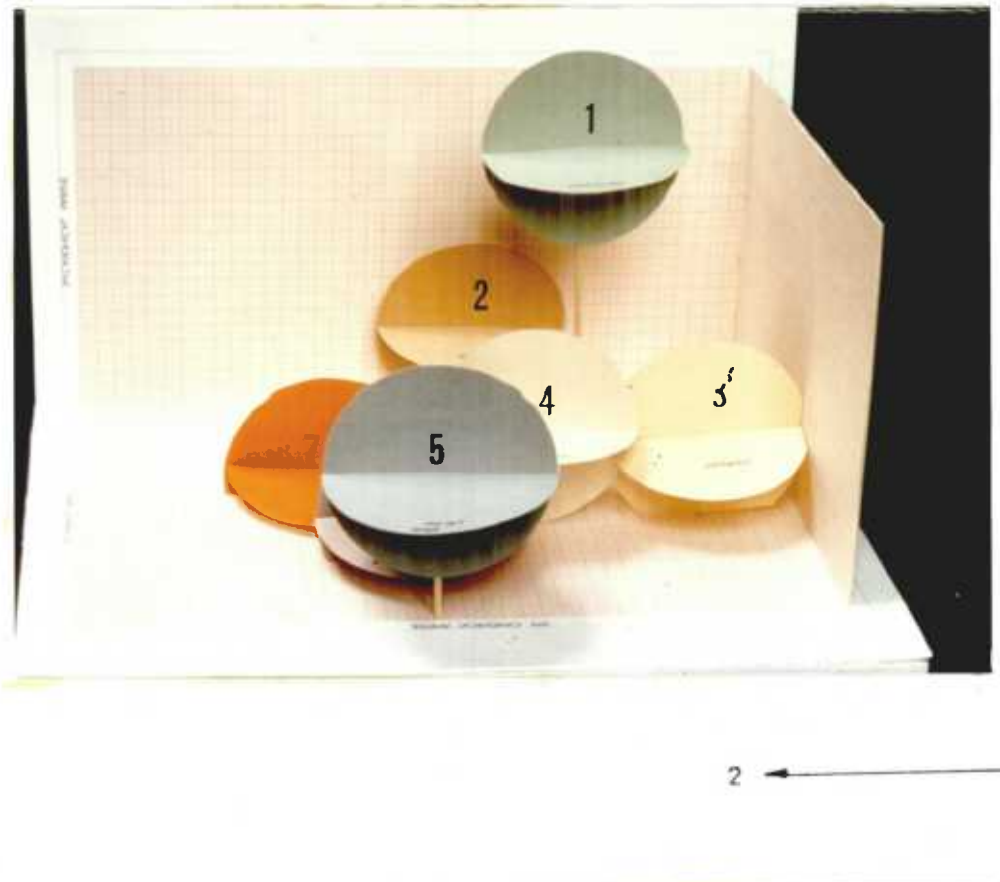
Summary of Canonical Variate Analysis : Seven taxa Based on Larger Samples

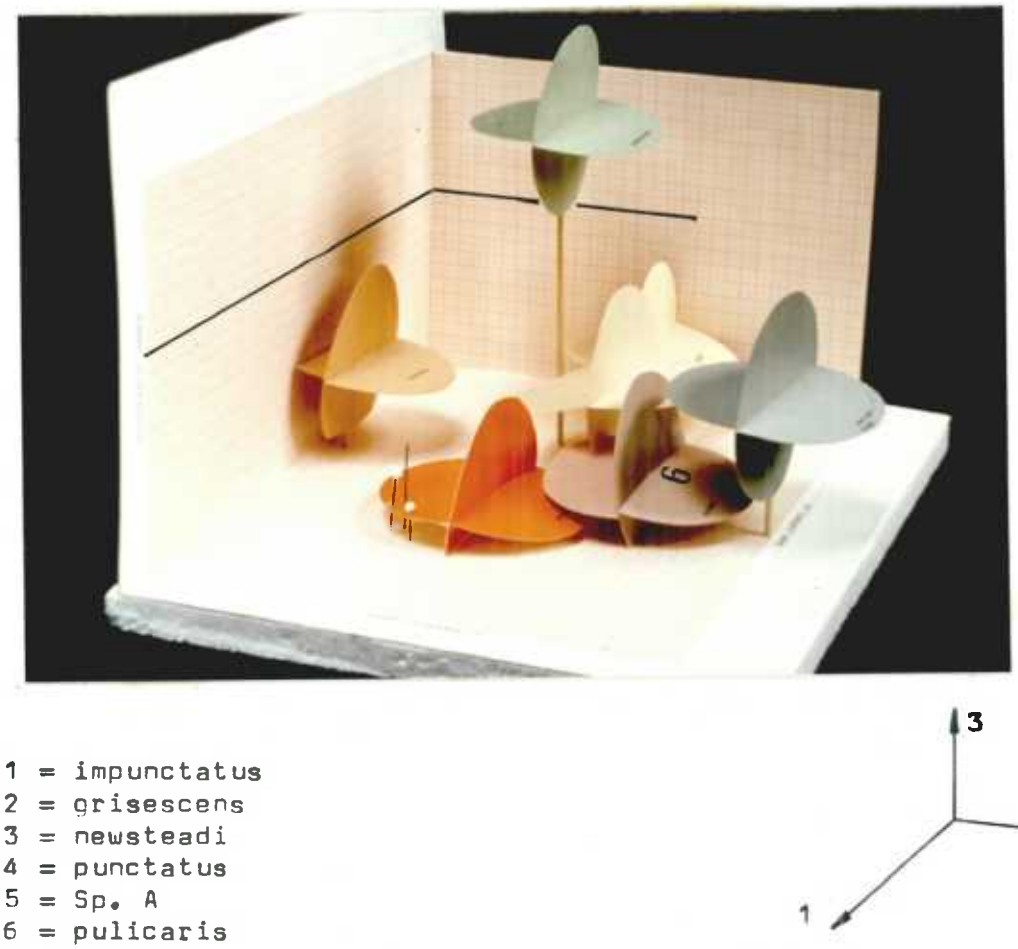
Variable	I Coeffic.	r	II Coeffic.	r	III Coeffic.	r
1	0.11	0.14	0.74	0.74	-0.03	-0.13
2	-19.67	0.03	-37.76	-0.04	-116.3	-0.48
17	0.15	0.00	-0.33	-0.08	1.15	0.50
18	-0.29	-0.04	0.47	0.09	0.33	0.20
37	-0.08	-0.15	0.11	-0.25	-0.83	-0.46
43	2.04	0.56	-0.58	-0.37	0.76	0.15
46	-0.38	-0.23	-0.48	-0.47	0.05	-0.06
65	1.28	-0.02	4.22	-0.01	-3.10	0.16
66	-1.78	-0.62	0.32	0.01	1.02	0.38
72	-6.69	-0.22	2.66	0.04	-3.07	-0.13

r = correlation of canonical variate with original variable

Variables may be identified from p. 284

	I	II	III
Cumulative Percentage of Trace	40.77	67.45	84.64





1 = impunctatus
2 = grisescens
3 = newsteadi
4 = punctatus
5 = Sp. A
6 = pulicaris
7 = delta

Fig. 75 Three Dimensional Diagram of First Three Canonical Variates : Analysis Based on Larger Samples of Seven Taxa

overlap with each other substantially. The 95% confidence limits are perhaps over-stringent; most traditional taxonomists probably work on 80-85% limits, although this is rarely stated. If the limits are relaxed slightly then spheres representing each species will contract, thus limiting some areas of overlap, e.g., between newsteadi and punctatus. There is little difference between those variables highly weighted in this analysis and those in the canonical variate analysis, based on smaller samples.

Canonical Variate Analysis Based on Six Taxa

As found in the analysis based on smaller samples, the difference between impunctatus and other species dominates the third canonical variate. In the hope of finding a taxonomically more useful set of canonical variates to separate the remaining six taxa, the analysis was repeated, omitting C. impunctatus. The removal of a distinctive group should allow attention to be focused on the variables which distinguish the remaining groups. A summary of this analysis is given in Table 44 . The group means are plotted on the first three canonical variates, together with their 90% confidence intervals in Figs 76 and 77 .

The first two axes of this analysis again account for 90% of the discriminatory power, providing a reasonable summary of the interspecific variation. One particularly interesting point is the extent of the overlap between pulicaris and punctatus. The means of the two species are within the 90% percentile of each other. As C. grisescens is so distinct in the first plot, it has not beendrawn in the second. The relatively slight overlap between sp A. and punctatus suggests that the former group is worthy of separate specific status. In Section 4, p.68, the difficulty in separating newsteadi from punctatus by traditional means has been discussed. The problem encountered was most disconcerting because these species breed in quite different habitats; newsteadi in saline mud and punctatus in marshes and bogs. However, the present analysis shows the species to be quite distinct and readily separated, thus concurring with the biological differences between them.

10.3. IDENTIFICATION USING THE RESULTS OF CANONICAL VARIATE ANALYSIS

General procedure

One of the main advantages of canonical variate analysis is

TABLE 44.

Summary of Canonical Variate Analysis of Six Taxa

Cumulative Percentage of Trace	I 55.85		II 89.02		III 95.93	
Variables	Coeffic.	r	Coeffic.	r	Coeffic.	r
1	0.24	0.29	0.59	-0.55	0.04	0.05
2	-14.79	0.04	-13.67	-0.07	-102.70	-0.04
17	-0.49	-0.08	-0.05	0.06	-0.25	0.02
18	-0.35	-0.05	0.41	0.16	1.25	0.59
37	-0.13	-0.22	0.25	-0.22	-0.58	-0.35
43	1.82	0.41	-0.53	-0.32	1.00	0.22
46	-0.51	-0.32	-0.63	-0.38	0.16	-0.11
65	1.18	-0.06	6.61	0.15	-13.29	-0.46
66	-1.71	-0.66	0.87	0.25	0.97	0.24
72	-2.77	-0.21	15.36	0.26	-1.83	0.01

Variables may be identified from p. 284.

r = correlation of canonical variate with original variable

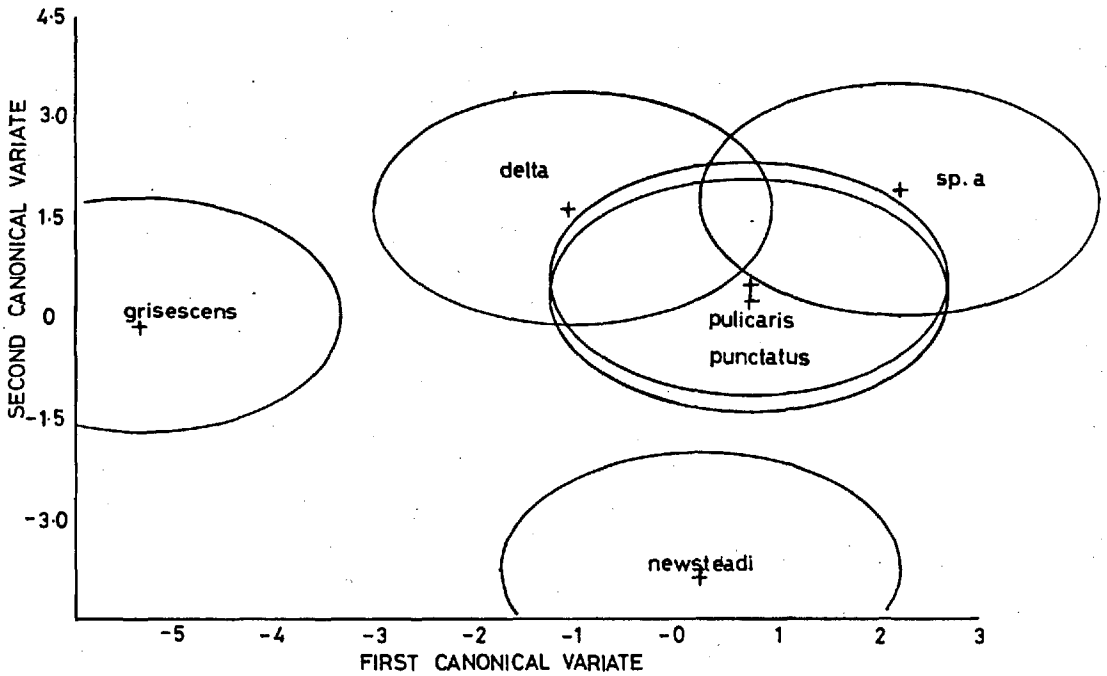


FIG. 76

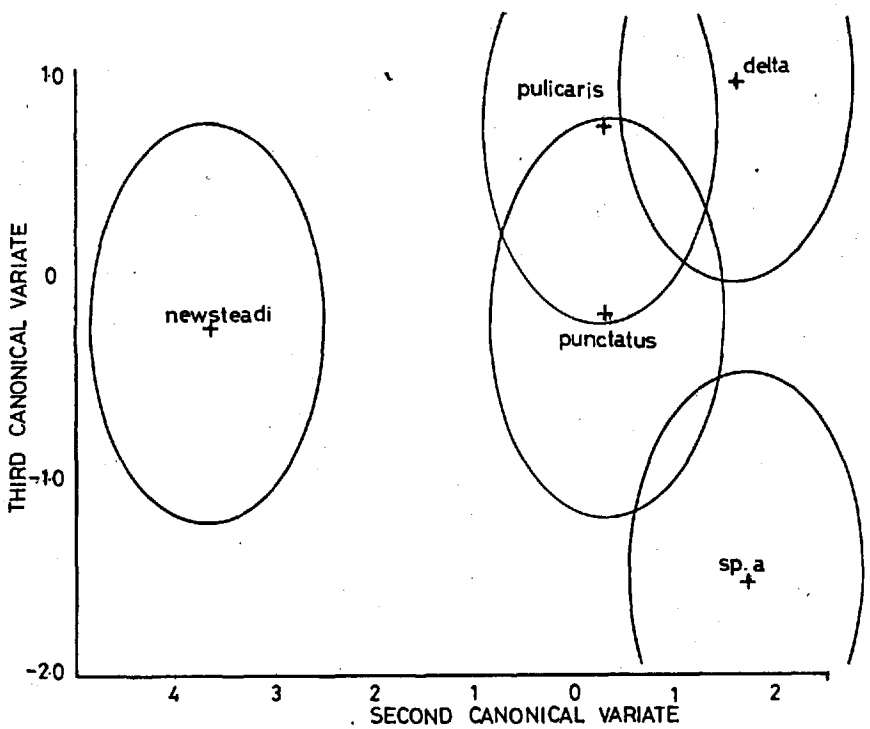


FIG. 77

FIGS 76-77
CANONICAL VARIATE PLOTS OF SIX TAXA INCORPORATING 90·1° PERCENTILES

that the results are easily used for the identification of specimens. Each canonical variate is defined as a linear combination of the original variables. The position of an individual on axis i may be calculated using the following linear equation:

$$y_i = \sum_{j=1}^{n} c_{ji} x_j$$

i.e., $y_i = c_{1i}x_1 + c_{2i}x_2 + \ldots\ldots c_{ni}x_n$, where c_{ji} is the j-th element in the i-th column of coefficients, and x_j is the j-th character value for the individual concerned. There is a different set of coefficients for each axis. For identification of an unknown, it is necessary to measure the n variables, and multiply each observed value by the corresponding coefficient. The process is repeated for the appropriate number of axes, to obtain a set of coordinates (the y values), which are used to place the specimens on a diagram. The position of a specimen is itself an accurate identification, but it is named according to its proximity to a group centroid.

Identification of Specimens in the C. pulicaris Complex

Specimens of C. fagineus are identified by the presence of cibarial teeth and therefore, unlike the other species in the complex, there is no need to use any of the canonical variate plots. Since the taxa cannot be separated by the use of only one plot, identification has to proceed in a series of steps.

Specimens of impunctatus are recognised first by using the third canonical variate from the analysis based on the seven taxa (Fig. 75) and Table 43 . Observed values of the ten characters given in Section 10.1., (denoted by subscripts) are substituted in the following equation:

$$y = -0.03X_1 - 116.3X_2 + 1.15X_{17} + 0.33X_{18} - 0.83X_{37} + 0.76X_{43} + 0.05X_{46} - 3.11X_{65} + 1.02X_{66} - 3.07X_{72}.$$

This places the specimens along the third axis. Although most specimens of impunctatus may be identified relatively easily by traditional means, the canonical variate diagram will identify those which lack a spot in wing cell 9. Such specimens may constitute 10% of a sample (see p. 182) and would not be accurately identified by traditional means. If the specimen is not impunctatus,

the next step is to attempt to place the specimen on the plot of canonical variate 1 versus 2 (Fig. 76), based on the analysis of six taxa. The position of the specimens along the first axis is calculated by the equation:

$$y_1 = 0.24X_1 - 14.8X_2 - 0.49X_{17} - 0.35X_{18} - 0.13X_{37} + 1.82X_{43} - 0.51X_{46} + 1.18X_{65} - 1.71X_{66} - 2.77X_{72}.$$

If the specimen is not grisescens, then one must proceed to use the plot of the second versus the third canonical variates (Fig. 77). The position along the second axis is calculated using the equation:

$$y_2 = 0.59X_1 - 13.7X_2 - 0.05X_{17} + 0.42X_{18} + 0.24X_{37} - 0.59X_{43} - 0.63X_{46} + 6.61X_{65} + 0.87X_{66} + 15.36X_{72}.$$

and the position along the third axis using the equation:

$$y_3 = 0.04X_1 - 102.7X_2 - 0.25X_{17} + 1.25X_{18} - 0.57X_{37} + 1.0X_{43} + 0.16X_{46} - 13.3X_{65} + 0.97X_{66} - 1.82X_{72}.$$

Using the calculated values of y_2 and y_3, the unknown may be placed on the diagram (Fig. 77). Specimens placed in areas of overlap cannot be positively identified. The fact that some taxa overlap does not preclude positive identification of a specimen as a member of a taxon, so long as it is placed in an area of non-overlap. For example, a specimen of delta may be positively identified as that species (and not pulicaris) if the value of y_2 is greater than 1.50.

The small white spheres in Fig. 75b represent specimens identified in this way. Although the method seems rather protracted, with the aid of the plots and a measuring eyepiece on a microscope, it is possible to identify a specimen accurately. Obviously the use of an electronic calculator makes the process easier because of rapidity in identification. Because of the simple equipment needed, and the limited skills required (for accurate measurement) this method of identification of a difficult complex is easier than techniques such as cytology or electrophoresis. Such a scheme for identification could easily be programmed so that a computer could undertake the calculations and determine the distance from

the nearest group centroid.

Two aspects of discriminant analysis that require further study are methods for selecting the best subset of variables, and the effect of adding new specimens to the diagram. Theoretically, incorporation of new specimens into the analysis should alter the estimate of the coefficients through the within and between-groups covariance matrices, and the means. In practice, this effect may be minimal, as found here when the sample sizes of some species were increased.

Section 11. GENERAL DISCUSSION

Discussion of the results and their inferred conclusions are organised around four main points: numerical techniques; taxonomy of the C. pulicaris complex; species complexes in Culicoides; and an alternative approach to Linnaean taxonomy for overlapping species.

11.1. EVALUATION OF NUMERICAL METHODS

(a) The C. pulicaris complex in relation to other taxonomic problems

The principal objective of this study was to assess the value of currently available multivariate methods in the analysis of groups which are difficult to separate into discrete clusters.

It was hoped that multivariate techniques would provide a clear summary of inter- and intra-specific variation in the C. pulicaris complex. The main methodological finding was however, that as the number of specimens included in the study increased, the interspecific relationships became, if anything, more obscure, and not clearer as had been hoped. The problems of discrimination are therefore more difficult than originally supposed - not an uncommon finding in the analysis of species complexes. To clarify this point and give some indication of how this taxonomic problem compares with others, it is convenient to outline the range of practical difficulties encountered in taxonomy. Whilst numerical analysis of taxonomic problems forms a spectrum from the orthodox to the statistical interpretation of a species, for clarity it is desirable to divide this range into ten classes.

Class i Groups are readily separated by traditional, univariate methods on a few obvious biometric characters or by qualitative non-biometric characters (the species of traditional taxonomy).

Class ii Groups can be separated as discrete clusters by simple bivariate plots of selected variables, without the need for derived variables.

Class iii Individuals fall into well defined, distinct, clusters when subjected to multivariate analysis not requiring prior definition of groups, e.g., principal component

analysis, principal coordinate analysis, hierarchical or non-hierarchical cluster analysis. Identification is achieved either by calculating the position of an individual in the first two methods, using the linear combination of variables, which define the principal axes, or by its inclusion in the primary data matrix and repeating the entire analysis.

Class iv Individuals can be assigned to groups on the basis of rather indistinct clusters from (iii). Using these groups discriminant techniques will then yield well defined clusters. The discriminant techniques may be straightforward, such as canonical variate analysis, or more sophisticated non-parametric versions. The latter may need development by a statistician.

Class v Individuals cannot be assigned reliably to groups as in (iii) but can be so assigned on the basis of some variable not included in the analysis, e.g., habitat, geographic locality, seasonal occurrence, or any qualitative non-structural character. When individuals are assigned to groups in this way, the discriminant techniques of (iv) work efficiently, but there is always the possibility that the defining variable is a result of inadequate or biased sampling. It should therefore be fully investigated.

Class vi As in (iv) or (v) but requiring an iterative re-location of individuals to form well defined clusters. Perhaps this is best undertaken in terms of probability of group membership.

Class vii Discriminant methods (used directly or iteratively) will not yield discrete clusters, though members of previously defined groups tend to occupy restricted zones in the discriminant space. Identification in this case may be made by calculating the distance to the nearest group centroid.

Class viii Like (vii) but groups are defined using numerical techniques that deliberately allow for overlapping, non-disjoint sets, e.g., B_k and C_u clustering techniques.

Class ix Individuals form an inextricable mixture, with no indication of groups, in some parts of the discriminant

space, though other individuals form well defined clusters. The latter are removed and the mixture re-investigated by methods described in Classes (iii) to (viii).

Class x No resolution possible by any of the above techniques. Variation is apparently all at the individual level. A search for new characters should be made of all kinds, or further dissection of the complex should be abandoned.

Originally the pulicaris complex was thought to be typical of Class iii, but when attempts were made to group the taxa by numerical methods and then identify new specimens, it was found to be more characteristic of Class vii. Thus, to some extent the complex remains unresolved in the traditional taxonomic manner, where specimens were identified as either species A or species B. Instead, identification was made in terms of distances from the nearest group centroid (i.e., by placing the specimen in the appropriate region of the discriminant space). In general, as one progresses through the categories outlined above, the less appears to be known empirically about the techniques, and the results become increasingly difficult to present in traditionally accepted taxonomic terms. This is one of the major reasons why alternative taxonomic concepts were sought and these are discussed below in relation to those in current use (i.e., a Linnaean or non-Linnaean system).

As far as the C. pulicaris complex is concerned, the capacity of multivariate statistical methods to resolve the taxonomic problem of this complex would have been regarded as most successful, if they had presented the species as discrete and well separated groups (category iii). Unfortunately this was not the case. The rather poorly defined boundaries between species, as defined by traditional characters, persist to some extent in the final results of the multivariate analyses. This appears to be a widespread taxonomic problem in insects, so that the results of this study have wider implications than merely resolving the difficulties encountered in the C. pulicaris complex.

(b) The number of characters used in numerical taxonomy

The appropriate number of characters to be used in a taxonomic study poses a difficult question, to which there appears to be no clear answer. Some aspects of this subject, together with some

techniques for reducing the number of variables have already been discussed in Sections 9 and 10. These may be expanded by a few more general comments.

In the early days of numerical taxonomy there was a general belief that the more characters on which a classification was based, the more reliable it is (Mayr, 1969). This view was later qualified by, among others, Rohlf (1962) and Sokal & Sneath (1962), who postulated the existance of various asymptotes. These provide an upper limit, beyond which additional characters contributed little, if anything, to the analysis. Rohlf (1962) suggested that the exact number of characters to be used in a numerical study depended on the precision required and that recourse to statistical theory would provide the answer. Recently, however, several workers (Blackith & Reyment, 1971; Clifford & Stephenson, 1975) have rejected the theoretical approach to selecting the appropriate number of variables, stating that "whether or not as more attributes are considered, the relative magnitudes of dissimilarity between populations undergoes drastic alteration, is entirely a matter for empirical investigation" (Jardine & Sibson, 1971:139).

The remarkable similarity found here between the classifications based on 72, 64 and 20 variables (Section 9) tends to contradict the recommendation that numbers of variables should be maximised in a numerical study. The classification based on 20 variables contained virtually all the types of characters originally considered. An important factor determining the number of variables for a numerical analysis is therefore not so much the absolute number, as the number of types of character, and presumably, the distribution of characters between these types. This conclusion is supported by several studies, typical of which is Stern*(1969). He found that from a total of 51 characters, a subset of 27 morphological, biochemical and karyological characters gave a classification very similar to that based on the complete set of characters. Furthermore, the reduced set produced a classification "which was certainly as satisfactory as one obtained by traditional means and possibly more suggestive of their [species] cause of development".

Another important factor controlling the optimum number of characters, is the nature of the taxonomic problem itself. Whenever all the OTU's associate into relatively homogeneous groups, with marked discontinuities between the groups, relatively few variables

*See Addenda

will be required to generate a useful classification. In contrast, if the data have a less well defined structure, the groups being less homogeneous, and not markedly different from one another, many variables are necessary to achieve a reasonable classification. (Clifford & Stephenson, 1975). To take this to its logical conclusion, if members of the pulicaris group were recognisable by different states of only one character (i.e., a monothetic method could be used to identify them), then perhaps only one character would be required to classify them! This situation is relatively common in insect taxonomy, where the structure of the male genitalia often affords specific separation and has consequently been used extensively in classification.

The number of variables used in a taxonomic work is inextricably tied up with character redundancy. Jardine & Sibson (1971) found it hard "to conceive of any general procedure for the elimination of redundancy in selections of attributes". One of the principal methods for assessing character redundancy is to investigate correlations between variables (see Section 9.4.4.). Although a study of correlations between variables proved a rather unsuccessful technique for reducing the number of variables, the information gained on the underlying structure of the data was very useful. Davies & Boratynski (1979) found that character association could be used in techniques for eliminating variables in a principal component analysis, if slightly different measures of similarity were used. The information gained from simple inspection of a matrix of character similarities is such that it should be an integral stage in most numerical taxonomic studies. This method revealed a substantial amount of character duplication, and hence character redundancy, in the data used here.

Blackith & Reyment (1971: 276) found that in studies of character redundancy in discriminatory topology and other descriptive numerical techniques, far fewer characters are needed in the former than the latter, i.e., identification requires fewer characters than classification. Although this observation was not tested specifically here, in general, the data for the C. pulicaris complex support it: 20 variables were required for an adequate classification, but a subset of only ten gave reasonable discrimination.

As in traditional taxonomic investigations, the deduction of homologies is important in determining not only the choice of characters, but the numbers used. This is perhaps more important

in numerical taxonomy per se , where an attempt is made to score each OTU for every character used. Sneath & Sokal (1973) give an extensive discussion of this topic in numerical taxonomy. When organisms are closely related, and members of the pulicaris complex certainly fall into this category, homologies are generally evident. But when the taxa are more distantly related, problems of homology arise and lead to the inclusion of characters which are not present in all the OTU's. In these circumstances the number of characters may have to be increased to include diverse structures which cannot be reasonably homologised. This problem is more frequently encountered in numerical taxonomy per se than in the multivariate morphometric approach employed here.

In conclusion, there are no definitive rules governing the appropriate number of variables to be used in a numerical study. The number actually required may be very much less than generally recommended, provided they are suitably chosen. Each problem should be considered in its own right and attempts made, particularly by inspecting the similarities between characters, to eliminate redundant variables.

(c) Establishing groups for discriminant analysis

In any study of closely related species, an overriding problem lies in establishing groups (species) prior to either their classification or discrimination. For most morphometric studies, the criteria for establishing groups may be divided into two broad categories - intrinsic and extrinsic. Methods using intrinsic criteria have only biometric characters or coded character states available, and therefore various numerical techniques must be used to arrange the specimens, or populations, into groups which may be given the status of species. The extrinsic criteria use characters from outside the analysis, usually non-morphological, to form the groups, which may then be separated solely on morphological grounds, or by a mixture of morphological and non-morphological characters.

'Intrinsic criteria' were used in this study for establishing groups: each specimen was treated as an individual OTU and a principal component analysis used to cluster them. The use of principal component analysis in this way has the advantage that it makes relatively few assumptions about the data (Blackith & Reyment, 1971; Marriott,* 1974:18) and is becoming a rather widely used

*See Addenda

technique in 'population phenetics'. However, there is a potential danger of circularity in using intrinsic criteria. Elmes*(1978) has drawn attention to the danger of using the same set of variables to discriminate taxa, as used in their recognition. In the present analysis, this circular argument was negated by using several sets of variables (from 20 to the complete set of 72) in the construction of groups (see Section 9). A smaller set of 10 variables was subsequently used to discriminate the groups so formed. The objective and subjective elements in this approach have been discussed in Section 9.

An alternative method for forming groups, and one that deserves further research, is that of Anderson*(1958). Initially, specimens are allocated to groups, perhaps on intuitive grounds, and then a set of discriminant functions are calculated. The probability of membership of a group is calculated for each specimen (based on scores on each of the discriminant functions). Thus, it is possible that a specimen provisionally allocated to species A is found to be nearer to species B. If the incorrectly allocated specimens are placed in the most appropriate groups, the technique can be used as an alternative to test the reliability of the groups formed here.

An abuse of multivariate methods in anthropology - of some interest to group formation by intrinsic methods - has been described by Corruccini*(1975). He refers to the apparently common procedure in the identification of fossil primate specimens, of producing discriminant functions based on modern taxa and then interpolating the fossil values into the function, to see which modern population they fall nearest to. Fisher's linear discriminant function was only designed to minimise the probability of mis-identification of an unidentified specimen into previously defined groups. It was not designed to indicate the relative affinities of parent populations (although it may provide a useful estimate) or to be applied to groups not included in the original function computations (Blackith & Reyment, 1971; Corruccini,*1975). The significance of this observation is that when groups are formed, by whatever intrinsic means, for subsequent inclusion in a discrimination analysis, the analysis will only allocate specimens to these groups. If another sample is believed to represent a new taxon, then a new set of discriminant functions should be computed, incorporating the new sample as a distinct group.

*See Addenda

It is possible that there are other ways of defining groups than those discussed here. If canonical variate analysis is to be used heuristically, there are no a priori objections to defining groups in whatever way one cares. The justification for this approach would be pragmatic: to produce an effective method of discrimination by canonical variate analysis. However, the disadvantage of such an approach is that groups may be formed which do not reflect any structure inherent in the data and, perhaps, therefore have little biological significance.

The problems of using extrinsic criteria, such as karyology, habitat, cross-mating tests, etc., for defining species in Culicoides have been discussed in Section 2.2. and of using morphology of the immature stages in Section 4.4. Despite these practical obstacles, alternative methods such as cytology or electrophoresis, so useful in other biting fly complexes, may give a more definitive result than that obtained from morphological data alone. An additional benefit of such techniques to obtain data for multivariate morphometric analysis, is their use of genetically homogeneous samples. When used in conjunction with canonical variate analysis, for example, it may be possible to produce an identification system expressed in morphological terms, but based on biologically defined groups. Alternatively, the electrophoretic and/or karyological data could supplement the biometric results in an analysis on the same line as that presented here. Blackith & Reyment (1971) have expressed some reservation however, concerning the inclusion of such drastically new characters into multivariate studies. They suggest that such data are likely to differentiate the material along new axes of variation, rather than add to the differentiation along the morphological axes of variation. This fear may be unfounded as Sneath & Sokal (1973:301), reviewing the congruence between numerical and biochemical studies, show that the two approaches complement each other rather well.

(d) Statistical assumptions of the multivariate methods used in this study

Numerical studies at low taxonomic levels, such as populations and variable species, emphasise a new series of problems from those encountered in the rest of numerical taxonomy. Such studies rely to a greater extent on the assumptions of multivariate

statistics, i.e., multivariate normality and equality of variance-covariance matrices (Sneath & Sokal, 1973). Jardine & Sibson (1971) maintain that statistical methods are only really applicable to subspecific concepts, because only at this level do relative probabilities of divergence or overlap bear on taxonomic decisions. Above the species level, relative degrees of difference are important, which formal statistics are not designed to indicate (Corruccini,* 1975). Despite these theoretical objections to the use of multivariate analysis on categories above the species level, reviews of numerical taxonomic studies show that, empirically at least, these techniques have been generally successful.

No useful test has been developed for comparing real data with the multivariate normal distribution (Cooley & Lohnes,* 1971; Blackith & Reyment, 1971) on which many of the techniques used in this study depend. Therefore it has either to be assumed that the data are normal or, if otherwise, that the statistical tests are not unduly affected. According to Burnaby (1966) the question of robustness in the presence of non-normality seems largely unexplored in multivariate statistics. This point is very important because there is evidence to suggest that multivariate normality is rarely found in biological data (Jardine & Sibson, 1971). Considerable disagreement exists as to whether multivariate techniques are robust enough to cope with differing data distributions. Reviews of numerical taxonomy (e.g., Sneath & Sokal, 1973) suggest the general success of these methods in diverse problems to be some indication of their robustness.

In addition to distributional normality, another major assumption made by multiple discriminant techniques is homogeneity of dispersion matrices. This is a stringent requirement and, in general practice, compliance is rarely found (Balakrishnant & Sanghvi,* 1968). Corruccini* (1975) provides a univariate analogy in the necessity for checking equality of sample variances prior to a 't' test for differences in sample means. This assumption (tested by an F test), like that of multivariate covariance homogeneity, is rarely tested. One reason why this issue of dispersion homogeneity is often ignored is that tests of the hypothesis are powerful and almost always reject it (Cooley & Lohnes,* 1971).

In this analysis, attempts were made to test the homogeneity of the dispersion matrices, but unfortunately some of the matrices were singular and/or had negative determinants, so the tests

*See Addenda

(due to Bartlett) were inapplicable. Therefore the assumption has to be made that the within-group dispersion matrices are homogeneous or the methods are robust. This result is particularly unfortunate because it is not possible to obtain an estimate of whether the canonical variates are the best discriminants, or if the technique is strictly applicable to these data. If the latter position is adopted, the choice of alternative discriminant techniques is very limited. A search for such techniques is beyond the scope of this study but nevertheless it remains a possibility. Non-parametric discriminant analysis is now often discussed in the statistical literature but there appear to be no worked examples in taxonomy or related biological fields. These methods will require development by statisticians before they can be used in taxonomy.

According to Blackith & Reyment (1971) there is "a body of empirical evidence available that suggests that this method [canonical variate analysis] may be moderately robust to departures from homogeneity". In a study of the evolutionary origins of the parasitic bees, Plowright & Stephen (1973) used canonical variate analysis exclusively and found that the within-group dispersion matrices were far from homogeneous ($p<0.01$). However, they regarded this as "irrelevent" because the results obtained were so reasonable (i.e., they duplicated the traditional arrangement of species). This is not as odd as it may seem, since, even with data that do not conform to the usual assumptions, satisfactory (though suboptimal) discrimination may be possible. Further, the significance tests based on the assumption of multivariate normality may not be required.

In conclusion therefore, the question of robustness of multiple discriminant analysis in relation to homogeneity of dispersion matrices and the application of non-parametric methods, are areas where further empirical work is of the utmost importance, though they do not deprive the application of standard methods of all value.

(e) Allometry and taxonomy

The taxonomic significance of allometry has already been discussed (p.122) both in general terms and in its implications for the taxonomy of the C. pulicaris complex. A few more comments

may be pertinant. Once allometry has been detected in the data it can be used in three ways:

- the allometric parameters can be used as taxonomic characters to distinguish groups or species.
- those variables which exhibit non-linear variation with size and are likely to be taxonomically unreliable, can be detected.
- attempts can be made to compensate for the effects of allometry and allow biometric comparisons between taxa based on 'pure shape' (sensu Corruccini, *1972).

The first two uses of allometric results have been applied to the pulicaris complex to provide valuable taxonomic information (see Section 7). The third application of these results however, is questionable, especially in taxonomic studies. Corruccini*(1972) boldly asserts that the influence af allometry in distorting data is an important reason why numerical taxonomy has failed in the eyes of some workers. He suggests that the use of correlation matrices in principal component analysis assumes the variables are linearly related, and if non-linear relationships occur (i.e., allometry) then the correlation coefficient will not quantify similarity of shape in a taxonomically accurate manner. Despite the relatively scarce occurrence of allometry (significant deviations from isometry) in the variables tested here, an attempt was made to implement the various adjustment methods of Corruccini* (1972). These normalisations may give an insight, in empirical terms, into the robustness of principal component analysis, when the variables are not linearly related.

Detailed examination of the two principal methods advocated by Corruccini shows them to be inappropriate. The first expresses each character in terms of a general size measure, and thus does not remove any allometric effect. The second method suggests replacing all the character states y_{ij} by $Y_{ij} = y_{ij}/x_i^{1-\alpha_j}$ (x_i is the general size measure for the OTU i). Algebraically it can be shown that this correction is the wrong way round, it increases the value of Y if α is greater than 1, instead of reducing it to eliminate the effect of allometry. The effect of this correction is to push the values away from the isometry line ($\alpha = 1$), rather than remove the allometric effects by pulling the values towards it. Later in the same paper (p. 381), Corruccini advocates replacing α_j, the allometry ratio for each character, by α/r_j (r_j not defined)

if the correlation between any character and overall size measure is low. Presumably, r_j signifies the correlation of a character and overall size. The use of this method cannot be justified unless all the variables are adjusted. A choice could be made with the use of significance tests, but these are sample-size dependent and may confuse the issue further.

Corruccini's methods were not applied to the Culicoides data because they were considered ineffectual. Apart from the possible problems which allometry may cause in applying the statistical methods (for which there is little empirical evidence), there is some doubt as to whether removing allometric trends has any biological (i.e., taxonomic) significance. For taxonomic purposes, the organisms or taxa should be compared as they are, not as they might be if they grew in a different way. Furthermore, difficulties may be encountered in making a biological interpretation of the axes of variation in a multivariate analysis, when corrected data have been used. In summary, it seems taxonomically inappropriate to seek further methods for removing the effects of allometry. Either the allometric parameters should be used as characters in their own right, or structures showing significant deviations from isometry should be omitted.

11.2. TAXONOMY OF THE C. PULICARIS COMPLEX

In the present work, numerical techniques have efficiently separated four of the species, but the cluster of pulicaris, punctatus and delta remains unresolved. A particularly successful achievement of this morphometric approach is the separation of punctatus and newsteadi (= halophilus) which, prior to the application of canonical variate analysis, were very difficult to distinguish. The similarity between pulicaris and punctatus has been confirmed by this study. Both show similar seasonal decrease in adult body size, a marked contrast to impunctatus, the only other species tested for this trait. The study of allometry of size, summarised by allometry gradients (p.135) reveals a remarkable likeness between pulicaris and punctatus. The allometric contrast between these species and impunctatus indicates that there is no uniform allometric description for the group as a whole. These 'univariate' studies, together with the multivariate analyses, cast doubt on the taxonomic distinction of pulicaris from punctatus.

The extent of their morphological variability, as described in Section 7, shows that the relative taxonomic status of pulicaris and punctatus cannot be resolved from morphological evidence. It may be that this pair of taxa (distinguished principally by wing venation) represents alternate morphs of a single species.

Many important characters for separating species of this and other complexes in Culicoides are on the antennae. The allometry studies revealed relatively little significant departure from isometry even though many calculated values of the allometric coefficient were larger or smaller than 1. This illustrates the importance of using significance tests in citing allometric exponents. The relative length of one antennal segment, segment iii, was found to be most important in the multivariate analyses used for classification and discrimination. This segment did not show allometric tendencies in any of the species tested. Another important antennal character in the analyses was the antennal ratio. This ratio is the sum of segments xi - xv (distal section) divided by the sum of segments iii - x (proximal section). The variances of the distal and proximal sections proved not to be significantly different and justifies the use of the ratio in statistical terms.

In previous taxonomic studies of the pulicaris group, considerabl emphasis has been placed on a spot in the cubital cell of the wing. The presence or absence of this spot was the basis of species identification. The results of the wing pattern analysis, and particularly of the discriminant analysis, showed the emphasis given to this character to be unjustified. Instead, the degree of pigmentation (an ordered multistate rather than binary character) of two other wing pattern elements around the medial fork were shown to be taxonomically more reliable.

11.3. SPECIES GROUPS IN CULICOIDES

As noted above, the results of this study, although equivocal for some species, enable one to discriminate between some taxonomically difficult species. The main advantages of using multivariate methods is their ability to summarise complex data and isolate those taxa for which non-morphological methods may have to be used for further separation. These conclusions are similar to those of Hensleigh & Atchley (1977) based on the economically

important C. variipennis complex in North America (see p. 99). Despite the limitations of multivariate morphometrics, their use would be most beneficial in other species groups for Culicoides for a more refined assessment of morphological variation. In the Palaearctic region, the nubeculosus and salinarius groups are possible candidates for a more comprehensive study. Perhaps a more pressing example is the austeni-milnei group in Africa. These species are important vectors of Dipetalonema in man and their taxonomy is very confused. There is evidence of morphological differences in this group associated with different breeding sites and distribution - milnei from inland sites and austeni predominating on the coast (Boorman, in press). In this case a multivariate morphometric study could be supplemented by ecological information, a combination which should prove most effective in understanding this complex. Another similar example is the C. imicola (= pallidipennis) group, which are vectors of bluetongue virus in Africa and the Middle East (Boorman, personal communication).

In those groups with sympatric species, and they are the majority, discrimination between the components may be enhanced by placing more emphasis on the use of specimens collected in the same habitat (locality). Such samples would presumably have been subjected to relatively similar environmental variables and yet are likely to exhibit character displacement. This phenomenon was defined by Brown & Wilson (1956) as "the situation in which, when two species overlap geographically, the differences between them are accentuated in the zone of sympatry and weakened or lost entirely in parts of their range outside this zone". In the present study the species were sympatric over most of their ranges and therefore the principle was not so easy to apply. This does not seem to be the case however, for some of the groups mentioned above, e.g., imicola and austeni groups.

11.4. AN ALTERNATIVE APPROACH - SPECIES AS NON-DISJOINT SETS

The taxonomic problems of the C. pulicaris complex are typical of those encountered in many other groups of insects. The conclusions drawn from this study may therefore be used as a basis for a discussion on more general apsects of morphologically indistinct species.

In both conventional taxonomy and most forms of numerical (phenetic) taxonomy, there is one unifying concept - that species

form disjoint (mutually exclusive) sets, which may be ordered in a hierarchical manner. Although this axiom has been the basis of most taxonomy, it does have considerable restrictions when dealing with highly variable species and species complexes. Usually all taxonomic observations are fitted into this model, even when the data suggest that it may not be wholly appropriate. Using the principle of mutually exclusive sets, naming a specimen (or population) as Ab implies that it is not a member of the species Ac or Ad, i.e., the identity population cannot be expressed in terms of more than one species name. Furthermore, the suggestion is made (although obliquely) that species are understood in such a manner that it is possible to decide whether an unknown population belongs to one or another - it cannot belong to more than one species at a time.

To some extent this axiom is modified by the definition of a species used by the taxonomist. A range of definitions and attitudes used in the study of the Diptera has been outlined in Section 1, p.18 . Although different criteria are used to define species, they are all based on the concept of mutually exclusive sets. When data are obtained which do not comply with a previous definition, instead of examining whether the basic axiom is appropriate there is a tendency for another specialised definition to be advanced. Despite continual efforts to discover more refined means of separating species (morphology, physiology, immunology, cytology) it has not been possible to find absolute criteria for their recognition. Species may be separated by one or more of these methods but none of them seems to apply to all species.

Dobzhansky (1972) suggests there are many different types of species and therefore searching for universal properties of all species is futile. Similar opinions are held by Scudder (1974).

The problems of overlapping species and the possibility of developing a non-Linnaean taxonomy raises two important questions.

1. Do species exist in nature as disjoint sets?

2. In our taxonomy - which is an attempt to model nature - need we restrict ourselves to using mutually discrete sets, when the data do not always suggest it necessary?

These questions correspond to the two principal levels on which species concepts are commonly used - the biological level and the taxonomic level, and therefore the two questions will be discussed

in relation to these topics. Although these are obviously related, the application of the points raised in one level does not preclude their application to the other.

Biological Aspects

With regard to the first question, much modern thought on speciation in genetical and behavioural terms tends towards considering the typical or centralising mechanisms. For example, Patterson (1978) defines a species in terms of a common mate-recognition-system, i.e., a system which keeps the members of each species together with the incidental effect of keeping the speciesthemselves apart. Although it has the same effect, Patterson's definition contrasts to Mayr (1969) who postulates the existance of isolating mechanisms between species. The basic difference between the two definitions is the emphasis given to the way in which the species maintain their integrity. With such concepts, it is possible to consider species in terms of their central tendencies and not in terms of their 'edges'. It may be useful to recognise the limits of a species but it is more important to know their centres. This point will be expanded later in this argument, but suffice to say at this juncture, that where species overlap significantly, for any type of character, the boundary between them is very difficult to define taxonomically. For practical taxonomic purposes the suggestion is therefore, that if mean values for a suite of characters are known (e.g., centroids in a canonical variate diagram) it is more convenient to identify new specimens relative to these central points, rather than trying to place the specimens on either side of an indistinct boundary.

Whether species exist in nature as disjoint sets is more difficult to answer. Clearly, many species do, but this need not always be the case. Whatever mode of speciation is accepted, allopatric, sympatric, parapatric, etc., there must always be some stage when 'species' begin to separate. Under these conditions, the subdivision of evolving lineages into successive species must, to a large extent, be arbitrary (at least for sexual species).

The problem of taxonomically defining such nebulous units as species is particularly difficult when geographic variation is considered. Some workers have a tendency to call every morphological

distinct geographical isolate a species. The problem of defining such nebulous units as conventional taxonomic species is similar to that outlined above for evolving lineages. There are, in principle, no clear divisions between populations, subspecies and species, they form a continuum which is cut up, arbitrarily or by extinctions, into discrete segments suitable for taxonomic purposes. According to some, geographic variation is the very fabric of evolution. It is now becoming increasingly clearer that neither species, nor subspecies evolve but only populations (Ehrlich & Raven, 1969). The fact that subspecies are not always, or even usually, discrete and that they may be connected by transitional populations is itself of considerable biological interest. Opinion differs on the extent of gene flow between populations (see Dobzhansky, 1970; Ehrlich & Raven,*1969), but again there do not appear to be any hard and fast rules. To think that subspecies do not exist purely because they are not fixed units is fallacious typological thinking. One of the most important lines of evidence in this field is the observed geographical variation in genes and character frequencies. In many cases, variation in gene frequencies have been found in a continuous manner along a geographical cline (e.g., in North and South American Drosophila, Dobzhansky, 1970). Thus an accurate definition of a subspecies is difficult because the gene pool of any one population is a genetic system adapted to the environment that population inhabits. The genetic systems of different subspecies are therefore adapted to different environments. Because of the contoversy over the extent of gene flow between populations, the emphasis given to genetic criteria, although objective, does not at first glance, provide an absolute criterion of a subspecies.

Mayr (1969, 1970) discusses the uses and abuses of the subspecific category, and suggests that, despite the shortcomings of the category, it should be retained as a means of referring to geographically isolated populations, which are distinguishable from other populations of the same species. To the taxonomist therefore, the problem is one of deciding when does intraspecific variation become interspecific variation.

The growing disparity between the absolute nature of taxonomic categories and the findings of modern biology was summarised by Sewall Wright*(1978 p. 8) as "hierarchical classification of organisms is a practical necessity for biologists but it does not

*See Addenda

accord with the continuity of life in space-time". The ever increasing emphasis on populations, and other infra-specific categories, in many biological disciplines puts a greater pressure on taxonomy to incorporate these ideas. It is therefore most important for taxonomic systems to be developed which are capable of recognising units smaller than a species as well as handling species which may overlap in some attributes.

Taxonomic Aspects

The preceding discussion has raised the general point that in nature, species (including subspecies and semispecies) show sufficient variation (either morphological or genetical) for them no longer always to be considered as disjoint sets.

Dobzhansky (1972), Scudder (1974) and others have reacted quite forceably against the rather naive belief that there is only one generally applicable definition of a species. There is no *a priori* reason to suggest that there should only be one type of species. Different sorts of species should be understood in relation to different strategies of evolutionary adaptation. Although neither Dobzhansky nor Scudder have suggested that the disjunct-set basis of species should be altered, it is precisely this point which may be modified for taxonomic purposes.

Whether or not species exist in nature, as disjoint sets, the taxonomic methods of manipulating species or 'species concepts' is a different (though related) question. Much confusion would be avoided if the duality of the species concept (the biological and taxonomic) were understood clearly (Dobzhansky, 1972; Paterson, 1978). Although these concepts were created for different purposes, the taxonomic use of the species category would be enhanced if it accords with nature, and yet remains practical. The effectiveness of the species concepts used by taxonomists is governed by their ability to incorporate some of the 'dynamic' aspects of species in nature, as well as providing a realistic description of the data available. Some of these data, including the results of the present study, indicate that the techniques and concepts of taxonomy should be expanded to incorporate ideas of sets (species, subspecies) without distinct boundaries. Traditional taxonomy works at a level which attempts to find the limits of species and distinguish between them (Fig.78). In contrast, biological trends

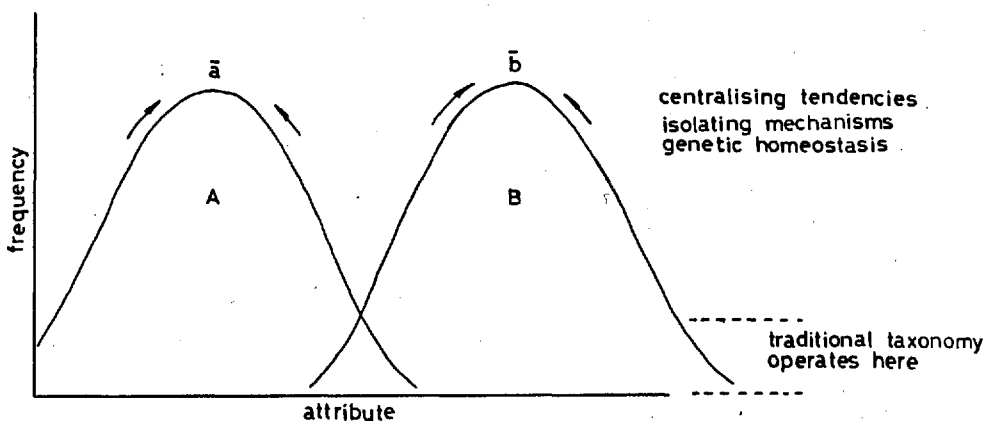
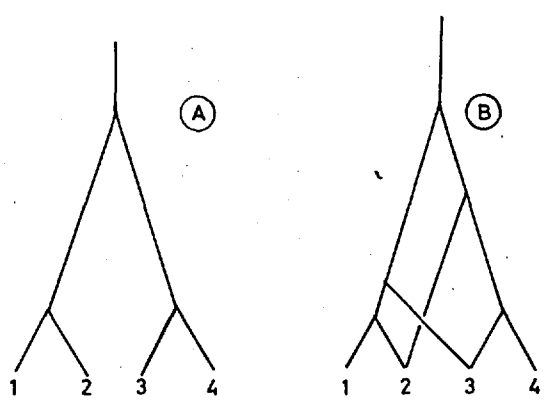


FIG. 78



A hiearchical non overlapping dendrogram

B overlapping cluster dendrogram

FIG. 79

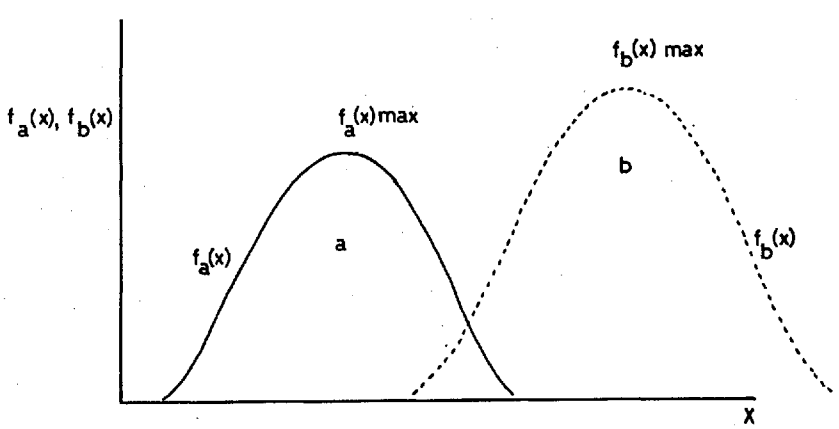


FIG. 80

FIGS 78-80 SEE TEXT(section 11.4) FOR EXPLANATION

are 'centralising' in so far as they attempt to consolidate the species (e.g., gentic homeostasis, specific mate recognition systems). The proposal advanced here therefore, is that for taxonomically difficult complexes of species, greater stress is given to the central features of species ($\bar{a}$, $\bar{b}$, in Fig. 78): the species are considered as groups with indistinct boundaries so that transition from one taxon to the next is gradual rather than abrupt.

To emphasise the impact on present day taxonomy of using sets with indistinct boundaries, the remainder of this discussion will outline how some available techniques may be used to put this suggestion into practice. Before these techniques are discussed, however, it should be made quite clear that the rejection of species as disjoint sets is not a disguised form of nominalism. Nominalists deny the existance of species, maintaining that they are only constructs of the human mind. A surprising number of biologists have supported this philosophy (see Ghiselin, 1974), but it has been strongly contested by, among others, Mayr (1969), on biological grounds, and Ghiselin (1974) on logical grounds. Treating species as groups with indistinct boundaries does not suggest that only individuals exist. The point of view adopted here is that species undoubtedly exist, the only difficulty being in defining them.

Species as Overlapping Sets

The first of three techniques discussed in this section is statistical in nature and is the most easily applied and interpreted.

Dupraw (1964, 1965a, 1965b) described a system of classification which he termed non-Linnaean taxonomy. Basically, this system used multiple discriminant function diagrams (similar to those produced in the preceding section on discrimination and identification, Section 10) as the classification. Individual specimens are points plotted on the diagrams. They usually form areas of high density linked to other similar areas by intermediate zones of lower density. Traditionally, the areas of high density would be circumscribed with more or less arbitrary boundaries. This generalisation from actual specimens to the concept of a

given species may introduce serious informational artifacts. In the case of distinct species (or a few individuals) with no observed intermediates there would be no problem. In this simple situation, the effort involved in constructing a discriminant diagram is unnecessary and unlikely to be attempted. The method is most useful in those cases where species overlap (either morphologically, physiologically, or cytologically). In such cases the groupings of specimens is poorly defined but the traditional taxonomic concept of the species would imply that they were sharply defined and mutually exclusive.

Identification of an unknown specimen is made using the linear combination of variables defining each discriminant axis (described in Section 10, p.297). The merit of this method over earlier methods is that if a specimen is identified, its position on the plot is fixed exactly by virtue of its character states. The identified specimen may be specifically referred to by its coordinates, or in more general terms, by its proximity to a nearby cluster. Under these circumstances, it might be referred to as Species A, if that was the nearest cluster with a binomial name. The coordinates of a specimen on the discriminant plots constitutes the nomenclature in the given classification, (both for the individual and the group to which it belongs). The main point is that this system does not rely on demarcating the edges of a cluster of a particular species, but only its centroid. Consequently this system is most useful for problems encountered in the C. pulicaris complex, in which our biological information is scanty and identification of specimens difficult. As stated above, the classification of the C. pulicaris complex, expressed in terms of multiple discriminant functions, is more accurate than previously available arrangements. Specimens or populations in areas of overlap between two or more species can be referred to objectively, without having to decide arbitrarily to which species the intermediate population belongs. Thus, although the identification and reference to a specimen takes longer than in conventional taxonomy, it gives a very practical and effective method for discussing taxonomically difficult species.

Another important advantage of this method is the transference of biological information associated with intermediate specimens. In traditional taxonomy, intermediate specimens are usually referred

to one or other of the relevant taxa, and in doing so, the biological characteristics of a population may be lost or become more difficult to retrieve later. This is because the information is 'filed' under the name of only one of the alternative taxa.

Gyllenburg*(1965) has developed an analagous method for the identification of microorganisms. A new specimen is identified by its distance from a group centroid in an ordination plot (in this case, a principal component plot).

Dupraw applied his technique most effectively to geographical variation in honey bees, where several subspecies are recognised. This technique has helped greatly in understanding the C. pulicaris complex, and at present is the most efficient method by which specimens may be identified and intermediates referred to. Using this system for the classification of the C. pulicaris complex would ease many of the current taxonomic problems, especially the status of intermediates. For example, a sample from Norway was difficult to place as either punctatus or newsteadi. Using discriminant function diagrams, it is possible to show exactly how this sample is related to the two species and therefore to avoid the necessity of assigning the sample to either group. As discussed above, the use of significance tests to augment a decision (if one is needed) strictly requires the variance-covariance matrices to be homogeneous, which places some restraint on this method. However, the results from dispersion matrices which are not strictly homogeneous would probably still provide a better system than is currently available. There are many examples of species groups to which this taxonomic approach could be profitably applied. An example is the Euxoa declarata group (Lep. Noctuidae), in which Harwick & Lefkovitch*(1973) found three sympatric species, by multiple discriminant function analysis. Despite repeated analysis of the data, there remained overlapping areas between the species, in which specimens could not be identified. If these workers had used the diagram as a classification, rather than just a tool of identification (like a diagnostic key), they could have referred to intermediates accurately. The intermediates are of some interest because, although naturally occurring hybrids have yet to be found, some crosses have produced vigorous hybrids in the laboratory.

Whether a system based on multiple discriminant functions is really non-Linnaean, as Dupraw suggests, is open to some doubt.

*See Addenda

The basis of the discriminant methods he advocates requires groups to be accurately defined, a step quite in keeping with Linnaean systematics. Subsequent identification however, does not require new material to beassigned unequivocally to a group, but allows it to be considered in its own right, perhaps as an intermediate. For the C. pulicaris complex, this system of classification and identification is most satisfactory, given the rather limited samples available. It allows such distinct species as impunctatus, grisescens, and newsteadi to be included in the same diagrams as the overlapping species, pulicaris, punctatus and delta. Furthermore, the slight morphological variation in samples from widely separated localities can be incorporated into the statistical limits of the groups, without having to propose new scietific names for them.

Although the taxonomy of the C. pulicaris complex has been considerably improved by the use of multiple discriminant analysis, it may be possible to improve it further by developing a truly non-Linnaean taxonomic system. Such a system would allow groups to have indistinct boundaries and individuals to have membership of more than one species. Because the concept of hierarchical classification seems to break down at the specific/intra-specific level, in practical terms at least (Sneath & Sokal, 1973), a non-hierarchical system would therefore be beneficial. The taxonomic system based on discriminant functions fulfils this requirement, but it does have the disadvantage of having to define groups initially. Two techniques which do not require definition at the outset are non-hierarchical cluster analysis (especially those of Jardine & Sibson, 1971) and the theory of 'fuzzy subsets'. Both these methods establish groups by adjusting parameters specific to the technique. Neither of them has been applied to the C. pulicaris complex, because of computational difficulties, but they are nevertheless worthy of further discussion as lines along which the taxonomy of species complexes may perhaps progress.

The first of these alternatives includes the family of clusterino techniques termed B_k and C_u methods by Jardine & Sibson (1971). Most algorithms for cluster analysis lose all record of the internal structure of the groups formed, by treating the groups as a new set of objects at each stage of the clustering process. The B_k and C_u methods were developed to produce a non-hierarchic arrangement of OTU's, in which a certain amount of information

about the internal structure of groups is retained. The B_k methods are based on single linkage cluster analysis in which some absolute restriction of the overlap between groups is permitted, e.g., the actual number of OTU's permitted in the overlap between clusters at any given rank. The dendrograms produced by single linkage cluster analysis and non-hierarchical clustering analysis are given in Fig. 79 for comparison. Although the B_k methods are theoretically elegant, several problems are encountered in the graphical presentation of the results, as the number of OTU's permitted to overlap at any given rank is increased. Jardine & Sibson(1971) suggest that some ordination may be necessary to portray the relationships between the vertices of the graphs effectively. These authors also describe another family of non-hierarchical techniques, C_u techniques, which control the extent of cluster overlap by relating the overlap diameter in terms of some intrinsic factor, such as the magnitude of the dissimilarity coefficient. This method appears to be most useful where well defined groups are present with only a few intermediates (Jardine & Sibson, 1971;157). Therefore, the choice of either B_k or C_u methods depends on the taxonomic problem in the species group being studied. For the C. pulicaris group, the B_k methods look most promising because of the extensive overlap between pulicaris, punctatus and delta. The type of problem for which the C_u methods seem most appropriate are probably those in which the species are fairly well defined, and therefore such a sophisticated technique is unnecessary. Traditional taxonomic methods or simple ordination techniques would probably suffice. Although these two clustering techniques represent overlapping clusters quite effectively, the difficulty of using them in species complexes is the subsequent identification of specimens. At present, the only way to identify a specimen is to include it in the primary data matrix and repeat the analysis.

Further work on adding individual specimens to the clusters, if only approximately, would contribute greatly to evaluating how useful these techniques would be in routine taxonomy.

The final method for discussion, and perhaps one with great potential after some development, is to consider species as 'fuzzy sets'. The theory of 'fuzzy sets' was developed mathematically by Zadeh*(1965) for electronic information transmission and has subsequently been expanded by Kaufmann (1975).

*See Addenda

The main difference between this system and some previously discussed (canonical variate analysis), is that the source of imprecision is the absence of sharply defined criteria of class membership, rather than the presence of random variables. Fuzzy sets are therefore completely non-statistical in nature. Basically, the accuracy of the methods depends on allowing individuals to have overlapping identities (in a mathematical sense). There is a philosophical, as well as procedural, difference between fuzzy sets and statistical methods. To explain these differences, it will be helpful to outline a basic property of fuzzy sets. For x individuals, a fuzzy set A is characterised by a membership function $F_A(x)$. This function associates each of the individual OTU's with a real number between 0 and 1. The value of $F_A(x)$ is the grade of membership of x in the fuzzy set A. This is shown graphically in Fig. 80. Returning to the comparison of fuzzy sets and statistics, although the membership function superficially resembles a probability of membership, it does in fact have quite a different meaning. For example, $F_A(x) = 0.25$ does <u>not</u> indicate a belief that the likelihood is 0.25 that x belongs entirely to class A, rather it states that x <u>shares</u> 0.25 of the qualities necessary for unequivocal membership of class A (Bezdek,*1975).

The membership function may have maximal values, which in the diagram represent the modes of A and B. Therefore, this complies with the suggestions made earlier concerning the need to concentrate on central properties of species (Fig. 77). If the concept is applied to an Euclidean space (see Kaufmann for further details) then the membership function is a measure of the distance from a group centroid. Recently, Bezdek*(1975) introduced the mechanics of the theory into numerical taxonomy, but made no comment on the significance of the method for systematic theory. The subject has since been discussed more fully by Zadeh*(1977). Despite the similarity of fuzzy set theory to the overlapping cluster analysis (B_k analysis) of Jardine & Sibson, little appears to be known of either theoretical, or empirical, relationships of the methods (Sokal,*1975;173).

One important point to make is that this theory is general, in that it also caters for the existence of quite discrete taxa. In these circumstances, the taxa would probably be termed mutually exclusive for practical purposes, in so far as membership of one

*See Addenda

species by a specimen, precludes almost certainly, membership of another species. This is the basis of much traditional taxonomy, which is clearly seen to be a special case of the fuzzy set concept. (the membership function for another set is very low, or zero, and the distances between centroids is correspondingly large).

Before this theory can be fully incorporated into the techniques of biological classification, further work needs to be done by methematicians, on adding new points to the analysis (identification). It was suggested early in this discussion that the present taxonomic system does not take adequate account of species in statu nascendi, geographically discrete populations, etc. It is exactly this area in which the theory of fuzzy sets seems to have most potential, theoretically at least. The significant point is that a segregate in the process of speciation is not believed by a biologist to be statistically part of the parent population, but to possess many features in common with it, as well as a number of unique properties.

Similar reasoning applies to the position of a hybrid relative to its two parents. Before this technique is applied in an analytical capacity in taxonomy, thorough empirical studies, using populations or specimens of known identity, will have to be undertaken.

To summarise, the results of this study indicate that traditional Linnaean taxonomy, based on disjoint sets, cannot always adequately accomodate morphologically variable species. Furthermore, if taxonomy is to benefit from the advances in evolutionary biology, then a system must be developed which is capable of recognising indistinct species as well as smaller units, such as populations, geographical isolates, etc. The traditional system was developed prior to Darwinian concepts of evolution, and is unable to incorporate incipient species in a realistic manner. It is suggested, therefore, that the axiom of mutually exclusive sets, on which the species category is based, be relaxed to allow for species overlapping in terms of morphological and other character suites. One procedure which would allow this depends on multiple discriminant analysis, the merits and limitations of which have been discussed in relation to the C. pulicaris complex. Other methods of overlapping cluster analysis are, however, becoming available and their potentialities are touched on. The alternative techniques proposed here do not alter the

status of disjoint sets, or question their widespread taxonomic applicability, but suggest a more general framework in which species with indistinct boundaries are included. Some such system could be used to augment Linnaean taxonomy at low taxonomic levels, and although it may change some nomenclatural and taxonomic procedures, the substantial improvement it brings in practical terms justifies these changes.

SUMMARY

This study assesses the value of currently available multivariate morphometric techniques in the analysis of the Culicoides pulicaris complex. This midge complex is typical of species groups which are difficult to separate into distinct species.

The introductory section outlines the relationships of Culicoides and their veterinary and medical importance. Species concepts applied in studies of the Diptera are reviewed, emphasising that the diversity of concepts are based to some extent on practical, rather than theoretical considerations, and that sibling species are not different in any way from other species, but merely near the invisible end of a broad spectrum of diminishing morphological differences between species (p.25).

Throughout a description of the morphology of Culicoides (Section 3), structure is related to function, particularly for the humeral pits (p. 49) and shape of the third palp segment (p.44). Scanning electron microscopy shows that pigmentation is responsible for wing pattern, and not spacing of microtrichia, as previously suggested (p.51).

Section 4 gives a formal taxonomic review of the eight nominal taxa in the pulicaris complex. The complex is defined and the taxonomic relationships to closely related species in both the Palaearctic and Nearctic regions are discussed (p.59,75). The complex is distributed throughout the Palaearctic region and the recorded distribution of each species is given. The record of this complex from the Afrotropical region is shown to be a misidentification of a closely related species - C. brucei (p. 74).

The necessity for accurately defining and standardising measurements is stressed in the outline of techniques (p. 82).

Prior to numerical classification and discrimination of species, the variation in several important quantitative characters is investigated, to determine their taxonomic reliability. The main findings are:-

1. C. pulicaris and punctatus show a seasonal decrease in body size (measured by wing length) during spring and early summer in Britain. A minimum size is reached during mid-summer, which is maintained for the remainder of the season. C. impunctatus shows no seasonal variation in size (p. 101).

2. A 'homeostatic' mechanism controlling the length of individual segments occurs in the antennae. An increase in the length of one segment is compensated for by a decrease in the length of a subsequent segment, ensuring that asymmetry in the proximal and distal sections of the antennae is minimised. This homeostasis is suggested to be a consequence of morphogenesis in the pupa, which may be of selective value because of the functional significance of the antennae as sound receptors (p.111).

3. Although there is some variation between the lengths of complementary segments of right and left antennae, the differences between antennae of one individual are smaller than between different species. Caution should be exercised in the use of such characters for distinguishing species - they should preferably be used in conjunction with other characters (p.119).

4. The expression of segment lengths as a proportion of total length is no more or less reliable than using absolute lengths. Therefore, for practical purposes, the use of proportional lengths is recommended.

5. Using large homogeneous samples, allometry of size is studied in the antennae, legs and palps of C. pulicaris, punctatus and impunctatus, to determine whether any taxonomically important characters show proportional changes in size (p.119). A tendency for isometry is found in the apical segments of the legs and palps in pulicaris and punctatus. In C. impunctatus, the only occurrence of allometry is seen in the apical segment of the antennae (p.130).

6. Allometric gradients for the three appendages are calculated, including the 95% confidence limits. Generally, these limits are rather large, and many apparent deviations from isometry are not significant. This provides a good example of the need to use confidence limits in studies on allometry. The allometric gradients are very similar in pulicaris and punctatus, in contrast to those of impunctatus, suggesting that no general description of allometry can be applied to the species complex.

Prompted by the general lack of absolute diagnostic characters for separating members of the complex, attempts are made to establish new and useful characters. The discovery of minute teeth on the cibarium of C. fagineus separates this species from the remainder of the complex (p.142). Chaetotaxy and various ratios are investigated, but give results of little practical use (p.145). Most species were originally defined by wing pattern and therefore

emphasis is given to alternative methods for coding and analysing these patterns: a mechanical scanning method producing 420 characters (p.151) is compared to a method based on extracting 13 pattern elements (p.163). The pattern element method proves more useful on the following grounds:- it operates on logically acceptable characters; reduces character redundancy; and allows easier and faster coding (p.162). Empirical studies show this method to be taxonomically useful, especially if combined with other characters (p.177).

The principal objective of Section 9 is to produce a numerical classification of the pulicaris complex. Two secondary objectives are: whether a large number of characters are required for a classification and secondly, whether the recognised species are homogeneous (p.183). Only when a reliable classification is produced and the taxa defined, can discrimination between taxa proceed (p.305).The main findings are:-

1. Several methods for reducing the number of variables in a multivariate study are tested by comparing the arrangement of OTU's, based on the subset of variables, to that based on all 72 variables. By far the most successful method uses the loadings from an R mode principal component analysis (p. 244. 266). A method using cluster analysis of a between-character distance matrix gives poor results (p.227). Three subsets are selected by subjective means, and although some of the subsets give results of taxonomic interest, they are not very effective in reproducing the arrangment of the reference classification (p.250). The poorest method for selecting variables used characters from only one body region, thus throwing doubt on the validity of the non-specificity hypothesis (p.186).

2. The remarkable similarity between classifications based on 72, 64 and 20 variables revokes the suggestion that the number of variables should be maximised (to a limit) in numerical studies. The classification based on 20 variables contains all the types of characters used in this study, suggesting that the number of character types is more important than the absolute number of characters themselves (p.268. 302).

3. All the experiments carried out in this section show two important taxonomic conclusions. Firstly, that as the number of specimens in this study is increased, the boundaries between the

taxa become less distinct, and overlap between the species increases (p.268, 300). Secondly, even when powerful techniques such as principal component analysis are used, a concise summary of the data is not possible. The relative magnitude of the principal axes in each multivariate analysis suggests that the overall geometric arrangement of specimens in hyperspace is very similar (p.272), the main difference between the results being the position of individual specimens.

4. In general, the results of the multivariate studies lend little support to the subdivision of the complex, as previously proposed (p.274).

5. Using the results of the multivariate analyses, the following species are recognised: C. delta, fagineus, grisescens, impunctatus, newsteadi, pulicaris, punctatus and a sample from Japan which is given specific status. One taxon, C. lupicaris, appears too heterogeneous to be considered as a valid species (p.276).

Having defined the groups (species), canonical variate analysis is used to discriminate between them. Unfortunately, tests of homogeneity of dispersion matrices are inapplicable because some matrices are singular or have negative determinants (p.281). Percentiles about the means of each species are incorporated into canonical variate diagrams. Specimens can be accurately identified by placing them on these diagrams, using the linear combination of 10 variables which define each axis (p.294).

The results of this study indicate that traditional Linnaean taxonomy, based on disjoint sets, cannot always adequately accomodate morphologically variable species(p.313.317). Furthermore, evidence is drawn from evolutionary studies to suggest that a taxonomic system should accomodate indistinct species as well as smaller units, such as populations, geographical isolates,etc. (p.315).

It is proposed therefore, that the axiom of mutually exclusive sets on which the species category is based, be relaxed for species overlapping in terms of morphological and other character suites. The merits and limitations of a classification system based on multiple discriminant analysis is discussed in relation to the C. pulicaris complex (p.391, 322). Two other methods, non-hierarchical overlapping cluster analysis (p.322) and theory of fuzzy sets (p.323)- in which transition from membership to non-membership of each set is gradual, rather than abrupt - are reviewed and related to present taxonomic requirements.

REFERENCES

AKIBA, K. 1960. Studies on the Leucocytozoon found in the chicken in Japan. II On the transmission of Leucocytozoon caulleryi by Culicoides arakawae.
Jap. J.vet.Sci. 22: 309-317.

AKIBA, K., KITAOKA, S. & YAJIMA, T. 1959. Studies on leucocytozoonosis in Japan. III Transmission of Leucocytozoon caulleryi by Culicoides arakawae.
Jap. J. vet. Sci. 21: 127-128.

ALBU, P., & GEORGESCU, A.D. 1971. [New contributions to the study of the Ceratopogonidae (Diptera) of Roumania].[in Roumanian].
Studii cerc. Biol. (Ser. Zool.). 23: 91-100.

AMADON, D. 1966. The superspecies concept.
Syst. Zool. 15: 245-249.

AMASOVA, I.S. 1957. [Some new or little known Culicoides Latr. (Dipt. Heleidae) from Ussuri Land]. [in Russian].
Ent. Obozr. 36: 233-247.

AREAN, V.M., & FOX, J. 1955. Dermal alterations in severe reactions to the bite of the sand fly Culicoides furens.
Am. J. clin. Path. 25: 1359-1366.

ARNAUD, P. 1956. The Heleid genus Culicoides in Japan, Korea and Ryuku Islands(Insecta: Diptera).
Microentomology 21: 84-207.

ATCHLEY, W.R. 1967. The Culicoides of New Mexico (Diptera: Ceratopogonidae).
Kans. Univ. Sci. Bull. 46: 937.

ATCHLEY, W.R. 1970. A biosystematic study of the subgenus Selfia of Culicoides (Diptera: Ceratopogonidae).
Kans. Univ. Sci. Bull. 49: 181-336.

ATCHLEY, W.R. 1971. Statistical analysis of geographic variation in pupae of three species of Culicoides (Diptera: Ceratopogonidae).
Evolution, Lancaster, Pa. 25: 51-74.

ATCHLEY, W.R. 1973. A quantitative separation of the females of Culicoides (Selfia) denningi, C. (S.) hieroglyphicus and C. (S.) jamesi (Diptera: Ceratopogonidae).
J. med. Ent. Honolulu 10: 629-632.

ATCHLEY, W.R. 1974. A quantitative taxonomic analysis of Leptoconops torrens and L. carteri (Diptera: Ceratopogonidae)
J. med. Ent. Honolulu 11: 467-470.

ATCHLEY, W.R., GASKINS, G.T. & ANDERSON, D. 1976. Statistical properties of ratios. I Empirical results.
Syst. Zool. 25: 137-148.

ATCHLEY, W.R. & JACKSON, R.C. 1968. Techniques and preliminary studies on the chromosomes of Nilobezzia and Bezzia (Diptera: Ceratopogonidae).
Ann. ent. Soc. Am. 61: 1524-1527.

AUTRUM, H. 1963. Anatomy and physiology of sound receptors in invertebrates.
In Busnel, R.G. (ed.) Acoustic behaviour of animals : 412-433. Elsevier, Amsterdam.

AYALA, F.J. 1971. Competition between species: frequency dependence.
Science, N.Y. 171: 820-824.

AYALA, F.J. & POWELL, J.R. 1972. Allozymes as diagnostic characters of sibling species of Drosophila.
Proc. natn. Acad. Sci. U.S.A. 69: 1094-1096.

BAILLY-CHOUMARA, H. & KREMER, M. 1970. Deuxieme contribution à l'étude des Culicoides du Maroc (Diptera: Ceratopogonidae).
Cah. Off. Rech. Sci. Tech. Outre-Mer (Ent. méd.) 8: 383 391.

BARTLETT, M.S. 1949. Fitting a straight line when both variables are subject to error.
Biometrics 5: 207-212.

BEALE, E.M.L., KENDALL, M.G. & MANN, D.W. 1967. The discarding of variables in multivariate analysis.
Biometrika 54: 357 - 366.

BELOKUR, V.M. 1960. On the fauna of bloodsucking Diptera in the Nenets National Region in the North of the Komi ASSR.
Ent. Rev., Wash. 39: 268-271.

BILINSKI, Z. 1968. Uzupelnienia do znajomosci kuczmanów Polski.
Annls Univ. Mariae Curie - Sklodowska (C) 23: 319-325.

BLACKITH, R.E., DAVIES, R.G. & MOY, E.A. 1963. A biometric analysis of development in Dysdercus fasciatus Sign. (Hemiptera: Pyrrhocoridae).
Growth 27: 317 - 334.

BLACKITH, R.E. & REYMENT, R.A. 1971. Multivariate morphometrics.
Academic Press, London: 412pp.

BLACKWELDER, R.E. 1967. Taxonomy: A text and reference book.
Wiley, New York: 698pp.

BOORMAN, J. 1974. Culicoides (Diptera: Ceratopogonidae) from Cyprus.
Cah. Off. Rech. Sci. Tech. Outre-Mer (Ent. med.) 12: 7-13.

BOORMAN, J. 1979. The transmission of animal virus diseases by Culicoides.
Annls Parasit. hum. comp. [in press].

BOORMAN, J. [in press]. Phlebotomidae and Ceratopogonidae as virus vectors.
Int. Congr. Parasitology 4: [in press].

BOORMAN, J. & GIBBS, E.P.J. 1973. Multiplication of the virus of Epizootic Haemorrhagic Disease of Deer in Culicoides species (Diptera: Ceratopogonidae). Arch. Virusforsch. 41: 259-266.

BOORMAN, J. & GOODARD, P. 1970. Observations on the biology of Culicoides impunctatus Goetgh. (Diptera: Ceratopogonidae) in Southern England. Bull. ent. Res. 60: 189-198.

BOORMAN, J. & LANE, R.P. 1979. A new genus and species of Ceratopogonidae (Diptera) closely related to Culicoides from West Africa. J. nat. Hist. [in press].

BORATYNSKI, K.L. 1952. Matsucoccus pini (Homoptera: Coccoidea: Margarodidae) Bionomics and external anatomy with reference to the variability of some taxonomic characters. Trans. R. ent. Soc. Lond. 103: 285-326.

BRAVERMAN, Y., BOORMAN, J. & KREMER, M. 1976. Faunistic list of Culicoides (Diptera: Ceratopogonidae) from Israel. Cah. Off. Rech. Sci. Tech. Outre-Mer (Ent. med.) 14: 179-185.

BRAVERMAN, Y., BOREHAM, P.F.L. & GALUM, R. 1971. The origin of blood meals of female Culicoides pallidipennis trapped in a sheepfold in Israel. J. med. Ent. Honolulu 8: 379-381.

BROWN, K.R. & SHIPP, E. 1977. Wing morphometrics of Australian Luciliini. Aust. J. Zool. 25: 765-777.

BROWN, K.R. & SHIPP, E. 1978. Wing morphometric analysis of Australian Sarcophaginae (Diptera: Sarcophagidae). Syst. Ent. 3: 179-188.

BROWN, V.K. 1969. Aspects of the biology and growth of 3 species of Ectobius (Dictyoptera: Blattidae). PhD thesis, University of London: 397pp.

BROWN, V.K. 1977. Metamorphosis: a morphometric description. Int. J. Insect Morphol. Embryol. 6: 221-223.

BROWN, V.K. & DAVIES, R.G. 1972. Allometric growth in two species of Ectobius (Dictyoptera: Blattidae). J. Zool., Lond. 166: 97-132.

BROWN, W.L. & WILSON, E.O. 1956. Character displacement. Syst. Zool. 5: 49-64.

BUCKLEY, J.C.C. 1934. On the development in Culicoides furens Poey of Filaria (Mansonella) ozzardi Manson. J. Helminth. 12: 99-118.

BUCKLEY, J.C.C. 1938. On Culicoides as a vector of Onchocerca gibsoni (Cleland & Johnston). J. Helminth. 16: 121-158.

BUXTON, P.A. 1938. Studies on the growth of Pediculus. Parasitology 30: 65-84.

CAIN, A.J. 1954. Animal species and their evolution. Hutchinson, London: 190pp.

CAIN, A.J. & HARRISON, G.A. 1958. An analysis of the taxonomist's judgement of affinity. Proc. zool. Soc. Lond. 131: 85-98.

CALLAHAN, P.S. 1975. Insect antennae with special reference to the mechanism of scent detection and the evolution of the sensilla. Int. J. Insect. Morph. Embryol. 4: 381-430.

CALLOT, J., KREMER, M. & BAILLY-CHOUMARA, H. 1968. Note faunistique sur les Culicoides (Dipt. Ceratopogonides) du Maroc. Gynandromorphisme de C. circumscriptus parasite par une Mermis. Bull. Soc. Path. exot. 61: 885-889.

CALLOT, J., KREMER, M. & BRAVERMAN, Y. 1969. Note sur des Culicoides récoltés en Israel (Diptera:Ceratopogonidae). Bull. Soc. Path. exot. 62: 118: 123.

CALLOT, J., KREMER, M. & COLUZZI, M. 1965. Nouvelle contribution á l'étude des Culicoides (Diptera: Ceratopogonidae) d'Italie. Parassitologia 7: 162 171.

CALLOT, J., KREMER, M. & DEDUIT, Y. 1962. Nouvelles espèces et nouvelles localisations de Culicoides. Annls Parasit. hum. comp. 37: 153

CALLOT,J., KREMER, M. & GEISS, J.L. 1972. Iconographie de l'armature cibariale de 22 espèces de Culicoides (Diptera: Ceratopogonidae). Annls Parasit. hum. comp. 47: 759-762.

CALLOT, J., KREMER, M., RAULT, B. & BACH, A. 1966. Contribution á l'étude des Culicoides de l'ouest de la France. Annls Parasit. hum. comp. 41: 513-521.

CALLOT, J., KREMER, M., RIOUX, J.A. & DESCOUS, S. 1967. Culicoides des Pyrenees Orientales description de C. caucoliberensis n.sp. Bull. Soc. zool. Fr. 92: 827-832.

CAMPBELL, J.A. & PELHAM-CLINTON, E.C. 1960. A taxonomic review of the British species of Culicoides Latreille (Diptera: Ceratopogonidae). Proc. R. Soc. Edinb. (B) 68: 181-302.

CARSON, H.L., CLAYTON, F.E. & STALKER, H.D. 1967. Karyotypic stability and speciation in Hawaiian Drosophila. Proc. natn. Acad. Sci. U.S.A. 57: 1280-1285.

CARTER, H.F., INGRAM, A. & MACFIE, J.W.S. 1920. Observations on the Ceratopogonine midges of the Gold Coast with descriptions of new species.
Ann. trop. Med. Parasit. 14: 187-210; 211-274; 308-331.

CARTER, H.F., INGRAM, A. & MACFIE, J.W.S. 1921. Observations on the Ceratopogonine midges of the Gold Coast with descriptions of new species.
Ann. trop. Med. Parasit. 15: 177-212.

CHAN, K.L. & SAUNDERS, L.G. 1965. Forcipomyia (Dacnoforcipomyia) anabaenae a new bloodsucking midge from Singapore described in all stages (Diptera: Ceratopogonidae).
Can. J. Zool. 43: 527-540.

CHANG, P.H. & WANG, C.Y. 1958. [A study of Lasiohelea taiwana Shiraki from Fukien Province] . [in Chinese with English Summary].
Acta microbiol. sin. 6: 140-144.

CHARDROME, M. & PEEL, E. 1951. Recherches sur la répartition de filairs dans la région de Coquilhatville et la transmission de Dipetalonema streptocerca par Culicoides grahami Austen.
Mem. Inst. r. colon. belge Sect. Sci. nat. med. 19: 3-83.

CHRISTOPHERS, S.R., SHORTT, M.E. & BARRAUD, P.J. 1925. Culicoides and Kala-azar in Assam.
Indian J. med. Res. 8: 175-176.

CHU, W., AXTELL, R.C. & KLINE, D.L. 1975. Antennal and palpal sensilla of the sand fly Culicoides furens (Poey) (Diptera: Ceratopogonidae).
Int. J. Insect morph. Embryol. 4 (2): 131-149.

CHUNG-WAN, L., ER-CHENG, T., LIEN-LAI, T. & YU-KUAN, L. 1964. [Observations on the breeding habits of Lasiohelea taiwana Shiraki]. [in Chinese with English Summary].
Acta ent. sin. 13: 760.

CLARK, L.B. & HERSH, A.H. 1939. A study of relative growth in Notonecta undulata.
Growth 3: 347-372.

CLASTRIER, J. 1957. Notes sur les Cératopogonidés. II Quelques Culicoides d'Algerie à ailes tachetées.
Archs Inst. Pasteur Alger. 35: 404-444.

CLASTRIER, J. 1959. Notes sur les Cératopogonidés. VI Cératopogonidés d'Afrique occidentale Francaise.
Archs Inst. Pasteur Alger. 37: 167-197.

CLASTRIER, J. & WIRTH, W.W. 1978. The Leptoconops kerteszi complex in North America (Diptera: Ceratopogonidae).
Tech. Bull. U.S. Dep. Agric. 1573: 58pp.

CLIFFORD, H.T. & STEPHENSON, W. 1975. An introduction to numerical classification.
Academic Press, London: 229pp.

COLUZZI, M. & KREMER, M. 1964. Contributo alla conoscenza dei culicoidi Italiani (Diptera: Ceratapogonidae). Parassitologia 6: 63-69.

COOLEY, W.W. & LOHNES, P.R. 1962. Multivariate procedures for the behavioural sciences. Wiley, New York - London: 211pp.

CORNET, M. 1974. Caracteres morphologique utilisés pour l'identification des Culicoides (Diptera: Ceratopogonidae). Cah. Off. Rech. Sci. Tech. Outre-Mer (Ent. med.) 22: 221-229.

CROWSON, R.A. 1970. Classification and Biology. Heinemann, London: 350pp.

DAVIDSON, G. & WHITE, G.B. 1972. The crossing characteristics of a new, sixth species in the Anopheles gambiae complex. Trans. R. Soc. trop. Med. Hyg. 66:531.

DAVIDSON, G. PATTERSON, H.E., COLUZZI, M., MASON, G.F. & MICKS, D.W. 1967. The Anopheles gambiae complex. In Wright, J.W. & Pal, R. (eds) Genetics of insect vectors of disease 211-250. Elsevier, Amsterdam.

DAVIES, F.G. & WALKER, A.R. 1974. The isolation of ephemeral fever virus from cattle and Culicoides midges in Kenya. Vet. Rec. 95: 63-64.

DAVIES, R.G. 1971. Computer programming in quantitative biology. Academic Press, London - New York: 492pp.

DAVIES, R.G. & BORATYNSKI, K.L. 1979. Character selection in relation to the numerical taxonomy of some male Diaspididae (Homoptera: Coccoidea). Biol. J. Linn. Soc. [in press].

DAVIES, R.G. & BROWN, V.K. 1972. A multivariate analysis of postembryonic growth in two species of Ectobius (Dictyoptera: Blattidae). J. Zool., Lond. 168: 51: 79.

DAVIS, J.C. 1972. Statistics and data analysis in Geology. Wiley, New York: 550pp.

DOBZHANSKY, T. 1937. Genetics and the origin of species. Columbia Universtity Press, New York: 364pp.

DOBZHANSKY, T. 1951. Experiments on sexual isolation in Drosophila. X. Reproductive isolation between Drosophila pseudoobscura and D. persimilis under natural and laboratory conditions. Proc. natn. Acad. Sci. U.S.A. 37: 792-796.

DOBZHANSKY, T. 1970. Genetics of the evolutionary process. Columbia Universtity Press, New York - London: 505pp.

DOBZHANSKY, T. 1972. Species of Drosophila - new excitement in an old field. Science, N.Y. 177: 664-669.

DODSON, P. 1978. On the use of ratios in growth studies. Syst. Zool. 27: 62-67.

DOHERTY, R.L., CARLEY, J.G., STANDFAST, H.A., DYCE, A.L. & SNOWDON, W.A. 1972. Virus strains from Arthropods during an epizootic of bovine ephemeral fever in Queensland. Aust. vet. J. 48: 81-86.

DOHERTY, R.L., CARLEY, J.G., STANDFAST, H.A., DYCE, A.L., KAY, B.H. & SNOWDON, W.A. 1973. Isolation of arboviruses from mosquitoes, biting midges , sand flies and vertebrates collected in Queensland 1969-1970. Trans. R. Soc. trop. Med. Hyg. 67: 536-543.

DOWNES, J.A. 1955. Observations onthe swarming flight and mating of Culicoides (Diptera: Ceratopogonidae). Trans. R. ent. Soc. Lond. 106: 213- 236.

DOWNES, J.A. 1958. The feeding habits of biting flies and their significance in classification. A. Rev. Ent. 3: 249-266.

DOWNES, J.A. 1969.. The swarming and mating flight of Diptera. A. Rev. Ent. 14: 271-298.

DOWNES, J.A. 1970. The ecology of bloodsucking Diptera: an evolutionary perspective. In Fallis, A.M. (ed.) Ecology and physiology of parasites, a symposium. Toronto University Press, Toronto: 232-258.

DOWNES, J.A. 1973. Biting flies: the need for a new systematics. In Hudson, A. (ed.) Biting fly control and environmental quality. Proceedings of a symposium held at the University of Alberta, Edmonton, May 16-18, 1972: 115-123. Defence Research Board, Ottawa.

DOWNES, J.A. 1977a. Evolution of feeding habits in Ceratopogonidae. Mosquito News 37: 279-280.

DOWNES, J.A. 1977b. Leptoconops - Leptoconopidae? Mosquito News 37: 277.

DOWNES, J.A. & KETTLE, D. 1952. Descriptions of three new species of Culicoides Latreille (Diptera: Ceratopogonidae) new to science, together with notes on, and a revised key to the British species of the pulicaris and obsoletus groups. Proc. R. ent. Soc. Lond. 21: 61-78.

DUKE, B.O.L. 1954. The uptake of the microfilariae of Acanthocheilonema streptocerca by Culicoides grahamii and their subsequent development. Ann. trop. Med. Parasit. 48: 416-420.

DUKE, B.O.L. 1956. The intake of microfilariae of Acanthocheilonema perstans by Culicoides grahamii and C. inornatipennis and their subsequent development. Ann. trop. Med. Parasit. 50: 32-38.

DUKE, B.O.L. 1958. The intake of the microfilariae of Acanthocheilonema streptocerca by Culicoides milnei with some observations on the potentialities of the fly as a vector.
Ann. trop. Med. Parasit. 52: 123-128.

DUPRAW, E.J. 1964. Non-Linnean taxonomy.
Nature, Lond. 202: 849-852.

DUPRAW, E.J. 1965a. Non-Linnean taxonomy and the systematics of honey bees.
Syst. Zool. 14: 1-24.

DUPRAW, E.J. 1965b. The recognition and handling of honey bee specimens in non-Linnean taxonomy.
J. apic. Res. 4:71-84.

DZHAFAROV, S.M. 1964. [Bloodsucking midges (Diptera: Heleidae) of Transcaucasia]. [in Russian].
Trudy Inst. Zool., Baku : 414pp.

EDWARDS, F.W. 1926. On the British biting midges (Diptera: Ceratopogonidae).
Trans. R. ent. Soc. Lond. 74: 389-426.

EDWARDS, F.W. 1939. Genitalia of the British Culicoides with notes on synonymy.
In Edwards, F.W., Oldroyd, H. & Smart, J. British Bloodsucking Flies: 129-148. British Museum (Natural History), London.

EIMER, G.H.T. 1897. Orthogenesis der Schmetterlinge.
Leipzig [not seen].

EWING, A.W. 1978. The antennae of Drosophila as a 'love song' receptor.
Physiol. Ent. 3: 33-36.

FALLIS, A.M. & BENNETT, G.F. 1960. Description of Haemoproteus canachites n.sp. (Sporozoa: Haemoproteidae) and sporogony in Culicoides (Diptera: Ceratopogonidae).
Can. J. Zool. 38: 455-464.

FALLIS, A.M. & BENNETT, G.F. 1961a. Ceratopogonidae as intermediate hosts for Haemoproteus and other parasites.
Mosquito News 21: 21-28.

FALLIS, A.M. & BENNETT, G.F. 1961b. Sporogony of Leucocytozoon and Haemoproteus in Simuliids and Ceratopogonids and a revised classification of the Haemosporidiida.
Can. J. Zool. 39: 215-228.

FALLIS, A.M. & WOODS, D.M. 1957. Biting midges (Diptera: Ceratopogonidae) as intermediate hosts for Haemoproteus of ducks.
Can. J. Zool. 35: 425-435.

FARRIS, J.S. 1969. On the cophenetic correlation coefficient.
Syst. Zool. 18: 279-285.

FARRIS, J.S. 1971. The hypothesis of non-specificity and taxonomic congruence.
A. Rev. Ecol. Syst. 2: 277-302.

FITTKAU, E.F. & REISS, F. 1978. Family Chironomidae.
In Limnofauna Europea: 404-440. Gustav Fischer Verlag, Stuttgart.

FLETCHER, T.B. 1924. Cited in Christophers, S.R., Shortt, M.E. & Barraud, P.J. 1925.

FOSTER. N.M., JONES, R.H. & McRORY, B.R. 1963. Preliminary investigations on insect transmission of bluetongue virus in sheep.
Am. J. vet. Res. 24: 1195-1200.

FOX, I. 1955. A catalogue of the bloodsucking midges of the Americas (Culicoides, Leptoconops and Lasiohelea) with keys to the subgenera and Nearctic species, a geographic index and bibliography.
J. Agric. Univ. P. Rico 39: 214-285.

FREDEEN, F.J.H. 1969. Culicoides (Selfia) denningi, a unique river-breeding species.
Can. Ent. 101: 539-544.

GAD, A.M. 1951. The head capsule and mouthparts of the Ceratopogonidae.
Bull. Soc. Fouad I. Ent. 44: 17-75.

GARNHAM, P.C.C. 1966. Malaria parasites and other Haemosporidia.
Blackwell, Oxford: 1114pp.

GARNHAM, P.C.C., HEISCH, R.B. & MINTER, D.M. 1961. The vector of Hepatocystis (Plasmodium) kochi, the successful conclusion in many parts of tropical Africa.
Trans. R. Soc. trop. Med. Hyg. 55: 487-502.

GEORGESCU, D.A. 1972. Nouvelle éspèces de Ceratopogonidae.
Revue roum. Biol. (ser. Zool.) 17: 15-21.

GHISELIN, M.T. 1974. A radical solution to the species problem.
Syst. Zool. 23: 536-544.

GIBSON, C.L. & ASCOLI, W.F. 1952. The relation of Culicoides (Diptera: Heleidae) to the transmission of Onchocerca volvulus.
J. Parasit. 38: 315-320.

GLUKHOVA, V.M. 1962. [Biting midges (Diptera: Heleidae) of Karelia]. [in Russian].
Trudy zool. Inst. Leningr. 31: 197-248.

GLUKHOVA, V.M. 1977. [The subgeneric classification of the genus Culicoides Latreille 1809 (Diptera: Ceratopogonidae) including morphological characters of the larvae]. [in Russian].
Parazit. Sb. 27: 112-118. [translation by D.S. Kettle, distributed through Ceratopogonidae information exchange No. 22].

GLUKHOVA, V.M. & BERZINA, A.N. 1963. [The blood-sucking midges (Diptera: Heleidae) of the middle course of the Kalcha River, Krass on Jarsk Region]. [in Russian]. Ent. Obozr. 40: 852-832.

GOETGHEBUER, M. 1920. Ceratopogonidae de la Belgique. Mem. Mus. r. Hist. nat. Belg. 8: 1-116.

GOULD, S.J. 1966. Allometry and size in ontogeny and phylogeny. Biol. Rev. 41: 587-640.

GOWER, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-328.

GOWER, J.C. 1971a. A general distance coefficient of similarity and some of its properties. Biometrics 27: 857-871.

GOWER, J.C. 1971b. Statistical methods of comparing different multivariate analysis of the same data. In Hodson, F.R., Kendall, D.G. & Tautu, P. (eds) Mathematics in the archeological and historical sciences: 138-149. Edinburgh University Press, Edinburgh.

GRAHAM, M.W.R. 1950. Postural habits and colour pattern evolution in Lepidoptera. Trans. Soc. Br. Ent. 10: 217-232.

GREINIER, C. & BENNETT, G.F. 1977. Culicoides as vectors of Haemosporidia. Mosquito News 37: 286.

GRIFFITHS, G.C.D. 1972. The phylogenetic classification of the Diptera: Cyclorrhapha, with special reference to the structure of the male post abdomen. Series Ent. 8: 1-340.

GUTSEVICH, A.V. 1952. [Contribution to the fauna of sandflies of the genus Culicoides of the forest zone (Diptera: Heleidae)]. [in Russian]. Parazit. Sb. 14: 75-94.

GUTSEVICH, A.V. 1964. Biting midges of the genus Culicoides (Diptera: Heleidae) of the Ukranian Carpathians (Transcarpathian Province). Ent. Rev., Wash. 43: 310-314.

GUTSEVICH, A.V. 1973. [The bloodsucking midges (Ceratopogonidae)]. [in Russian]. Fauna SSSR 3 (5): 1-270.

GUTSEVICH, A.V. & VIGOVSKII, A.I. 1960. [The isolation of neurotropic virus from Culicoides]. [in Russian , with English summary]. Vop. Virus. 5:657.

GYLLENBERG, H.G. & NIEMELA, T.K. 1975. New approaches to automatic identification of micro-organisms. In Pankhurst, R.J. (ed.) Biological identification with computers : 137-152. Academic Press, London.

HAIR, J.A. & TURNER, E.C. 1968. Preliminary host preference studies on Virginia Culicoides (Diptera: Ceratopogonidae). Mosquito News 28: 103-107.

HARTLEY, W.J., WANNER, R.A., DELLA-PORTA, A.J. & SNOWDON, W.A. 1975. Serological evidence for the association of Akabane virus with epizootic bovine congenital arthrogryposis and hydrancephaly syndromes in N.S.W. Aust. vet. J. 5: 103-104.

HAVELKA, P. 1976. Limnologische and systematische studien an Ceratopogoniden. Beitr. Ent. 26: 211-305.

HAVELKA, P. 1979. Die Ceratopogoniden forschung auf der Iberischen Halbinsel. Annls Parasit. hum. comp. (in press).

HENNIG, W. 1968. Phylogenetic Systematics. University of Illinois Press: 263pp.

HENSLEIGH, D.A. & ATCHLEY, W.R. 1977. Morphometric variability in natural and laboratory populations of Culicoides variipennis (Diptera: Ceratopogonidae). J. med. Ent. Honolulu 14: 379-386.

HEUVAL, M.J. van den. 1963. The effect of rearing temperature on the wing length, thorax length, leg length and ovariole number of the adult mosquito Aedes aegypti. Trans. R. ent. Soc. Lond. 115 (7): 197-216.

HILL, M.A. 1947. The life cycle and habits of Culicoides impunctatus Goetghebuer and Culicoides obsoletus Meigen together with some observations on the life cycle of Culicoides odibilis Austen, Culicoides pallidicornis Kieffer, Culicoides cubitalis Edwards and Culicoides chiopterus Meigen. Ann. trop. Med. Parasit. 41: 55-115.

HILLS, M. 1978. On ratios - a response to Atchley, Gaskins and Anderson. Syst. Zool. 27: 61-62.

HOPKINS, C.A. 1952. Notes on the biology of certain Culicoides studied in the British Cameroons, West Africa, together with observations on their possible rôle as vectors of Acanthocheilonema perstans. Ann. trop. Med. Parasit. 46: 165-172.

HOPKINS, C.A. & NICHOLAS W.L. 1952. Culicoides austeni, the vector of Acanthocheilonema perstans. Ann. trop. Med. Parasit. 46: 276-283.

HOWELL, C.J. 1970. Leptoconops mange. Jl S. Afr. vet. med. Ass. 41: 71-73.

HULL, D.L. 1965. The effect of essentialism on taxonomy - two thousand years of stasis. Br. J. Phil. Sci. 15: 314-326 and 16: 1-18.

HUQ, M. 1961. African Horse Sickness.
Vet. Rec. 73: 123.

HUXLEY, J.S. 1924. Constant differential growth ratios and their significance.
Nature, Lond. 114: 895-896.

HUXLEY, J.S. 1932. Problems of relative growth.
Methuen, London: 276pp.

IMMS, A.D. 1944. On the constitution of the maxillae and labium in Mecoptera and Diptera. Q. Jl microsc. Sci. 85: 73-96.

IVANOV, K.S. & GLUKHOVA, V.M. 1967. New data on the fauna of bloodsucking midges (Diptera: Ceratopogonidae) from the coastal zone of the Primorje Territory.
Rev. Ent., Wash. 46: 808-813.

JAMNBACK, H. 1965. The Culicoides of New York State.
Bull. N.Y. St. Mus. Sci. Serv. 399: 1-154.

JARDINE, N. & SIBSON, R. 1968. The construction of hierarchic and non-hierarchic classifications.
Comput. J. 11: 177-184.

JARDINE, N. & SIBSON, R. 1971. Mathematical taxonomy.
Wiley, London: 286pp.

JEFFERS, J.N.R. 1967. The study of variation in taxonomic research.
Statistician 17: 29-43.

JENKINS, D.W. 1948. Ecological observations on the blackflies and punkies of Central Alaska.
Mosquito News 8: 148-154.

JOBLING, B. 1928. The structure of the head and mouthparts of Culicoides pulicaris.
Bull. ent. Res. 18: 211-236.

JOBLING, B. 1976. On the fascicle of bloodsucking Diptera.
J. nat. Hist. 10: 457-461.

JOHNSTON, C.G. 1939. Taxonomic characters, variability and relative growth in Cimex.
Trans. R. ent. Soc. Lond. 89: 543-568.

JOLICOEUR, P. 1963. The multivariate generalisation of the allometry equation.
Biometrics 19: 497-499.

JOLLIFFE, I.T. 1972. Discarding variables in principal component analysis. I Artificial data.
Appl. Statist. 21: 160-173.

JOLLIFFE, I.T. 1973. Discarding variables in a principal component analysis. II Real data.
Appl. Statist. 22: 21-31.

JONES, R.H. 1960. Mass-production methods for the colonisation of Culicoides variipennis sonorensis.
J. econ. Ent. 53: 731-735.

JONES, R.H., POTTER, H.W. & RHODES, H.A. 1972. Ceratopogonidae attacking horses in South Texas during the 1971 V.E.E. epidemic.
Mosquito News 32: 507-509.

KARSTAD, L.H., FLETCHER, O.K., SPALATIN, J., ROBERTS, R. & HANSON, R.P. 1957. Eastern Equine encephalomyelitis isolated from three species of Diptera from Georgia.
Science, N.Y. 125: 395-396.

KAUFMANN, A. 1975. Introduction to the theory of fuzzy subsets Vol. 1 Fundamental theoretical elements.
Academic Press, New York: 416pp.

KELLOG, F.E. 1970. Water vapor and carbon dioxide receptors in Aedes aegypti (L).
J. Insect. Physiol. 16: 99-108.

KENDALL, M. 1975. Multivariate analysis.
Griffin & Co, London: 210pp.

KERMACK, K.A. & HALDANE, J.B.S. 1950. Organic correlation and allometry.
Biometrika 37: 30-41.

KETTLE, D.S. 1949. An attempt to control Culicoides impunctatus Goetghebuer in Scotland by barrier-spraying.
Ann. trop. Med. Parasit. 43: 284-296.

KETTLE, D.S. 1950. The seasonal distribution of Culicoides impunctatus Goetghebuer (Diptera: Ceratopogonidae) with a discussion on the possibility that it may be composed of two or more biological races.
Trans. R. ent. Soc. Lond. 101: 125-145.

KETTLE, D.S. 1951a. Some factors affecting the population density and flight range of insects.
Proc. R. ent. Soc. Lond. (A) 26: 59-63.

KETTLE, D.S. 1951b. The spatial distribution of Culicoides impunctatus Goetghebuer under woodland and moorland conditions and its flight range through woodland.
Bull. ent. Res. 42: 239-292.

KETTLE, D.S. 1952. Midge (Culicoides) control in Scotland.
Int. Congr. Ent. 9: 935-938.

KETTLE, D.S. 1956. Rainfall and larval density of Culicoides impunctatus Goetghebuer (Diptera: Ceratopogonidae).
Proc. R. ent. Soc. Lond. (A) 31: 173-177.

KETTLE, D.S. 1957. Preliminary observations on weather conditions and the activity of biting flies.
Proc. R. ent. Soc. Lond. (A) 32: 13-20.

KETTLE, D.S. 1960. The flight of Culicoides impunctatus Goetghebuer (Diptera: Ceratopogonidae) over moorland and its bearing on midge control.
Bull. ent. Res. 51: 461-489.

KETTLE, D.S. 1962. The bionomics and control of Culicoides and Leptoconops (Diptera: Ceratopogonidae).
A. Rev. Ent. 7: 401-418.

KETTLE, D.S. 1965. Biting ceratopogonids as vectors of human and animal diseases.
Acta trop. 22: 356-362.

KETTLE, D.S. & LAWSON, J.W.H. 1952. The early stages of the British biting midges Culicoides Latreille (Diptera: Ceratopogonidae) and allied genera.
Bull. ent. Res. 43: 421-467.

KHALAF, K. 1961. More Culicoides from Iraq (Diptera: Heleidae).
Beitr. Ent. 11: 450-470.

KHAMALA, C.P.M. & KETTLE, D.S. 1971. The Culicoides Latreille (Diptera: Ceratopogonidae) of East Africa.
Trans. R. ent. Soc. Lond. 123: 1-95.

KIDWELL, J.F. & CHASE, H.B. 1967. Fitting the allometric equation - a comparison of ten methods by computer simulation.
Growth 31: 165-179.

KIMURA R. 1959. Studies on dermatitis due to Leptoconops nipponensis Tok. and the abatement of it.
J. Yonago med. Assoc. 10: 904-929.

KITZMILLER, J.B., FRIZZI, G. & BAKER, R.H. 1967. Evaluation and speciation within the maculipennis complex of the genus Anopheles.
In Wright, J.W. & Pal, R. (eds) Genetics of Insect vectors of disease: 151-210. Elsevier, Amsterdam.

KNOZ, J. 1977. Přispěvek k poznáni rodu Culicoides (Ceratopogonidae: Diptera) na územi cech a moravy.
Scr. Fac. Sci. nat. Univ. purkyn. brun. 8: 55-72.

KORNURBAEV, E.O. 1965. [Biting midges (Diptera: Heleidae) of the Issyk-kul Depression in Kirghizia]. [in Russian].
Ent. Obozr. 44: 32-40.

KREMER, M. 1965. Contribution a l'étude du genre Culicoides Latreille.
Encycl. ent. (A) 39: 1-299.

KREMER, M. & CALLOT, J. 1961. Sur quelques Culicoides (Diptera: Ceratopogonidae) du Haut-Valais.
Acta trop. 18: 255-256.

KREMER, M. & DELECOLLE, J.C. 1974. Variabilité des caractères morphologiques des Culicoides.
Annls Parasit. hum. comp. 49: 617-619.

KREMER, M., DOBY, J.M. & SKIERSKA, B. 1965. Capture de Culicoides (Diptères Nématoceres Ceratopogonidés) dans le Nord et l'Est de la Pologne.
Acta parasit. pol. 13: 109-116..

KREMER, M., GEISS, J.C. & HOMMEL, M. 1975. Localisations nouvelles de Culicoides (Diptères: Ceratopogonidés).
Annls Parasit. hum. comp. 50: 245-246.

KREMER,M., HOMMEL, M. & BAILLY-CHOUMARA, H. 1971. Troisieme contribution a l'étude faunistique des Culicoides du Maroc.
Annls Parasit. hum. comp. 46: 661-670.

KREMER, M., LEBERRE, G. & BEAUCOURNU, F. 1971. Notes sur les Culicoides (Diptera: Ceratopogonidae) de Corse. Description de C. corsicus n.sp.
Annls Parasit. hum. comp. 46: 653-660.

KREMER, M., REBHOLTZ, C. & DELECOLLE, J.C. 1975. Étude des types de Culicoides (Diptera: Ceratopogonidae) de Goetghebuer et des autres Ceratopogonidae déposés au Musee de Tervuren.
Revue Zool. Bot. afr. 89: 769 -820.

LANE, R.P. 1975. An investigation into blowfly (Diptera: Calliphoridae) succession on corpses.
J. nat. Hist. 9: 581-588.

LANE, R.P. 1978. The Diptera (two winged flies) of Lundy Island.
Rep. Lundy Fld Soc. [28] : 15-31.

LANE, R.P. 1979. The micromorphology of the humeral pits in Culicoides and their possible function.
Annls Parasit. hum. comp. [in press].

LAVEN, H. 1967. Speciation and evolution in Culex pipiens.
In Wright, J.W. & Pal, R. Genetics of insect vectors of disease: 251-275. Elsevier, Amsterdam.

LAWRENCE, P.A. 1970. Polarity and patterns in the postembryonic development of insects.
In Beament, J.W.L., Treherne, J.E. & Wigglesworth, V.B. (eds) Advances in insect physiology 7: 197-267. Academic Press, London.

LAWRENCE, P.A. & MORATA, G. 1976. The compartment hypothesis.
Symp. R. ent. Soc. Lond. 8: 132-149.

LEE, D.J., REYE, E.J. & DYCE, A.L. 1962. 'Sandflies' as possible vectors of disease in domesticated animals in Australia.
Proc. Linn. Soc. N.S.W. 77: 364-376.

LEE, R.M. 1974. Structure and function of the fascicular stylets, and the labral and cibarial sense organs of male and female Aedes aegypti (L) (Diptera: Culicidae).
Quaest. ent. 10: 187-215.

LEE, R.M. & DAVIES, D.M. 1978. Cibarial sensilla of Toxorhynchites mosquitoes (Diptera: Culicidae).
Int. J. Insect Morph. Embryol. 7: 189-194.

LEGAY, J.M. 1977. Allometry amd systematics. Insect egg form.
J. nat. Hist. 11: 493-499.

LENTZ, F.R. 1934. Die metamorphose der Heleidae.
Fliegen palaearkt. Reg. 3: 95-128.

LEON, N. L'armature bucale et la pompe salivoire des Culicoides pulicaris. 1924.
Bull. Sect. scient. Aead. roum. 9: 17-23.

LEWIS, D.J. 1973. Phlebotomidae and Psychodidae (Sand flies and moth flies).
In Smith, K.G.V. (ed.) Insects and other arthropods of of medical importance : 155-179. British Museum (Natural History), London.

LEWIS, D.J. 1974. The biology of Phlebotomidae in relation to leishmaniasis.
A. Rev. Ent. 19: 363-384.

LEWIS, D.J. 1975. Functional morphology of the mouthparts in New World phlebotomine sandflies (Diptera: Psychodidae).
Trans. R. ent. Soc. Lond. 126: 497-532.

LIDICKER, W.Z. 1973. A phenetic analysis of some New Guinea rodents.
Syst. Zool. 22: 36-45.

LINDEBERG, B. 1971. The study of populations of Tanytarsus gracilentus (Diptera: Chironomidae).
Limnologica, Berlin 8: 49-50.

LINLEY, J.R. 1968a. Studies on the larval biology of Leptoconops bequarti (Kief) (Diptera: Ceratopogonidae).
Bull. ent. Res. 58: 1-24.

LINLEY, J.R. 1968b. Autogeny and polymorphism for wing length in Leptoconops bequarti (Kief.) (Diptera: Ceratopogonidae).
J. med. Ent. Honolulu 5: 53-66.

LINLEY, J.R. 1969. Studies on larval development in Culicoides furens Poey (Diptera: Ceratopogonidae).
Ann. ent. Soc. Amer. 62: 702-711.

LINLEY, J.R. 1976. Biting midges of mangrove swamps and salt marshes (Diptera: Ceratopogonidae).
In Cheng, L. (ed.) Marine insects : 335-376. North Holland Publishing Co., Amsterdam.

LINLEY, J.R. & ADAMS, G.M. 1972. A study of the mating behaviour of Culicoides melleus (Coquillett) (Diptera: Ceratopogonidae).
Trans. R. ent. Soc. Lond. 124: 81-121.

LINLEY, J.R. & CHENG, L. 1974. The grooming organs of Culicoides (Diptera: Ceratopogonidae).
Mosquito News 34: 204-206.

LINLEY, J.R. & DAVIES, J.B. 1971. Sandflies and tourism in Florida, and the Bahamas and Caribbean area. J. econ. Ent. 64: 264-278.

LINLEY, J.R., EVANS, H.T. & EVANS F.D.S. 1970. A quantitative study of autogeny in a naturally ocurring population of Culicoides furens (Diptera: Ceratopogonidae). J. anim. Ecol. 39: 169-183.

LINLEY, J.R. & HINDS, M.J. 1976. Seasonal changes in size, female fecundity and male potency in Culicoides melleus (Diptera: Ceratopogonidae). J. med. Ent. Honolulu 13: 151-156.

LLOYD, L.I. & GOLIGHTLY, W.H. 1939. Insect size and temperature. Nature, Lond. 144: 155.

LOCKE, M. 1967. The development of patterns in the integument of insects. Adv. Morphogen. 6: 33-88.

MACFIE, J.W.S. 1937. Ceratopogonidae (Diptera) from Ethiopia and British Somaliland. Proc. R. ent. Soc. Lond. 6: 73-79.

MATHER, K. 1943. Polygenic inheritance and natural selection. Biol. Rev. 18: 32-64.

MATHER, K. & JINKS, J.L. 1971. Biometrical genetics. 2nd edition, Chapman & Hall, London: 321pp.

MATSUDA, R. 1961. Studies of relative growth in Gerridae (Hemiptera: Heteroptera). Ann. ent. Soc. Am. 54: 578-598.

MATSUDA, R. 1963. A study of relative growth in two strains of Pycnoscelus surinamensis (Linnaeus) (Panchloridae: Blattaria). Growth 26: 129-135.

MATSUDA, R. 1965. Morphology and evolution of the insect head. Mem. Am. ent. Inst. 4: 1-335.

MAYER, K. 1934a. Die metamorphose der Ceratopogonidae (Diptera). Arch. Naturgesch. (N.F.) 3: 205-288.

MAYER, K. 1934b. Die nahrung der Ceratopogoniden larven (Diptera). Arch. Hydrobiol. 27: 564-570.

MAYR, E. 1942. Systematics and the origin of species. Columbia University Press, New York: 334pp.

MAYR, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Mass.: 797pp.

MAYR, E. 1969. Principles of systematic zoology. McGraw-Hill, New York: 428pp.

MAYR, E. 1970. Populations, species and evolution. Belknap Press, Cambridge, Mass.: 453pp.

McGREEVY, P.B., BRYAN, J.H., OOTHUMAN, P. & KOLSTRUP, N. 1978. The lethal effects of the cibarial and pharyngeal armatures of mosquitoes on microfilariae. Trans. R. Soc. trop. Med. Hyg. 72: 361-368.

McGUIRE, J.U. & WIRTH, W.W. 1958. Discriminant function in taxonomic research. Int. Congr. Ent. 10: 387-393.

McIVER, S.B. 1971. Comparative studies on the sense organs on the antennae and maxillary palps of selected male culicine mosquitoes. Can. J. Zool. 49: 235-359.

McMAHON, T.A. 1973. Shape and size in Biology. Science, N.Y. 179: 1201-1204.

McMAHON, T.A. 1977. Allometry: 48-57. McGraw-Hill 1977 Yearbook of Science and Technology, N.Y.

MEGAHEAD, M.M. 1956. Anatomy and histology of the alimentary tract of the female of the biting midge Culicoides nubeculosus (Meigen) (Diptera: Heleidae = Ceratopogonidae). Parasitology 46: 22-47.

MELLOR, P.S., BOORMAN, J. & LOKE, R. 1974. The multiplication of main drain virus in two species of Culicoides. Arch. Virusforsch. 46: 105-110.

MELTZER, B., SEARLE, N.H. & BROWN, R. 1967. Numerical specification of biological form. Nature, Lond. 216: 32-36.

MERCER, K.L. & McIVER, S.B. Sensilla on the palps of black flies (Diptera: Simuliidae). J. med. Ent. Honolulu 10: 236-239.

MESGHALI, A. 1963. Heleidae (Diptera) of Iran. Bull. Soc. Path. exot. 56: 1063-1069.

MIRZAYEVA, A.G. 1963. On the bloodsucking midges of West Siberia. Ent. Rev.,Wash. 43: 108-110.

MOORHOUSE, D.E. 1978. [Bovine onchocerciasis in Queensland]. Ceratopogonidae information exchange June 1978 (21): 2-3.

MORII, T. & KITAOKA, S. 1968a. Relationship between the course of gametocyte appearance of Akiba caulleryi in the chicken and sporozoite production in Culicoides arakawae. Natn. Inst. Anim. Hlth Q., Tokyo 8: 204-209.

MORII, T. & KITAOKA, S. 1968b. Influence of temperature on the sporogony of Akiba caulleryi in three Culicoides species. Natn. Inst. Anim. Hlth Q., Tokyo 8: 210-216.

MORII, T., KITAOKA, S. & AKIBA, K. 1965. Some investigations on the sporogony of Leucocytozoon caulleryi in laboratory reared biting midges of four Culicoides species. Natn. Inst. Anim. Hlth Q., Tokyo. 5: 109-110.

MOSS, W.W. & POWER, D.M. 1975. Semi automatic data recording. Syst. Zool. 24: 199-208.

MUNRO, H.K. 1947. African Trypetidae (Diptera). Mem. ent. Soc. sth Afr. 1: 1-284.

MURADOV, S. M. 1965. Biting midges (Diptera: Heleidae) of Turkmenia - fauna and ecology. Ent. Rev., Wash. 44: 70-74.

NATHAM, M.8. 1978. Culicoides phlebotomus, a vector of Mansonella ozzardi in coastal north Trinidad, West Indies. Trans. R. Soc. trop. Med. Hyg. 72: 436-437.

NIELSEN, B.O. 1963. The biting midges of Lyngby aamose (Culicoides: Ceratopogonidae). Natura jutl. 10: 1-45.

NIELSEN, B.O. 1964. Studies on the Danish biting midges, Culicoides Latreille (Diptera: Ceratopogonidae). Ent. Meddr. 32: 261-276.

NIELSEN, B.O. 1971. Some observations on biting midges (Diptera: Ceratopogonidae) attacking grazing cattle in Denmark. Entomologica scand. 2: 94-98.

NIJHOUT, H.F. 1977. Control of antennal hair erection in male mosquitoes. Biol. Bull. 153: 591-603.

OLDROYD, H. 1952. The horse flies (Diptera: Tabanidae) of the Ethiopian region. I Haematopota and Hippocentrum. British Museum (Natural History), London: 211pp.

OLDROYD, H. 1966. The future of taxonomic entomology. Syst. Zool. 15: 253-260.

OLDROYD, H. 1977. The suborders of Diptera. Proc. ent. Soc. Wash. 79: 3-10.

ONYIAH, J.A. 1971. Studies on the biology and ecology of Ceratopogonidae midges in southern England. PhD Thesis. University of London: 163pp.

ORSZÁGH, I. 1977. Die Gattung Culicoides Latr. 1809 (Diptera: Ceratopogonidae) in Slowakei. Acta Fac. Rerum. nat. Univ. comen., Bratisl. Zool. 21: 1-89.

ORTIZ, I. 1952. Note on the presence in Venezuela of a blood-sucking ceratopogonid Lasiohelea stylifer (Lutz). Revta Sanid. Asist. soc., Caracus 17: 241-244.

PANKHURST, R. (ed.). 1975. Biological identification with computers. Academic Press, London: 366pp.

PARKER, A.H. 1949. Observations on the seasonal and daily incidence of certain biting midges (Culicoides Latreille-Diptera: Ceratopogonidae) in Scotland.
Trans. R. ent. Soc. Lond. 100: 179-190.

PARKER, A.H. 1950. Studies on the eggs of certain biting midges (Culicoides Latreille)occurring in Scotland.
Proc. R. ent. Soc. Lond. (A) 25: 43-52.

PATERSON, H.E. 1978. Thoughts on species and speciation.
Lecture given to the Linnean Society of London, January 1978.

PETERSEN, B. 1952. Studies on geographic variation of allometry in some European Lepidoptera.
Zool. Bidr. Upps. 29: 1-38.

PLOWRIGHT, R.C. & STEPHEN, W.P. 1973. A numerical taxonomic analysis of the evolutionary relationships of Bombus and Psithyrus (Apidae: Hymenoptera).
Can. Ent. 105: 733-743.

POMERANTZEV, B.I. 1932. [Morphology and anatomy of the genitalia of Culicoides]. [in Russian].
Parazit. Sb. 3: 183-214.

PRICE, D.A. & HARDY, W.T. Isolation of Bluetongue virus from Texas sheep - Culicoides shown to be a vector.
J. Am. vet. med. Ass. 124: 255-261.

RAY, C. 1960. The application of Bergmann's and Allen's rules to the poikilotherms.
J. Morph. 106: 85-108.

REES, J.W. 1969. Morphological variation in the mandibles of white tailed deer (Odocoileus virginianus): a study of populational and skeletal variation by principal component and canonical variate analysis.
J. Morph. 128: 113-130.

REEVE, E.C.R. & HUXLEY, J. 1945. Some problems in the study of allometric growth.
In Gros Clarke, W.E. & Medewar, P.B. (eds) Essays on growth and form: 121-156.
Clarendon Press, Oxford.

REID, J.A. 1942. The relative sizes of different parts in beetles of the genus Laemophloeus (Coleoptera: Cucujidae).
Proc. R. ent. Soc. Lond. (A) 17: 19-26.

REMM, K. 1956. On the midge fauna of the genus Culicoides (Diptera: Heleidae) of Esthonia.
Rev. Ent., Wash. 35: 172-183.

REMM, K. 1973. Beitrage zur Ceratopogoniden fauna Ungarns (Diptera).
Folia ent. hung. 26: 349-357.

REMM, K. 1975. [A contribution to the classification of the biting midges]. [in Russian].
Parazitologiya 9: 393-397.[Translated by Translation Bureaux, Dept. of Secretary of State, Canada. No. 767324.].

REMM, K. 1976. Late cretaceous biting midges (Diptera: Ceratopogonidae) from fossil resins of the Khatanga depression.
Paleont. J. 10: 344-351.

REMM, K. & ZHOGOLOV, D.T. 1968. On the biting midge fauna (Diptera: Ceratopogonidae) of the Crimea.
Ent. Rev., Wash. 47: 503-513.

REUBEN, R. 1959. The ecology of biting midges (Culicoides species) in Scotland.
PhD Thesis, University of Edinburgh. [not seen].

REUBEN, R. 1963. A comparison of trap catches of Culicoides impunctatus Goetghebuer (Diptera: Ceratopogonidae) with meteorological data.
Proc. R. ent. Soc. Lond. (A) 38: 181-193.

RIEB, J.P. & KREMER, M. 1977. Preliminary note on the Culicoides of the Ried of Alsace, with special reference to halophilic species.
Mosquito News 37: 288.

RIEK, R.F. 1954. Studies on allergic dermatitis (Queensland Itch) of the horse: the aetiology of the disease.
Aust. J. agric. Res. 5: 109-129.

ROBACK, S.S. & MOSS, W.W. 1978. Numerical taxonomic studies on the congruence of classifications for the genera and subgenera of Macropelopiini and Anatopyniini (Diptera: Chironomidae: Tanypodinae).
Proc. Acad. nat. Sci. Philad. 129: 125-150.

ROHLF, F.J. 1962. A numerical taxonomic study of the genus Aedes (Diptera: Culicidae) with emphasis on the congruence of larval and adult classifications.
PhD Thesis. University of Kansas: 98pp.

ROHLF, F.J. 1963. Congruence of larval and adult classifications in Aedes (Diptera: Culicidae).
Syst. Zool. 12: 97-117.

ROHLF, F.J. 1970. Adaptive hierarchical clustering schemes.
Syst. Zool. 19: 58-82.

ROHLF, F.J. 1972. An empirical comparison of three ordination techniques in numerical taxonomy.
Syst. Zool. 21: 271-280.

ROHLF, F.J. & SOKAL, R.R. 1967. Taxonomic structure from randomly and systematically scanned biological images.
Syst. Zool. 16: 246-260.

ROSTRON, J. 1972. A multivariate statistical study of skull measurements of five taxa of gazelles.
Zool. J. Linn. Soc. 4: 1-14.

ROTH, L.M. 1948. A study of mosquito behaviour. An experimental laboratory study of the sexual behaviour of Aedes aegypti.
Am. Midl. Nat. 40: 265-352.

ROTHFELS, K.H. 1956. Black flies: siblings, sex, and species grouping.
J. Hered. 47: 113-122.

ROTHFELS, K.H., FARADAY, R. & KANEPS, A. 1978. A cytological description of sibling species of Simulium venustum and S. verecundum with standard maps for the subgenus Simulium Davies (Diptera).
Can. J. Zool. 56: 1110-1128.

ROWLEY, W.A. & CORNFORD, M. 1972. Scanning electron microscopy of the maxillary palp of selected species of Culicoides.
Can. J. Zool. 50: 1207-1210.

RUSSELL, B. 1918. The philosophy of logical atomism. [transcripts of 8 lectures given in London during 1918].
In Pears, D. (ed.) Fontana/Collins, London: 174pp.

SAILER, R.I., MARKS, E.P. & LIENK, S. 1956. Notes on Culicoides in Alaska (Diptera: Ceratopogonidae).
Mosquito News 16: 270-278.

SANDS, W.A. 1972. The soldierless termites of Africa (Isoptera: Termitidae).
Bull. Br. Mus. nat. Hist. (Ent.), Suppl. 18: 1-244.

SASA, M. 1976. Human filariasis.
University Park Press, Baltimore: 819pp.

SCHEVCHENKO, A.K. 1967. [Biting midges of the genus Culicoides (Diptera: Ceratopogonidae) from the valley of the middle reaches of the Desna]. [in Russian].
Ent. Obozr. 46: 163-179.

SCHNEIDERMAN, H.A. 1977. New ways to probe pattern formation and determination in insects.
Symp. R. ent. Soc. Lond. 8: 3-34.

SCHWANWITSCH, B.N. 1924. On the ground plan of wing-pattern in Nymphalids and certain other families of the rhopalocerus Lepidoptera.
Proc. Zool. Soc. Lond. for 1924: 509-528.

SCUDDER, G.G.E. 1974. Species concepts and speciation.
Can. J. Zool. 52: 1121-1134.

SEAL, H.L. 1964. Multivariate statistical analysis for Biologists.
Methuen, London: 207pp.

SELEDTSOV, I.I. & KOSTYRYO, I.N. 1969. [Preliminary data on the possible role of Culicoides nubeculosus in the circulation of West Nile virus] . [in Russian].
Medskaya Parasit. 38: 496.

SERVICE, M.W. 1968. Blood digestion and oviposition in Culicoides impunctatus Goetghebuer and C. obsoletus (Meigen).
Ann. trop. Med. Parasit. 62: 325-330.

SERVICE, M.W. 1969a. Light-trap catches of Culicoides spp. (Diptera: Ceratopogonidae) from southern England. Bull. ent. Res. 59: 317-322.

SERVICE, M.W. 1969b. Studies on the biting habits of Culicoides impunctatus Goetghebuer, C. obsoletus (Meigen) and C. punctatus (Meigen) (Diptera: Ceratopogonidae) in southern England. Proc. R. ent. Soc. Lond. (A) 44: 1-10.

SERVICE, M.W. 1971. Adult flight activities of some British Culicoides species. J. med. Ent. Honolulu 8: 605-609.

SHARP, N.A. & DYCE, A.L. 1928. Filaria perstans: its development in Culicoides austeni. Trans. R. Soc. trop. Med. Hyg. 21: 371-396.

SHELFORD, V.E. 1917. Color and color-pattern mechanisms of tiger beetles. Illinois biol. Monogr. 3: 395-532.

SHERLOCK, I.A. 1963. O problema do Culicoides em Salvador, Bahia. Revta bras. Malar. Doenc. trop. 15: 567-591.

SHERLOCK, I.A. 1964. Estudo clinico da dematozoonose pela picada de Culicoides. Revta. bras. Malar. Doenc. trop. 16: 51-68.

SHIRAKI, T. 1913. [Investigations on general injurious insects]. [in Japanese and English]. Taiwan Sotokufu noji Shikenjo Tokubetsu Hokoku 8: 286-297.

SIMPSON, G.G. 1961. Principles of animal taxonomy. Columbia University Press, New York: 247pp.

SIMPSON, G.G., ROE, A. & LEWONTIN, R.C. 1960. Quantitative Zoology. Harcourt, New York: 440pp.

SKIERSKA, B. 1972. The biting midges of the North Western Territories of Poland. Wiad. parazyt. 18: 815-818.

SLIFER, E. 1970. The structure of Arthropod chemoreceptors. A. Rev. Ent. 15: 121-142.

SNEATH, P.H.A. 1967. Trend-surface analysis of transformation grids. J. Zool., Lond. 151: 65-122.

SNEATH, P.H.A. & SOKAL, R.R. 1973. Numerical taxonomy. Freeman, San Francisco: 573pp.

SNODGRASS, R. 1944. The feeding apparatus of the biting and sucking insects affecting man and animals. Smithson. misc. Collns 104: 1-113.

SOKAL, R.R. & CROVELLO, T.J. 1970. The biological species concept: a critical evaluation. Am. Nat. 104: 127-153.

SOKAL, R.R. & SNEATH, P.H.A. 1963. Principles of Numerical Taxonomy. Freeman, San Francisco: 359pp.

SOLOVEY, V.Y.A. & LIKHODED, V.G. 1966. The blood-sucking Diptera of the north western part of Murmansk province. Ent. Rev., Wash. 45: 319-321.

SØRENSEN, G.H. 1974. Studies on the aetiology and transmission of summer mastitis. Nord. Vet. Med. 26: 122-132.

SPRATT, D.M., DYCE, A.L. & STANDFAST, H.A. 1978. Onchocerca sweetae (Nematoda: Filariodea): notes on intermediate hosts. J. Helminth. 52: 75-81.

STANDFAST, H.A. & DYCE, A.L. 1968. Attacks on cattle by mosquitoes and biting midges. Aust. vet. J. 44: 585-586.

STANDFAST, H.A. & DYCE, A.L. 1972. Potential vectors of cattle and buffalo in Australia. Aust. vet. J. 48: 224-227.

STEWARD, J.S. 1933. Onchocerca cervicalis (Railliet & Henry) and its development in Culicoides nubeculosus (Mg). Rep. Inst. Anim. Path. Univ. Camb. 3: 272-284.

SZABÓ, J.B. 1965. Measures taken against the damage to health of Culcioides nubeculosus (Mg.) 1818, in the area of Hajdúböszörmény. Opusc. zool. Bpest. 2: 229-234.

TAKAHASHI, H. 1941. Notes on some species of the genus Culicoides from Manchoukuo with a description of a new species (Diptera: Ceratopogonidae). Insecta matsum. 14: 80-84.

TEISSIER, G. 1960. Relative growth. In Waterman, T.H. (ed.) The physiology of the Crustacea. Vol. 1, chapt. 16. Academic Press, London.

THIENEMANN, A. 1954. Chironomus. Binnengewasser 20:1-834.

THORPE, R.S. 1976. Biometric analysis of geographic variation and racial affinity. Biol. Rev. 51: 407-452.

TISCHNER, H. 1953. Uber den Gehorsinn von Stechmücken. Acustica 3: 335-343. [not seen].

duTOIT, R.M. 1944. The transmission of blue-tongue and horse sickness by Culicoides. Onderstepoort J. vet. Res. 19: 7-16.

TOKUNAGA, M. 1937. Sandflies (Ceratopogonidae: Diptera) from Japan. Tenthredo 1: 233-338.

TOKUNAGA, M. 1940. Biting midges from Japan and neighbouring countries including Micronesian Islands, Manchuria, North China and Mongolia.
Tenthredo 3: 58-186.

TOKUNAGA, M. 1941. Biting midges from Manchuria (Ceratopogonidae: Diptera).
Insecta matsum. 15: 89-102.

TOWNSEND, C.H.T. 1935. Manual of Myiology in twelve parts. Part II Muscoid classification and habits.
Townsend, Itaquaquecetuba, São Paulo: 296pp.

TURNER, E.C., WEHRHEIM, N.L., MESSERSMITH, D.H. & DAVIS, J.W. 1963. Transmission studies of Avian infectious synovitis by selected arthropod vectors.
Poult. Sci. 42: 1434-1441.

VAJIME, C.G. & DUNBAR, R.W. 1975. Chromosomal identification of of eight species of the subgenus Edwardsellum near and including Simulium (Edwardsellum) damnosum Theobold (Diptera: Simuliidae).
Tropenmed. Parasit. 26: 111-138.

VANE-WRIGHT, R.I. & HUGGINS, C.F. 1972. The superspecies Ethope himachala (Moore), and the identity of Zethera noire Janet (Lepidoptera: Nymphalidae, Satyrinae).
J. Ent. (B) 41 (1) : 1-22.

VIOLOVICH, N.A. 1965. The blood sucking Diptera of Tuva.
Ent. Rev., Wash. 44: 489-492.

WALKER, A.R. & BOREHAM, P.F.L. 1976. Blood feeding of Culicoides (Diptera: Ceratopogonidae) in Kenya in relation to the epidemiology of bluetonge and ephemeral fever.
Bull. ent. Res. 66: 181-188.

WALKER, A.R. & DAVIES, F.G. 1971. A preliminary survey of the epidemiology of bluetongue in Kenya.
J. Hyg., Camb. 69: 47-60.

WHITE, G.B. 1977. The place of morphological studies in the investigation of Anopheles species complexes.
Mosquito Syst. 9: 1-24.

WHITE, G.B. 1979. The identification of mosquitoes as vectors of malaria and filariasis.
Symp. Br. Soc. Parasit. 17: 103-143.

WHITE, G.B., COLUZZI, M. & ZAHAR, A.R. 1975. Review of cytogenetic studies on anopheline vectors of malaria.
Unpublished document WHO/VBC/75.538 - WHO/MAL/75.849.
Geneva: World Health Organisation, 36pp.

W.H.O. 1977. Species complexes in insect vectors of disease (blackflies, mosquitoes, tsetse flies).
Unpublished document WHO/VBC/77.656 - WHO/ONCHO/77.131.
Geneva: World Health Organisation, 56pp.

WILLIAMS, C.B. 1972. A graphical method of demonstrating allometric growth, with special reference to the antennae of Thysanoptera.
J. Ent. (B) 41: 151-153.

WILLIAMS, W.T. 1971. Principles of clustering.
A. Rev. Ecol. Syst. 2: 303-326.

WINNERTZ, J. 1852. Beitrag zur kenntniss der gattung Ceratopogon Meigen.
Linn. ent. 6: 1-80.

WIRTH, W.W. 1952. The Heleidae of California.
Univ. Calif. Publs Ent. 9: 95-266.

WIRTH, W.W. & ARNAUD, P.H. 1969. Polynesian biting midges of the genus Culicoides (Diptera: Ceratopogonidae).
Pacif. Insects 11: 507-520.

WIRTH, W.W. & ATCHLEY, W.R. 1973. A review of the North American Leptoconops (Diptera: Ceratopogonidae).
Grad. Stud. Texas Tech. Univ. 5: 1-57.

WIRTH, W.W. & BLANTON, F.S. 1969. North American Culicoides of the pulicaris group (Diptera: Ceratopogonidae).
Fla. Ent. 52: 207-243.

WIRTH, W.W. & JONES, R. 1957. The North American sub-species of Culicoides variipennis.
Tech. Bull. U.S. Dept. Agric. 1170: 1-35.

WIRTH, W.W. & LEE, D.J. 1959. Australasian Ceratopogoninae (Diptera: Ceratopogonidae) pt 8: A new genus from Western Australia attacking man.
Proc. Linn. Soc. N.S.W. 83: 337-339.

WIRTH, W.W., RATANAWORABHAN, N.C. & BLANTON, F.S. 1974. Synopsis of the genera of Ceratopogonidae (Diptera).
Annls Parasit. hum. comp. 49: 595-613.

WISHART, G. & RIORDAN, D.F. 1959. Flight responses to various sounds by adult males of Aedes aegypti (Diptera: Culicidae).
Can. Ent. 91: 181-191.

WOLFE, J.H. 1970. Pattern clustering by multivariate mixture analysis.
Multivariate Behav. Res. 5(3): 329-350.

WU, C.J. & WU, S.Y. 1957. Isolation of the virus B type encephalitis from Lasiohelea taiwana Shiraki, a blood-sucking midge.
Acta Microbiol. sin. 5: 22-26.

YOUNKER, J.L. & EHRLICH, R. 1977. Fourier biometrics: harmonic amplitudes as multivariate shape descriptors.
Syst. Zool. 26: 336-342.

ZILAHI-SEBESS, G. 1933. Verszivo Chironomidaink.
Allatt. Közl. 30: 146-151.

ZILAHI-SEBESS, G. 1936. A balanton partvidék Heleidai.
Arb. ung. biol. Forsch Inst. 8: 196-206.

ADDENDA

ANDERSON, T.W. 1958. An introduction to multivariate statistical analysis.
Wiley, New York: 374pp.

BALAKRISHNANT, V. & SANGHVI, L.D. 1968. Distance between populations on the basis of attribute data.
Biometrics 27: 825-840.

BEZDEK, J.C. 1975. Mathematical models for systematics and taxonomy.
Int. Cong. num. Taxon. 8: 141-166.

BURNABY, T.P. 1966. Growth-invariant discriminant functions and generalised distances.
Biometrics 22: 96-110.

COOLEY, W.W. & LOHNES, P.R. 1971. Multivariate data analysis.
Wiley, New York: 364pp.

CORRUCCINI, R.S. 1972. Allometry correction in taximetrics.
Syst. Zool. 21: 375-383.

CORRUCCINI, R.S. 1975. Multivariate analysis in biological anthropology: some considerations.
J. hum. Evol. 4: 1-19.

EHRLICH, P.R. & RAVEN, P.H. 1969.
Differentiation of populations.
Science, N.Y. 165: 1228-1232.

ELMES, G.W. 1978. A morphometric comparison of three closely related species of Myrmica (Formicidae), including a new species from England.
Syst. Ent. 3: 131-145.

GYLLENBURG, H.G. 1965. A model for computer identification of microorganisms.
J. gen. Microbiol. 39: 401-405.

HARWICK, D.F. & LEFKOVITCH, L.P. 1973. Species separation in the declarata group of the genus Euxoa, a computer analysis based on structural characters.
Can. Ent. 105: 501-508.

MARRIOTT, F.M.C. 1974. The interpretation of multiple observations.
Academic Press, London: 118pp.

SOKAL, R. 1975. Discussion of preceding papers.
Int. Congr. num. Taxon. 8: 167-175.

STERN, W.T. 1969. The Jamaicam species of Columnea and Alloplectus (Gesneriaceae).
Bull. Br. Mus. nat. Hist. (Bot.) 4: 181-236.

WRIGHT, S. 1978. Evolution and the genetics of populations Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago: 580pp.

ZADEH, L.A. 1965. Fuzzy sets. Inf. Control 8: 338-353.

ZADEH, L.A. 1977. Fuzzy sets and their application to pattern classification and analysis. In van Ryzin, J. (ed.) Classification and Clustering: 251-299. Academic Press, New York.

APPENDIX

Included in the appendix are the following primary data matrices:

1. Scans of 23 wings (Pages 360 - 362).

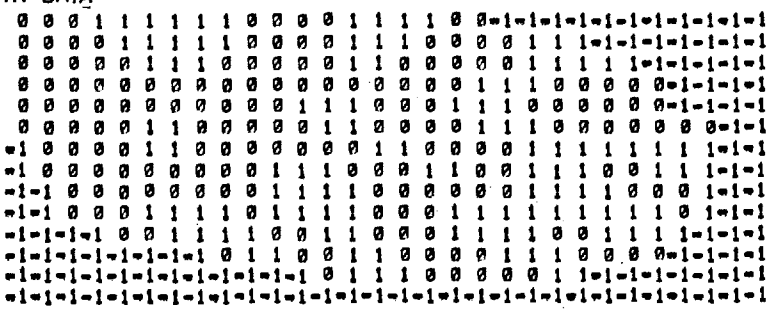
2. 13 pattern elements scored for 23 wings (P.363).

3. 72 variables scored for 84 specimens (Pages 364 - 371).
   The specimens used in the analysis of 53 OTU's are indicated by an asterisk and the code number used in these analyses is given in brackets.
   Variables may be identified from Table 21 (p.188) and collection data for the specimens in Table 27 (p.215).

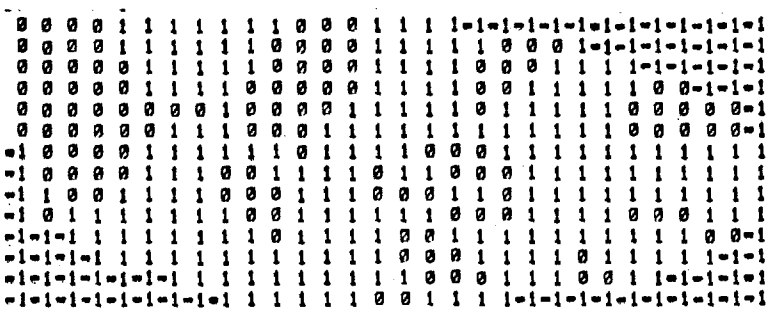
4. Data matrix and summary of seven taxa used in the canonical variate analysis (Section 10). (Pages 372 - 375).

WING SCAN DATA

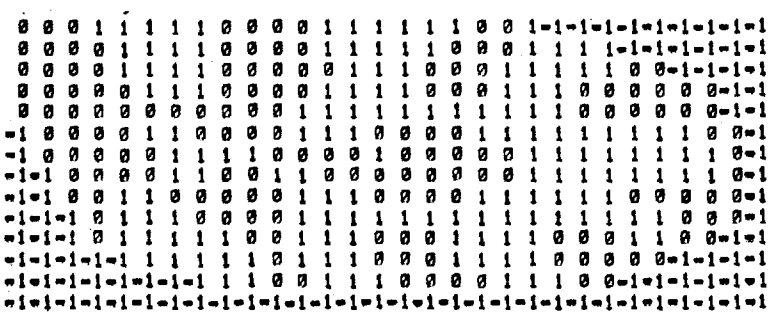
OTU 1



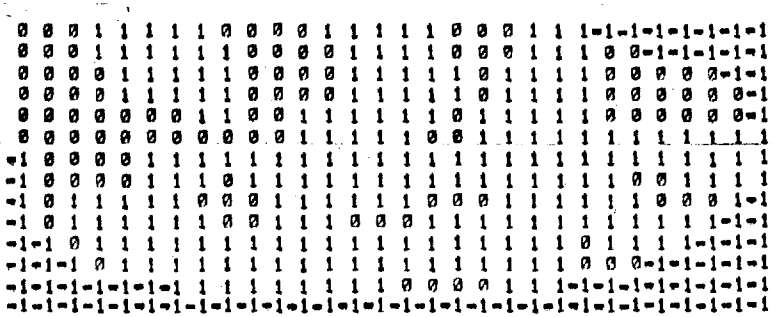
OTU 2



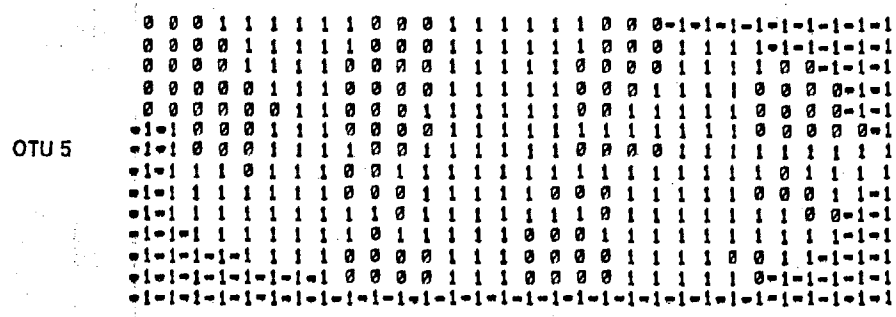
OTU 3



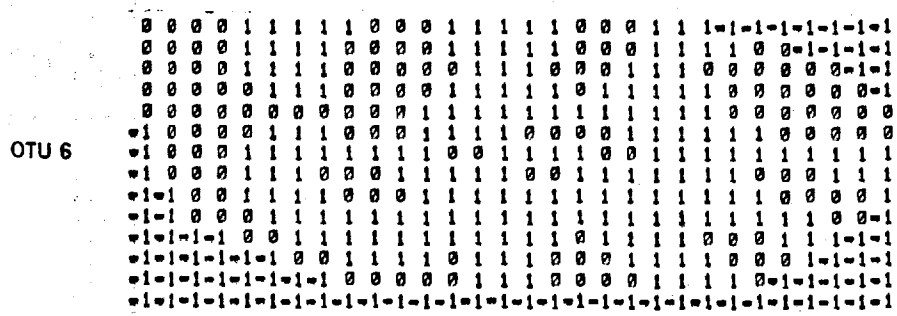
OTU 4



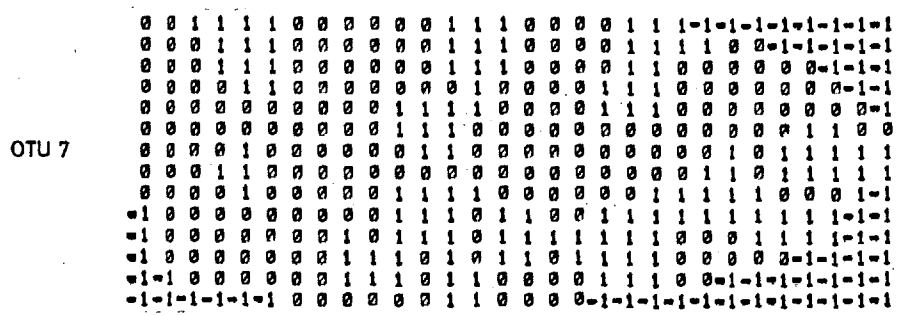
OTU 5



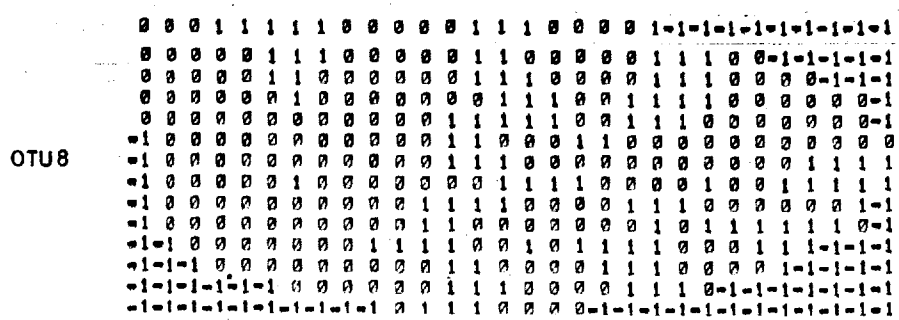
OTU 6



OTU 7

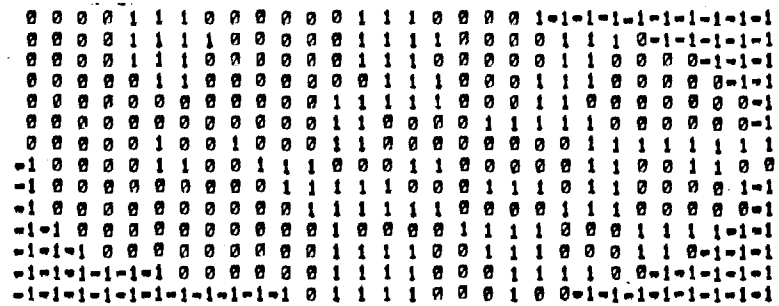


OTU 8

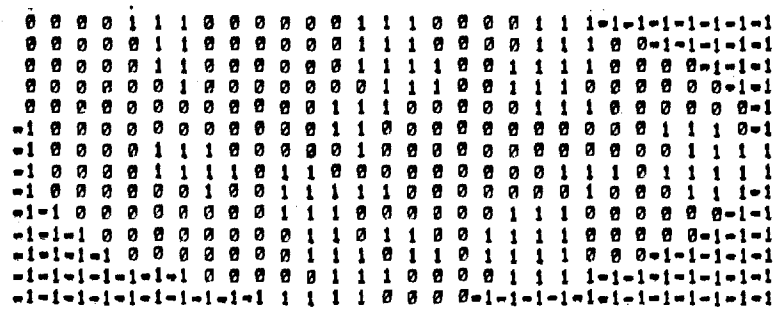


WING SCAN DATA

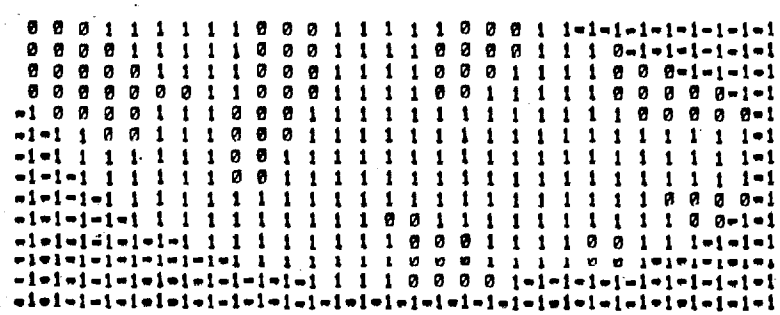
OTU 9



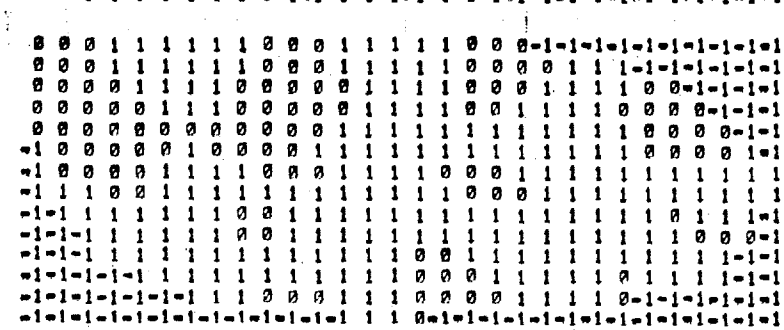
OTU 10



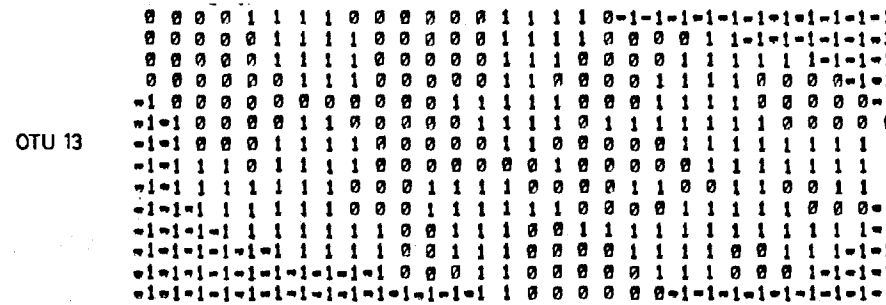
OTU 11



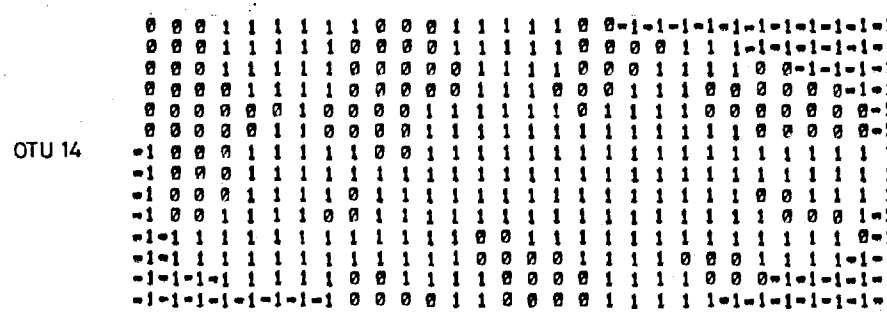
OTU 12



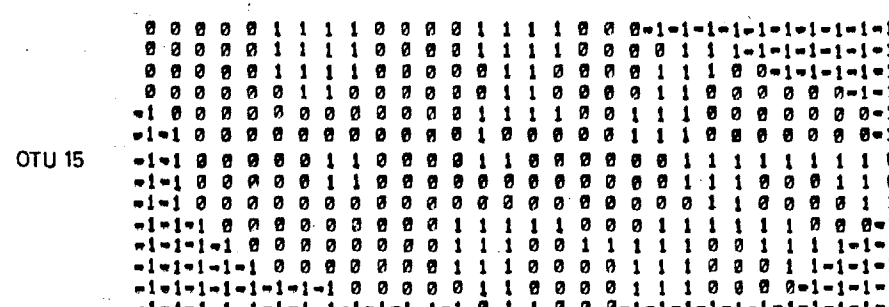
OTU 13



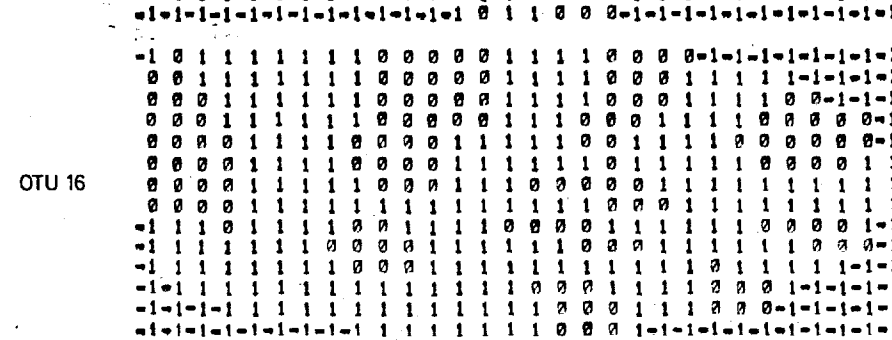
OTU 14



OTU 15

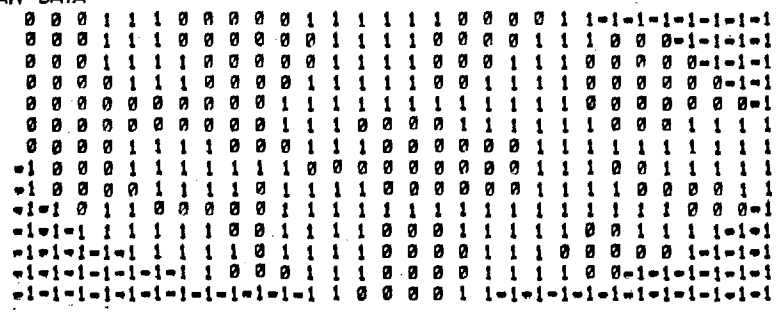


OTU 16

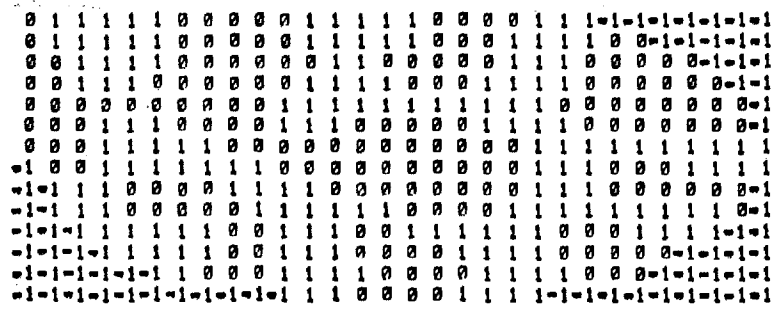


WING SCAN DATA

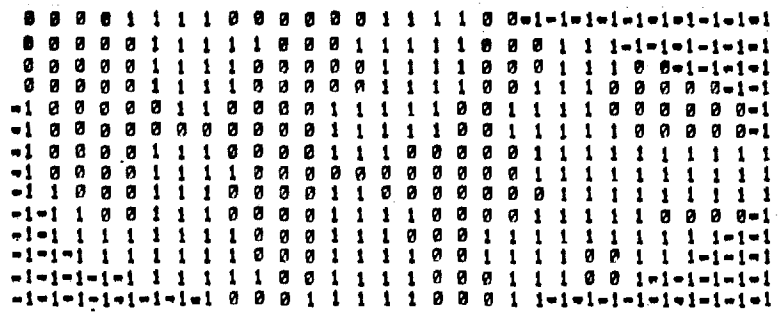
OTU 17



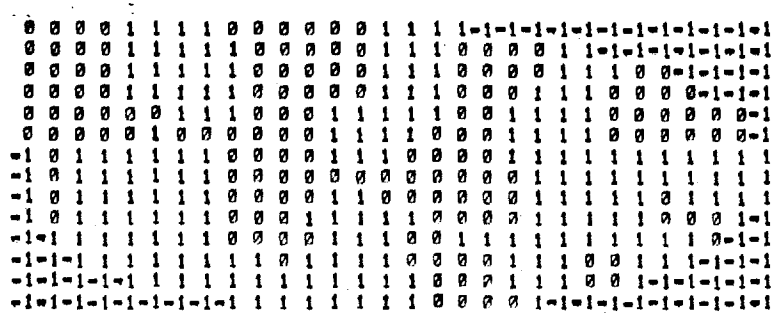
OTU 18



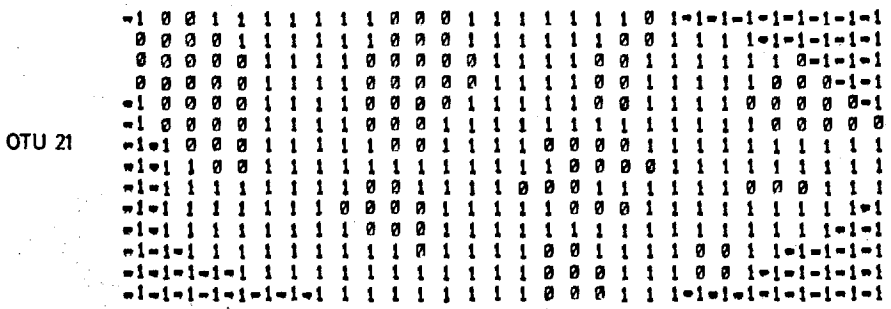
OTU 19



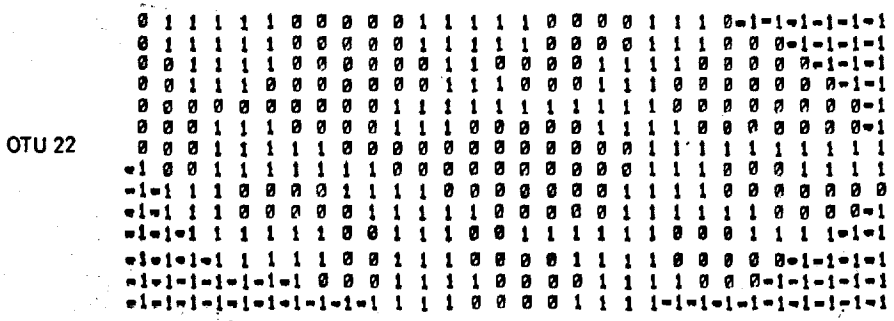
OTU 20



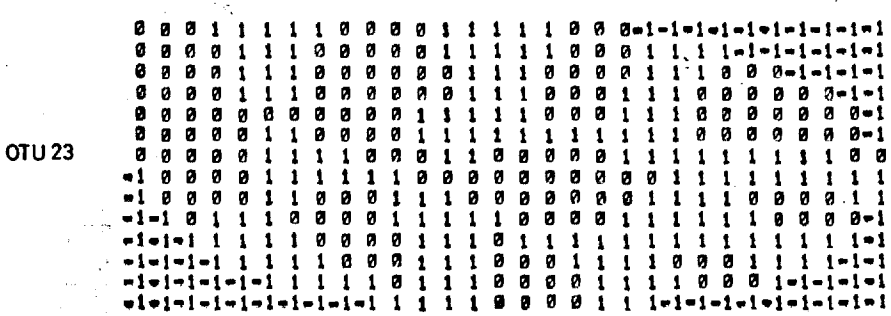
OTU 21



OTU 22



OTU 23



PRIMARY DATA MATRIX 2. Wing pattern elements coded for 23 wings.

OTU'S	PATTERN ELEMENTS												
	2.	1.	2.	2.	3.	2.	3.	2.	0.	2.	1.	2.	1.
	3.	3.	4.	4.	5.	2.	3.	3.	0.	2.	2.	5.	2.
	2.	3.	4.	4.	2.	2.	1.	3.	0.	2.	3.	2.	2.
	3.	3.	4.	3.	4.	2.	3.	3.	1.	2.	2.	5.	2.
	2.	3.	4.	4.	4.	2.	3.	3.	0.	2.	2.	4.	2.
	2.	3.	3.	4.	3.	2.	2.	3.	2.	2.	3.	3.	3.
	1.	1.	2.	1.	3.	1.	2.	2.	1.	1.	2.	1.	1.
	1.	2.	2.	3.	2.	1.	1.	1.	1.	2.	1.	1.	1.
	2.	3.	3.	3.	2.	1.	1.	1.	1.	2.	2.	0.	1.
	1.	2.	2.	1.	3.	1.	0.	2.	1.	2.	1.	0.	1.
	3.	3.	3.	5.	5.	2.	3.	3.	0.	2.	2.	5.	2.
	3.	4.	3.	4.	5.	2.	3.	3.	0.	2.	2.	4.	2.
	2.	3.	3.	3.	4.	1.	3.	3.	0.	1.	2.	3.	2.
	3.	3.	3.	5.	5.	2.	3.	3.	0.	2.	3.	4.	3.
	2.	2.	1.	1.	3.	2.	2.	2.	0.	2.	1.	0.	1.
	3.	3.	4.	4.	4.	2.	2.	3.	0.	2.	3.	4.	2.
	1.	2.	1.	2.	3.	3.	2.	2.	0.	2.	2.	2.	1.
	1.	2.	1.	2.	3.	3.	2.	3.	0.	3.	2.	2.	2.
	3.	4.	4.	3.	5.	2.	3.	3.	0.	2.	2.	4.	2.
	3.	3.	4.	2.	5.	2.	3.	3.	0.	2.	2.	5.	2.
	3.	4.	4.	4.	4.	2.	3.	3.	0.	2.	3.	5.	1.
	1.	2.	1.	2.	3.	3.	2.	3.	0.	3.	2.	2.	2.
	1.	2.	1.	2.	3.	3.	2.	2.	0.	2.	2.	2.	2.

PRIMARY DATA MATRIX 3

OTU 1	4.0000	.0768	.0560	.0560	.0622	.0622	.0622	.0622	.0622	.072
(1)	.0747	.0934	.1037	.1556	2.0000	0.0000	1.0000	1.0000	1.0000	2.000
	63.8000	71.9200	25.5200	29.0000	27.8400	213.7499	166.2500	80.7500	123.5000	15.000
	15.0000	119.9999	51.6000	68.4000	3.0000	3.0000	4.0000	5.0000	5.0000	2.000
	3.0000	3.0000	2.0000	2.0000	3.0000	3.0000	3.0000	346.7498	360.9998	185.249
	99.7500	66.5000	451.2498	460.7498	223.2499	104.5000	71.2500	451.2498	451.2498	199.499
	128.2499	80.7500	5.0000	0.0000	1.0000	2.5833	1.2857	1.1714	3.4429	.653
	1.0000	.5700								
OTU 2	4.0000	.0720	.0554	.0572	.0590	.0646	.0646	.0590	.0590	.077
(2)	.0812	.0941	.0996	.1568	3.0000	1.0000	1.0000	1.0000	1.0000	2.000
	75.4000	76.5600	34.8000	37.1200	30.1600	242.2499	199.4999	104.5000	151.9999	15.000
	15.0000	136.7999	62.4000	74.4000	3.0000	4.0000	4.0000	4.0000	5.0000	2.000
	3.0000	3.0000	2.0000	2.0000	2.0000	3.0000	3.0000	43$.9998	455.9997	223.249
	128.2499	80.7500	593.7500	593.7500	275.4999	128.2499	90.2500	555.7498	574.7495	261.250
	161.4999	95.0000	6.0000	0.0000	1.0376	2.5385	1.2143	1.1488	3.2262	.687
	1.0000	.5439								
OTU 3	3.0000	.0803	.0592	.0634	.0571	.0529	.0529	.0550	.0592	.069
(3)	.0846	.0930	.1036	.1691	2.0000	0.0000	1.0000	1.0000	1.0000	2.000
	69.6000	58.0000	25.5200	26.6800	34.8000	204.2499	166.2500	90.2500	113.9999	17.000
	16.0000	109.2000	46.8000	57.6000	2.0000	3.0000	3.0000	4.0000	5.0000	2.000
	3.0000	3.0000	2.0000	2.0000	3.0000	4.0000	3.0000	342.0000	342.0000	166.250
	99.7500	61.7500	441.7497	427.4999	223.2499	109.2500	66.5000	436.9998	14.2500	204.249
	128.2499	76.0000	6.0000	0.0000	1.0837	1.6667	1.2286	1.1071	3.3786	.791
	1.0625	.5275								
OTU 4	3.0000	.0624	.0546	.0585	.0604	.0604	.0585	.0604	.0604	.078
(4)	.0838	.0994	.1072	.1559	2.0000	0.0000	1.0000	1.0000	2.0000	2.000
	77.7200	69.6000	23.2000	31.3200	32.4800	208.9999	180.4999	104.5000	142.4999	17.000
	13.0000	131.9999	57.6000	74.4000	2.0000	2.0000	2.0000	4.0000	4.0000	3.000
	3.0000	3.0000	1.0000	1.0000	3.0000	3.0000	2.0000	451.2498	475.0000	237.500
	147.2499	80.7500	484.4999	475.0000	251.7499	118.7500	76.0000	451.2498	479.7498	237.500
	132.9999	85.5000	6.0000	0.0000	1.1025	2.1429	1.1579	1.1447	3.3750	.733
	1.3077	.5636								
OTU 5	5.0000	.0734	.0570	.0571	.0571	.0603	.0619	.0603	.0571	.075
	.0888	.0978	.1027	.1517	2.0000	1.0000	0.0000	1.0000	1.0000	2.000
	87.0000	92.8000	32.4800	41.7600	40.6000	261.2500	213.7499	109.2500	166.2500	19.000
	16.0000	161.9999	72.0000	101.9999	3.0000	4.0000	3.0000	4.0000	3.0000	2.000
	3.0000	3.0000	2.0000	2.0000	3.0000	4.0000	2.0000	475.0000	475.0000	270.750
	132.9999	95.0000	593.7500	603.2497	308.7499	142.4999	104.5000	569.9996	588.9997	303.999
	175.7500	118.7500	6.0000	0.0000	1.0639	2.2837	1.2220	1.2166	3.4055	.657
	.8421	.6296								
OTU 6	5.0000	.0778	.0584	.0564	.0583	.0631	.0622	.0622	.0603	.075
	.0836	.1011	.1031	.1400	3.0000	1.0000	1.0000	1.0000	1.0000	2.000
	69.6000	69.6000	16.2400	30.1600	32.4800	227.9998	171.0000	95.0000	132.9999	16.000
	16.0000	129.5999	57.6000	72.0000	3.0000	3.0000	3.0000	4.0000	4.0000	3.000
	3.0000	3.0000	1.0000	1.0000	3.0000	4.0000	3.0000	356.2500	379.9999	204.249
	104.5000	61.7500	451.2498	475.0000	246.9999	104.5000	71.2500	451.2498	475.0000	237.500
	118.7500	80.7500	6.0000	0.0000	1.0156	2.1428	1.3333	1.1111	3.5694	.714
	1.0000	.5555								
OTU 7	7.0000	.0773	.0562	.0548	.0562	.0615	.0579	.0579	.0579	.080
	.0826	.0949	.1036	.1581	2.0000	1.0000	1.0000	1.0000	1.0000	3.000
	81.1999	78.8800	29.0000	37.1200	32.4800	237.5000	180.4999	118.7500	161.4999	17.000
	14.0000	149.9999	60.0000	93.6000	2.0000	3.0000	3.0000	4.0000	4.0000	3.000
	3.0000	3.0000	1.0000	2.0000	3.0000	4.0000	3.0000	408.4999	427.4999	227.999
	128.2499	85.5000	546.2497	541.4997	284.9999	132.9999	85.5000	522.4996	522.4996	265.999
	171.0000	104.5000	6.0000	0.0000	1.0842	2.4285	1.3157	1.2828	3.7430	.735
	.8235	.6240								
OTU 8	7.0000	.0716	.0459	.0596	.0596	.0630	.0630	.0630	.0650	.080
	.0885	.0936	.1022	.1448	3.0000	1.0000	1.0000	1.0000	1.0000	2.000
	87.0000	84.6799	32.4800	37.1200	37.1200	246.9999	199.4999	104.5000	156.7500	18.000
	13.0000	149.9999	66.0000	83.9999	2.0000	2.0000	3.0000	4.0000	4.0000	3.000
	3.0000	3.0000	1.0000	2.0000	3.0000	4.0000	3.0000	413.2499	422.7499	223.249
	123.5000	76.0000	522.4996	531.9997	284.9999	132.9999	95.0000	522.4996	531.9997	284.999
	161.4999	104.5000	5.0000	0.0000	1.0380	2.2812	1.2380	1.2320	3.4940	.666
	.7220	.5600								
OTU 9	5.0000	.0728	.0607	.0546	.0566	.0607	.0607	.0607	.0607	.074
	.0910	.0951	.0951	.1558	3.0000	1.0000	1.0000	1.0000	1.0000	2.000
	63.8000	69.6000	23.2000	34.8000	31.3200	213.7499	175.7500	80.7500	118.7500	17.000
	16.0000	131.9999	56.4000	74.4000	3.0000	3.0000	4.0000	5.0000	5.0000	3.000
	3.0000	3.0000	1.0000	2.0000	4.0000	5.0000	3.0000	370.4998	384.7499	189.999
	95.0000	71.2500	460.7498	475.0000	237.5000	109.2500	71.2500	451.2498	484.4999	227.999
	142.4999	90.2500	6.0000	0.0000	1.0497	2.2222	1.2160	1.1148	3.3370	.680
	.9411	.5636								
OTU 10	5.0000	.0772	.0579	.0579	.0579	.0637	.0617	.0579	.0617	.081
	.0868	.0926	.1003	.1428	3.0000	1.0000	1.0000	1.0000	1.0000	2.000
	78.8800	71.9200	25.5200	34.8000	32.4800	213.7499	199.4999	85.5000	104.5000	15.000
	16.0000	134.3999	57.6000	78.0000	3.0000	3.0000	3.0000	5.0000	5.0000	3.000
	3.0000	3.0000	1.0000	2.0000	4.0000	5.0000	3.0000	379.9999	379.9999	199.499
	109.2500	71.2500	475.0000	475.0000	237.5000	109.2500	76.0000	451.2498	475.0000	227.999
	142.4999	80.7500	6.0000	0.0000	1.0155	2.2142	1.0714	1.0833	3.0833	.818
	1.0666	.5803								
OTU 11	5.0000	.0609	.0513	.0513	.0513	.0626	.0577	.0577	.0577	.077
	.0930	.1075	.1075	.1573	3.0000	1.0000	1.0000	1.0000	2.0000	2.000
	87.0000	83.5200	34.8000	38.2800	40.6000	261.2500	213.7499	113.9999	171.0000	19.000
	16.0000	173.9999	74.4000	101.9999	3.0000	3.0000	3.0000	4.0000	5.0000	3.000
	3.0000	3.0000	1.0000	2.0000	4.0000	5.0000	3.0000	498.7499	512.9995	261.250
	142.4999	99.7500	626.9998	674.4995	332.5000	166.2500	99.7500	598.4996	641.2496	308.749
	189.9999	118.7500	6.0000	0.0000	1.1859	2.0571	1.2222	1.1666	3.4611	.666
	.8421	.5862								

OTU 12	5,0000	,0771	,0514	,0531	,0582	,0600	,0600	,0600	,0583	,0771
	,0840	,1080	,1020	,1492	3,0000	1,0000	1,0000	1,0000	2,0000	2,0000
	78,8800	75,4000	34,8000	35,9600	29,0000	227,9998	189,9999	95,0000	151,9999	17,0000
	15,0000	149,9999	63,6000	83,9999	3,0000	3,0000	4,0000	4,0000	5,0000	3,0000
	3,0000	3,0000	1,0000	2,0000	3,0000	5,0000	3,0000	451,2498	451,2498	237,5000
	123,5000	80,7500	546,2497	522,4996	284,9999	142,4999	95,0000	569,9996	555,7498	284,9999
	171,0000	95,0000	6,0000	0,0000	1,0896	2,6000	1,2000	1,2125	3,6437	,6250
	,8823	,5600								
OTU 13	4,0000	,0705	,0508	,0564	,0592	,0635	,0592	,0564	,0550	,0790
(5)	,0832	,1016	,1100	,1551	3,0000	1,0000	1,0000	1,0000	3,0000	3,0000
	104,3999	106,7200	40,6000	47,5600	26,6800	294,4999	284,9999	132,9999	189,9999	20,0000
	15,0000	179,9999	83,9999	117,5999	3,0000	4,0000	4,0000	4,0000	4,0000	2,0000
	3,0000	3,0000	0,0000	2,0000	3,0000	5,0000	1,0000	522,4996	484,4999	280,2498
	161,4999	99,7500	641,2496	626,9998	342,0000	166,2500	99,7500	664,9999	669,7494	365,7499
	208,9999	113,9999	6,0000	0,0000	1,1228	4,0000	1,0333	1,0750	2,9542	,7000
	1,3333	,6533								
OTU 14	4,0000	,0711	,0558	,0586	,0600	,0628	,0572	,0558	,0558	,0851
(6)	,0879	,0948	,1060	,1492	3,0000	1,0000	1,0000	1,0000	3,0000	3,0000
	110,1999	110,1999	40,6000	52,2000	26,6800	284,9999	294,4999	118,7500	189,9999	21,0000
	15,0000	185,9999	78,0000	119,9999	3,0000	3,0000	4,0000	4,0000	4,0000	2,0000
	2,0000	3,0000	0,0000	2,0000	3,0000	4,0000	2,0000	512,9995	508,2496	256,4998
	142,4999	95,0000	626,9998	617,4997	322,9999	166,2500	95,0000	664,9999	664,9999	346,7498
	204,2499	113,9999	6,0000	0,0000	1,0965	4,1304	,9677	1,0887	2,8911	,6250
	1,4000	,6452								
OTU 15	4,0000	,0709	,0506	,0532	,0582	,0595	,0570	,0570	,0570	,0848
(7)	,0886	,1025	,1063	,1544	3,0000	1,0000	1,0000	1,0000	2,0000	3,0000
	115,9999	113,6799	47,5600	52,2000	30,1600	308,7499	308,7499	132,9999	199,4999	21,0000
	15,0000	201,5998	91,1999	129,5999	3,0000	4,0000	4,0000	3,0000	5,0000	2,0000
	3,0000	3,0000	0,0000	2,0000	2,0000	4,0000	2,0000	569,9996	565,2496	294,4999
	171,0000	113,9999	721,9995	712,4998	379,9999	189,9999	109,2500	702,9996	759,9996	356,2500
	218,4999	132,9999	6,0000	0,0000	1,1585	3,7692	1,0000	1,0923	3,0385	,6667
	1,4000	,6429								
OTU 16	4,0000	,0618	,0484	,0538	,0565	,0605	,0565	,0551	,0551	,0780
(8)	,0847	,1089	,1102	,1707	3,0000	1,0000	1,0000	1,0000	2,0000	3,0000
	115,9999	110,1999	40,6000	51,0400	24,3600	299,2499	294,4999	128,2499	194,7499	21,0000
	14,0000	183,5998	83,9999	119,9999	3,0000	3,0000	4,0000	4,0000	4,0000	2,0000
	3,0000	3,0000	0,0000	2,0000	2,0000	5,0000	3,0000	522,4996	527,2497	275,4999
	161,4999	104,5000	664,9999	617,4997	337,2498	161,4999	109,2500	664,9999	655,4996	351,5000
	208,9999	118,7500	7,0000	0,0000	1,2342	4,5238	1,0161	1,1048	3,0000	,6585
	1,5000	,6536								
OTU 17	6,0000	,0675	,0519	,0545	,0597	,0610	,0597	,0532	,0545	,0831
(9)	,0870	,1052	,1091	,1532	3,0000	1,0000	2,0000	2,0000	3,0000	3,0000
	110,1999	117,1599	46,4000	52,2000	30,1600	318,2498	303,9999	137,7499	199,4999	24,0000
	16,0000	185,9999	88,7999	125,9999	3,0000	3,0000	4,0000	2,0000	5,0000	2,0000
	3,0000	3,0000	0,0000	2,0000	2,0000	5,0000	2,0000	555,7498	593,7500	299,2499
	166,2500	109,2500	726,7496	688,7496	332,5000	180,4999	95,0000	688,7496	736,2499	394,2499
	204,2499	118,7500	6,0000	0,0000	1,1629	3,8846	1,0469	1,0977	3,0078	,6905
	1,5000	,6774								
OTU 18	7,0000	,0779	,0554	,0599	,0614	,0584	,0554	,0554	,0569	,0809
	,0869	,0884	,1049	,1574	2,0000	1,0000	1,0000	1,0000	3,0000	2,0000
	110,1999	113,6799	40,6000	46,4000	25,5200	246,9999	256,4998	123,5000	166,2500	23,0000
	16,0000	167,9998	78,0000	110,3999	3,0000	3,0000	4,0000	4,0000	4,0000	3,0000
	3,0000	3,0000	0,0000	2,0000	3,0000	5,0000	2,0000	484,4999	484,4999	256,4998
	128,2499	71,2500	607,9998	598,4996	318,2498	156,7500	95,0000	617,4997	607,9998	332,5000
	189,9999	104,5000	6,0000	0,0000	1,0778	4,4545	,9629	1,2407	3,0879	,7428
	,6956	,6571								
OTU 19	7,0000	,0728	,0500	,0576	,0606	,0606	,0561	,0561	,0606	,0880
	,0880	,0955	,0986	,1547	3,0000	1,0000	1,0000	1,0000	3,0000	4,0000
	110,1999	110,1999	31,3200	40,6000	25,5200	284,9999	261,2500	137,7499	166,2500	19,0000
	16,0000	167,9998	78,0000	105,6000	3,0000	3,0000	4,0000	4,0000	4,0000	3,0000
	3,0000	3,0000	0,0000	2,0000	3,0000	5,0000	2,0000	475,0000	484,4999	246,9999
	142,4999	71,2500	607,9998	598,4996	318,2498	156,7500	95,0000	593,7500	617,4997	308,7499
	180,4999	99,7500	6,0000	0,0000	1,1054	4,3181	1,0909	1,1454	2,9950	,8285
	,8421	,6285								
OTU 20	7,0000	,0695	,0491	,0573	,0573	,0573	,0573	,0559	,0559	,0845
	,0886	,1009	,1159	,1500	3,0000	1,0000	1,0000	1,0000	3,0000	3,0000
	104,3999	110,1999	39,4400	55,6800	31,3200	256,4998	275,4999	123,5000	180,4999	18,0000
	16,0000	191,9998	90,0000	119,9999	3,0000	3,0000	2,0000	4,0000	5,0000	3,0000
	3,0000	3,0000	0,0000	2,0000	4,0000	5,0000	3,0000	522,4996	546,2497	284,9999
	156,7500	99,7500	655,4996	626,9998	356,2500	171,0000	104,5000	664,9999	679,2497	379,9999
	208,9999	109,2500	6,0000	0,0000	1,1750	3,5185	,9310	1,1508	3,1590	,6842
	,9269	,6250								
OTU 21	5,0000	,0780	,0497	,0538	,0578	,0592	,0592	,0565	,0578	,0767
	,0874	,1022	,1103	,1507	3,0000	1,0000	1,0000	1,0000	3,0000	3,0000
	113,6799	106,7200	39,4400	52,2000	32,4800	256,4998	270,7500	132,9999	189,9999	18,0000
	16,0000	185,9999	62,4000	122,3999	1,0000	3,0000	3,0000	5,0000	5,0000	3,0000
	3,0000	3,0000	0,0000	2,0000	4,0000	5,0000	3,0000	522,4996	531,9997	294,4999
	156,7500	389,4999	679,2497	655,4996	346,7498	171,0000	104,5000	688,7496	688,7496	379,9999
	213,7499	118,7500	7,0000	0,0000	1,1168	3,2857	,9473	1,1798	3,2587	,7000
	,8888	,6580								
OTU 22	5,0000	,0715	,0556	,0572	,0556	,0572	,0588	,0577	,0572	,0827
	,0842	,0938	,1049	,1637	3,0000	1,0000	1,0000	1,0000	3,0000	3,0000
	104,3999	100,9199	41,7600	40,6000	23,2000	251,7499	251,7499	128,2499	166,2500	19,0000
	15,0000	152,3999	72,0000	98,4000	1,0000	3,0000	4,0000	4,0000	5,0000	3,0000
	3,0000	3,0000	0,0000	2,0000	4,0000	5,0000	3,0000	451,2498	460,7498	242,2499
	118,7500	76,0000	550,9996	546,2497	294,4999	142,4999	85,5000	550,9996	584,2496	284,9999
	171,0000	95,0000	5,5000	0,0000	1,1250	4,3500	1,0000	1,1698	2,9669	,7714

OTU 23	5.0000	.0711	.0496	.0536	.0577	.0604	.0604	.0590	.0563	.0805
	.0859	.0953	.1114	.1583	3.0000	1.0000	2.0000	1.0000	3.0000	3.0000
	115.9999	107.8799	44.0800	53.3600	29.0000	275.4999	294.4999	142.4999	166.2500	19.0000
	15.0000	137.9999	62.4000	122.3999	2.0000	2.0000	3.0000	4.0000	5.0000	3.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	3.0000	2.0000	531.9997	550.9996	284.9999
	161.4999	104.5000	612.7496	655.4996	332.5000	175.7500	109.2500	688.7496	688.7496	370.4998
	208.9999	171.0000	7.0000	0.0000	1.1346	3.7200	.9354	1.1169	3.0040	.8571
	.7894	.6580								
OTU 24	9.0000	.0774	.0557	.0666	.0697	.0681	.0650	.0666	.0681	.0820
(10)	.0836	.0944	.0867	.1161	4.0000	1.0000	2.0000	2.0000	2.0000	2.0000
	83.5200	109.0399	42.9200	51.0400	34.8000	275.4999	232.7499	109.2500	161.4999	21.0000
	16.0000	134.3999	60.0000	61.2000	2.0000	2.0000	1.0000	2.0000	3.0000	3.0000
	2.0000	2.0000	0.0000	2.0000	2.0000	1.0000	1.0000	432.2498	451.2498	227.9998
	123.5000	66.5000	560.4995	550.9996	303.9999	123.5000	80.7500	522.4996	522.4996	270.7500
	161.4999	80.7500	5.0000	1.0000	.8617	3.1333	1.1837	1.2602	3.2959	.6765
	1.3125	.4554								
OTU 25	8.0000	.0614	.0477	.0587	.0587	.0600	.0600	.0559	.0573	.0969
(11)	.0914	.1023	.1023	.1473	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	84.6799	98.6000	40.6000	48.7200	27.8400	284.9999	261.2500	113.9999	261.2500	22.0000
	17.0000	153.5999	64.8000	94.7999	3.0000	3.0000	4.0000	3.0000	4.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	2.0000	2.0000	465.4998	470.2499	242.2499
	128.2499	90.2500	593.7500	579.4996	308.7499	123.5000	95.0000	522.4996	536.7498	284.9999
	161.4999	85.5000	6.0000	1.0000	1.1751	3.5417	1.0909	1.0682	3.3318	.4364
	1.2941	.6172								
OTU 26	9.0000	.0647	.0485	.0566	.0580	.0606	.0580	.0593	.0606	.0889
(12)	.0889	.1011	.1092	.1456	4.0000	1.0000	1.0000	1.0000	3.0000	3.0000
	90.4800	87.0000	37.1200	41.7600	33.6400	261.2500	227.9998	113.9999	156.7500	20.0000
	16.0000	140.3999	60.0000	87.5999	1.0000	2.0000	1.0000	2.0000	3.0000	3.0000
	2.0000	2.0000	0.0000	2.0000	2.0000	2.0000	1.0000	427.4999	451.2498	223.2499
	118.7500	76.0000	555.7498	541.4997	303.9999	123.5000	80.7500	503.4996	536.7498	261.2500
	161.4999	85.5000	6.0000	1.0000	1.1445	2.5862	1.1458	1.1510	3.8646	.7273
	1.2500	.6239								
OTU 27	9.0000	.0584	.0480	.0584	.0597	.0623	.0584	.0597	.0610	.0908
(13)	.0856	.1038	.1115	.1427	4.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	92.8000	102.0800	44.0800	40.6000	25.5200	275.4999	237.5000	113.9999	171.0000	24.0000
	18.0000	143.9999	62.4000	90.0000	1.0000	2.0000	1.0000	2.0000	3.0000	3.0000
	2.0000	3.0000	0.0000	3.0000	2.0000	2.0000	2.0000	455.9997	489.2498	256.4998
	128.2499	80.7500	574.7495	593.7500	332.5000	118.7500	80.7500	546.2497	560.4995	284.9999
	166.2500	95.0000	5.0000	1.0000	1.1476	4.0000	1.1600	1.2050	3.8550	.6667
	1.3333	.6250								
OTU 28	9.0000	.0570	.0598	.0570	.0613	.0627	.0598	.0627	.0598	.0926
(14)	.0855	.0969	.1026	.1425	4.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	92.8000	92.8000	41.7600	37.1200	26.6800	246.9999	223.2499	109.2500	156.7500	21.0000
	17.0000	167.9998	60.0000	95.9999	1.0000	2.0000	1.0000	2.0000	3.0000	3.0000
	2.0000	2.0000	0.0000	2.0000	2.0000	2.0000	2.0000	408.4999	408.4999	218.4999
	109.2500	71.2500	522.4996	541.4997	294.4999	99.7500	76.0000	512.9995	498.7499	265.9999
	199.4999	85.5000	6.0000	1.0000	1.0831	3.4783	1.1064	1.2128	3.7340	.6970
	1.2353	.5714								
OTU 29	8.0000	.0673	.0538	.0565	.0592	.0606	.0592	.0592	.0579	.0875
(15)	.0875	.1009	.1050	.1454	2.0000	2.0000	1.0000	2.0000	3.0000	4.0000
	119.4799	95.1199	37.1200	48.7200	33.6400	299.2499	265.9999	123.5000	189.9999	28.0000
	16.0000	191.9998	85.2000	113.9999	1.0000	2.0000	2.0000	2.0000	1.0000	1.0000
	0.0000	2.0000	1.0000	1.0000	1.0000	1.0000	1.0000	546.2497	546.2497	327.7499
	142.4999	95.0000	664.9999	664.9999	356.2500	166.2500	85.5000	674.4995	674.4995	356.2500
	204.2499	109.2500	6.0000	0.0000	1.1108	2.8276	1.1250	1.1562	3.3170	.6500
	1.2500	.5938								
OTU 30	7.0000	.0671	.0534	.0549	.0579	.0610	.0595	.0579	.0579	.0838
(16)	.0838	.1067	.1067	.1494	3.0000	2.0000	1.0000	2.0000	5.0000	3.0000
	115.9999	97.4399	34.8000	41.7600	29.0000	261.2500	275.4999	128.2499	142.4999	16.0000
	15.0000	173.9999	76.8000	107.9999	3.0000	3.0000	4.0000	3.0000	4.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	3.0000	3.0000	2.0000	451.2498	475.0000	261.2500
	128.2499	80.7500	555.7498	593.7500	313.5000	151.9999	85.5000	598.4996	593.7500	313.5000
	166.2500	104.5000	6.0000	0.0000	1.1299	3.3600	.9483	1.0776	2.8276	.9000
	1.0667	.6207								
OTU 31	7.0000	.0703	.0523	.0583	.0643	.0628	.0568	.0568	.0538	.0837
(17)	.0897	.1016	.0972	.1525	3.0000	1.0000	2.0000	1.0000	2.0000	5.0000
	121.7999	87.0000	38.2800	40.6000	20.8800	289.7499	237.5000	132.9999	161.4999	19.0000
	17.0000	177.5999	73.2000	101.9999	3.0000	3.0000	4.0000	4.0000	4.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	3.0000	2.0000	498.7499	484.4999	261.2500
	113.9999	71.2500	631.7495	584.2496	337.2498	128.2499	80.7500	617.4997	598.4996	313.5000
	180.4999	104.5000	5.0000	0.0000	1.1038	4.1667	1.2200	1.2400	3.3450	.8235
	1.1176	.5743								
OTU 32	8.0000	.0685	.0527	.0553	.0593	.0685	.0632	.0659	.0593	.0777
(18)	.0856	.0962	.1028	.1449	3.0000	2.0000	1.0000	3.0000	4.0000	5.0000
	115.9999	115.9999	40.6000	52.2000	34.8000	308.7499	299.2499	142.4999	185.2499	19.0000
	15.0000	191.9998	67.5999	113.9999	2.0000	2.0000	2.0000	3.0000	3.0000	1.0000
	2.0000	3.0000	0.0000	2.0000	2.0000	2.0000	2.0000	522.4996	546.2497	294.4999
	166.2500	104.5000	712.4998	688.7496	356.2500	175.7500	113.9999	702.9996	712.4998	370.4998
	213.7499	118.7500	7.0000	0.0000	1.0294	3.3333	1.0317	1.1111	3.0119	.7692
	1.2667	.5938								
OTU 33	7.0000	.0719	.0523	.0588	.0601	.0654	.0601	.0588	.0588	.0797
(19)	.0902	.0980	.1033	.1425	2.0000	1.0000	1.0000	2.0000	4.0000	5.0000
	121.7999	105.5599	38.2800	52.2000	30.1600	313.5000	284.9999	142.4999	147.2499	15.0000
	16.0000	176.3999	91.1999	122.3999	2.0000	2.0000	2.0000	3.0000	3.0000	1.0000
	2.0000	3.0000	0.0000	2.0000	2.0000	3.0000	2.0000	546.2497	555.7498	308.7499
	161.4999	104.5000	712.4998	712.4998	379.9999	175.7500	109.2500	721.9995	736.2499	379.9999
	199.4999	128.2499	5.0000	0.0000	1.0565	3.5000	1.1000	1.1417	3.1875	.9677
	.9375	.6939								

OTU 34	8.0000	.0720	.0533	.0576	.0576	.0648	.0634	.0591	.0605	.0850
(20)	.0850	.0922	.1023	.1470	2.0000	1.0000	1.0000	1.0000	3.0000	3.0000
	110.1999	91.6400	44.0800	46.4000	29.0000	289.7499	275.4999	113.9999	189.9999	21.0000
	17.0000	167.9998	90.0000	119.9999	2.0000	2.0000	2.0000	4.0000	3.0000	1.0000
	2.0000	3.0000	0.0000	2.0000	2.0000	2.0000	1.0000	522.4996	546.2497	270.7500
	161.4999	104.5000	693.4995	688.7496	365.7499	175.7500	109.2500	688.7496	688.7496	356.2500
	218.4999	118.7500	6.0000	0.0000	1.0472	3.1600	1.0517	1.0862	2.9914	.6000
	1.2353	.7143								
OTU 35	9.0000	.0693	.0539	.0585	.0600	.0616	.0600	.0600	.0554	.0832
	.0812	.0970	.1016	.1587	2.0000	1.0000	1.0000	2.0000	4.0000	4.0000
	102.0800	80.0400	34.8000	44.0800	31.3200	284.9999	246.9999	118.7500	180.4999	20.0000
	16.0000	175.1998	78.0000	105.6000	2.0000	3.0000	2.0000	3.0000	3.0000	3.0000
	2.0000	3.0000	1.0000	2.0000	2.0000	2.0000	1.0000	475.0000	475.0000	237.5000
	128.2499	90.2500	607.9998	579.4996	342.0000	151.9999	85.5000	603.2497	617.4997	318.2498
	189.9999	95.0000	6.0000	0.0000	1.0868	2.5555	1.1538	1.0817	3.1201	.6578
	.8000	.6027								
OTU 36	9.0000	.0716	.0531	.0543	.0601	.0612	.0601	.0612	.0624	.0890
	.0832	.1040	.1063	.1328	2.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	138.0399	115.9999	49.8800	54.5200	34.8000	346.7498	308.7499	128.2499	213.7499	20.0000
	16.0000	209.9998	98.4000	123.6000	2.0000	2.0000	2.0000	2.0000	3.0000	2.0000
	2.0000	3.0000	1.0000	1.0000	2.0000	2.0000	1.0000	617.4997	617.4997	322.9999
	171.0000	118.7500	793.2495	783.7497	408.4999	199.4999	113.9999	769.4997	783.7497	403.7498
	261.2500	137.7499	6.0000	0.0000	1.0644	3.3333	1.1230	1.1884	3.3280	.6000
	.8000	.5885								
OTU 37	9.0000	.0759	.0601	.0659	.0659	.0659	.0630	.0616	.0616	.0873
	.0902	.1074	.1060	.1547	2.0000	1.0000	1.0000	2.0000	3.0000	2.0000
	100.9199	98.6000	40.6000	46.4000	34.8000	308.7499	261.2500	118.7500	166.2500	19.0000
	15.0000	179.9999	81.5999	104.3999	1.0000	1.0000	2.0000	1.0000	1.0000	2.0000
	1.0000	2.0000	1.0000	1.0000	1.0000	1.0000	1.0000	522.4996	536.7498	275.4999
	156.7500	95.0000	679.2497	674.4995	356.2500	180.4999	99.7500	674.4995	688.7496	356.2500
	213.7499	113.9999	6.0000	0.0000	1.2018	2.8333	1.1818	1.1227	3.1727	.7142
	.7894	.5800								
OTU 38	6.0000	.0724	.0526	.0559	.0576	.0576	.0609	.0592	.0576	.0757
(21)	.0905	.0954	.1069	.1579	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	95.1199	78.8800	26.6800	41.7600	30.1600	275.4999	223.2499	113.9999	142.4999	19.0000
	16.0000	167.9998	68.4000	83.9999	2.0000	3.0000	3.0000	3.0000	2.0000	1.0000
	1.0000	1.0000	1.0000	2.0000	2.0000	0.0000	1.0000	441.7497	451.2498	237.5000
	137.7499	85.5000	560.4995	598.4996	322.9999	147.2499	80.7500	617.4997	607.9998	318.2498
	132.9999	71.2500	6.0000	0.0000	1.1111	2.6154	1.2340	1.1117	3.2340	.8000
	1.1875	.5000								
OTU 39	7.0000	.0714	.0510	.0569	.0612	.0641	.0583	.0612	.0583	.0802
(22)	.0889	.0918	.0991	.1574	3.0000	1.0000	1.0000	1.0000	1.0000	4.0000
	92.8000	87.0000	41.7600	41.7600	32.4800	299.2499	246.9999	123.5000	175.7500	19.0000
	16.0000	173.9999	81.5999	104.3999	1.0000	2.0000	2.0000	3.0000	2.0000	1.0000
	1.0000	1.0000	1.0000	2.0000	1.0000	1.0000	1.0000	451.2498	475.0000	246.9999
	151.9999	95.0000	593.7500	626.9998	327.7499	156.7500	104.5000	664.9999	664.9999	327.7499
	180.4999	109.2500	6.0000	0.0000	1.0725	2.6786	1.2115	1.0913	3.2981	.7027
	1.1875	.6000								
OTU 40	7.0000	.0688	.0550	.0581	.0581	.0612	.0581	.0550	.0581	.0872
(23)	.0872	.0948	.1070	.1514	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	110.1999	90.4800	32.4800	46.4000	29.0000	284.9999	256.4998	118.7500	161.4999	22.0000
	16.0000	161.9999	72.0000	95.9999	1.0000	1.0000	2.0000	1.0000	2.0000	0.0000
	1.0000	1.0000	1.0000	0.0000	1.0000	1.0000	1.0000	451.2498	498.7499	237.5000
	128.2499	95.0000	617.4997	569.9996	308.7499	151.9999	99.7500	598.4996	593.7500	294.4999
	171.0000	95.0000	5.0000	0.0000	1.1165	3.1200	1.1111	1.1157	3.0278	.7353
	1.3750	.5926								
OTU 41	7.0000	.0731	.0526	.0599	.0614	.0614	.0614	.0614	.0614	.0804
(24)	.0892	.1023	.1023	.1330	2.0000	2.0000	1.0000	2.0000	2.0000	2.0000
	98.6000	90.4800	34.8000	33.6400	30.1600	299.2499	227.9998	118.7500	166.2500	16.0000
	16.0000	165.5998	74.4000	95.9999	1.0000	1.0000	2.0000	1.0000	2.0000	1.0000
	1.0000	2.0000	1.0000	1.0000	1.0000	1.0000	1.0000	522.4996	527.2497	265.9999
	142.4999	90.2500	650.7495	641.2496	308.7499	151.9999	85.5000	641.2496	641.2496	308.7499
	189.9999	104.5000	6.0000	0.0000	1.0297	3.0000	1.3125	1.1562	3.5625	.7143
	1.0000	.5797								
OTU 42	8.0000	.0746	.0559	.0559	.0593	.0576	.0559	.0559	.0542	.0814
(25)	.0831	.1000	.1085	.1576	3.0000	1.0000	2.0000	1.0000	4.0000	4.0000
	100.9199	81.1999	34.8000	40.6000	26.6800	261.2500	237.5000	113.9999	142.4999	17.0000
	16.0000	145.1999	68.4000	83.9999	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000
	1.0000	2.0000	1.0000	1.0000	2.0000	1.0000	1.0000	408.4999	403.7498	213.7499
	109.2500	76.0000	512.9995	498.7499	261.2500	118.7500	80.7500	512.9995	531.9997	270.7500
	156.7500	95.0000	6.0000	0.0000	1.1300	3.0435	1.1000	1.1100	2.9500	.8000
	1.0625	.5785								
OTU 43	9.0000	.0722	.0562	.0562	.0578	.0578	.0562	.0578	.0562	.0803
(26)	.0883	.1043	.1124	.1445	2.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	75.4000	77.7200	30.1600	37.1200	25.5200	261.2500	223.2499	137.7499	99.7500	22.0000
	16.0000	148.7999	54.0000	95.9999	2.0000	2.0000	2.0000	3.0000	4.0000	3.0000
	2.0000	3.0000	2.0000	2.0000	3.0000	4.0000	2.0000	436.9998	427.4999	213.7499
	128.2499	66.5000	569.9996	536.7498	284.9999	128.2499	85.5000	550.9996	522.4996	270.7500
	166.2500	95.0000	6.0000	0.0000	1.1263	3.0455	1.1702	1.0106	3.3138	1.3810
	1.3750	.6452								
OTU 44	7.0000	.0719	.0497	.0548	.0565	.0565	.0548	.0565	.0548	.0856
(27)	.0942	.1027	.1147	.1473	2.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	64.9600	69.6000	29.0000	40.6000	24.3600	261.2500	204.2499	99.7500	132.9999	19.0000
	16.0000	146.3999	63.6000	95.9999	3.0000	3.0000	3.0000	4.0000	4.0000	3.0000
	3.0000	3.0000	2.0000	2.0000	3.0000	4.0000	2.0000	413.2499	417.0998	204.2499
	123.5000	80.7500	531.9997	546.2497	275.4999	137.7499	76.0000	531.9997	522.4996	275.4999
	156.7500	85.5000	6.0000	0.0000	1.1955	2.8571	1.2791	1.0233	3.3953	.7500
	1.1875	.6557								

OTU 45	8.0000	.0681	.0511	.0596	.0545	.0562	.0545	.0579	.0596	.0869
(28)	.0886	.1022	.1107	.1499	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	58.0000	69.6000	29.0000	40.6000	24.3600	251.7499	208.9999	99.7500	85.5000	13.0000
	16.0000	143.9999	61.2000	85.2000	1.0000	4.0000	4.0000	4.0000	5.0000	3.0000
	3.0000	3.0000	1.0000	1.0000	3.0000	3.0000	2.0000	417.9998	417.9998	218.4999
	118.7500	80.7500	493.9998	484.4999	275.4999	132.9999	85.5000	522.4996	522.4996	251.7499
	161.4999	95.0000	6.0000	0.0000	1.1661	2.8571	1.2045	.9659	3.3352	1.1667
	.8125	.5917								
OTU 46	8.0000	.0730	.0520	.0540	.0610	.0610	.0610	.0610	.0610	.0910
	.0910	.0890	.1010	.1430	3.0000	1.0000	1.0000	1.0000	2.0000	1.0000
	69.6000	52.2000	29.0000	40.6000	25.5200	251.7499	199.4999	104.5000	142.4999	15.0000
	15.0000	149.9999	72.0000	98.4000	2.0000	3.0000	3.0000	4.0000	4.0000	2.0000
	2.0000	3.0000	1.0000	2.0000	3.0000	2.0000	2.0000	394.2499	408.4999	237.5000
	128.2499	80.7500	508.2496	531.9997	261.2500	118.7500	76.0000	522.4996	522.4996	261.2500
	142.4999	95.0000	6.0000	0.0000	1.0699	2.0450	1.2619	.9821	3.5230	.7330
	1.0000	.6560								
OTU 47	8.0000	.0698	.0582	.0532	.0565	.0582	.0565	.0549	.0549	.0881
	.0865	.1031	.1098	.1497	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	69.6000	64.9600	31.3200	34.8000	25.5200	237.5000	213.7499	95.0000	128.2499	15.0000
	15.0000	149.9999	62.4000	92.3999	3.0000	3.0000	3.0000	4.0000	5.0000	3.0000
	3.0000	3.0000	2.0000	2.0000	4.0000	4.0000	2.0000	408.4999	451.2498	223.2499
	123.5000	76.0000	517.7497	522.4996	284.9999	142.4999	85.5000	522.4996	522.4996	261.2500
	147.2499	95.0000	6.0000	0.0000	1.1618	2.5454	1.1111	.9611	3.3380	.7407
	1.0000	.6160								
OTU 48	8.0000	.0713	.0563	.0577	.0577	.0577	.0577	.0577	.0577	.0814
	.0828	.1001	.1050	.1510	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	69.6000	55.6800	29.0000	40.6000	23.2000	227.9998	189.9999	95.0000	142.4999	18.0000
	16.0000	155.9999	66.0000	98.4000	3.0000	3.0000	3.0000	4.0000	4.0000	3.0000
	3.0000	3.0000	2.0000	2.0000	4.0000	5.0000	2.0000	342.0000	403.7498	223.2499
	123.5000	76.0000	512.9995	522.4996	275.4999	123.5000	76.0000	531.9997	522.4996	270.7500
	151.9999	95.0000	6.0000	0.0000	1.1111	2.4000	1.2000	1.0500	3.6812	.6666
	.8888	.6307								
OTU 49	8.0000	.0726	.0536	.0519	.0553	.0570	.0588	.0605	.0553	.0830
	.0865	.1072	.1072	.1505	3.0000	1.0000	0.0000	1.0000	2.0000	3.0000
	75.4000	60.3200	25.5200	40.6000	20.8800	227.9998	199.4999	85.5000	142.4999	17.0000
	15.0000	141.5999	62.4000	90.0000	2.0000	3.0000	3.0000	4.0000	4.0000	3.0000
	1.0000	3.0000	1.0000	2.0000	4.0000	4.0000	2.0000	379.9999	408.4999	204.2499
	118.7500	80.7500	508.2496	508.2496	251.7499	132.9999	90.2500	522.4996	527.2497	251.7499
	142.4999	90.2500	6.0000	0.0000	1.1486	2.8888	1.1666	1.0357	3.4404	.6330
	.8823	.6355								
OTU 50	8.0000	.0734	.0505	.0555	.0555	.0555	.0555	.0570	.0570	.0848
	.0880	.1060	.1141	.1468	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	75.4000	63.8000	32.4800	47.5600	20.8800	251.7499	204.2499	85.5000	142.4999	18.0000
	16.0000	153.5999	66.0000	97.2000	2.0000	3.0000	2.0000	3.5000	3.0000	3.0000
	3.0000	3.0000	1.0000	2.0000	3.0000	2.0000	2.0000	394.2499	394.2499	213.7499
	113.9999	76.0000	508.2496	498.7499	261.2500	123.5000	76.0000	522.4996	522.4996	261.2500
	161.4999	76.0000	6.0000	0.0000	1.1734	2.8888	1.2325	1.0988	3.5639	.0666
	.8888	.6328								
OTU 51	6.0000	.0766	.0487	.0557	.0557	.0599	.0599	.0599	.0557	.0808
(29)	.0877	.0989	.1072	.1532	2.0000	2.0000	1.0000	1.0000	2.0000	4.0000
	107.8799	88.1599	38.2800	45.2400	33.6400	299.2499	256.4998	147.2499	189.9999	17.0000
	17.0000	185.9999	82.7999	115.1999	3.0000	3.0000	2.0000	3.0000	2.0000	2.0000
	1.0000	3.0000	1.0000	1.0000	2.0000	1.0000	1.0000	522.4996	522.4996	289.7499
	151.9999	90.2500	688.7496	664.9999	389.4999	156.7500	104.5000	664.9999	664.9999	360.9998
	166.2500	104.5000	5.0000	0.0000	1.1180	2.6207	1.1667	1.1157	3.3241	.7750
	1.0000	.6194								
OTU 52	11.0000	.0750	.0560	.0580	.0600	.0620	.0600	.0630	.0620	.0909
	.0940	.1070	.1050	.1620	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	81.1999	69.6000	27.8400	38.2800	31.3200	261.2500	189.9999	76.0000	123.5000	17.0000
	16.0000	119.9999	52.8000	72.0000	1.0000	2.0000	3.0000	3.0000	0.0000	2.0000
	0.0000	1.5000	1.0000	1.0000	3.0000	2.0000	2.0000	356.2500	370.4998	194.7499
	99.7500	61.7500	475.0000	475.0000	246.9999	104.5000	71.2500	475.0000	451.2498	237.5000
	132.9999	71.2500	5.0000	0.0000	1.0075	2.1428	1.3750	1.1680	3.3250	.6150
	.9411	.6000								
OTU 53	9.0000	.0660	.0460	.0490	.0500	.0530	.0530	.0546	.0550	.0850
	.0940	.1060	.1250	.1600	3.0000	1.0000	1.0000	0.0000	3.0000	3.0000
	81.1999	63.8000	34.8000	40.6000	30.1600	246.9999	185.2499	109.2500	156.7500	17.0000
	17.0000	134.3999	60.0000	80.3999	1.0000	2.0000	3.0000	2.0000	0.0000	2.0000
	0.0000	1.5000	1.0000	1.0000	3.0000	2.0000	2.0000	403.7498	403.7498	213.7499
	109.2500	71.2500	522.4996	522.4996	265.9999	113.9999	109.2500	498.7499	498.7499	251.7499
	109.2500	80.7500	5.0000	0.0000	1.3300	2.1150	1.3333	1.2180	4.1090	.6970
	1.0000	.5980								
OTU 54	9.0000	.0730	.0580	.0590	.0580	.0560	.0580	.0580	.0560	.0820
	.0990	.0990	.0990	.1410	3.0000	1.0000	1.0000	1.0000	3.0000	3.0000
	69.6000	63.8000	30.1600	29.0000	30.1600	237.5000	180.4999	104.5000	147.2499	16.0000
	17.0000	117.5999	49.2000	72.0000	2.0000	2.0000	2.0000	2.0000	0.0000	.5000
	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	1.0000	356.2500	356.2500	189.9999
	95.0000	57.0000	436.9998	427.4999	213.7499	123.5000	76.0000	451.2498	451.2498	223.2499
	128.2499	76.0000	5.0000	0.0000	1.0990	2.1153	1.3159	1.0920	3.6320	.7096
	1.0600	.6120								
OTU 55	9.0000	.0743	.0520	.0560	.0580	.0580	.0580	.0560	.0590	.0800
	.0910	.1010	.1050	.1490	3.0000	2.0000	1.0000	1.0000	3.0000	3.0000
	87.0000	68.4400	23.2000	37.1200	25.5200	256.4998	199.4999	95.0000	166.2500	17.0000
	15.0000	125.9999	60.0000	74.4000	2.0000	2.0000	3.0000	3.0000	2.0000	2.0000
	1.0000	3.0000	1.0000	1.0000	3.0000	1.0000	1.0000	379.9999	379.9999	189.9999
	85.5000	57.0000	498.7499	475.0000	232.7499	113.9999	80.7500	475.0000	484.4999	242.2499
	118.7500	85.5000	6.0000	0.0000	1.1189	2.6818	1.2850	1.1070	3.3920	.5714
	.8823	.5904								

OTU 56	9.0000	.0740	.0480	.0530	.0540	.0590	.0580	.0620	.0610	.093
	.0880	.0990	.0990	.1520	3.0000	1.0000	1.0000	1.0000	3.0000	3.000
	87.0000	71.9200	30.1600	37.1200	33.6400	270.7500	227.9998	109.2500	161.4999	21.000
	16.0000	141.5999	56.4000	81.5999	1.0000	2.0000	2.0000	3.0000	0.0000	2.000
	0.0000	1.0000	1.0000	1.0000	3.0000	1.0000	1.0000	360.9998	408.4999	223.249
	118.7500	76.0000	531.9997	522.4996	265.9999	118.7500	80.7500	531.9997	522.4996	270.750
	137.7499	80.7500	6.0000	0.0000	1.1330	2.1379	1.1870	1.0156	3.2550	.676
	.7619	.5762								
OTU 57	9.0000	.0711	.0520	.0580	.0599	.0560	.0560	.0599	.0599	.088
	.0860	.1020	.1020	.1460	3.0000	1.0000	1.0000	1.0000	3.0000	3.000
	76.5600	53.3600	23.2000	32.4800	27.8400	237.5000	166.2500	95.0000	156.7500	17.000
	15.0000	107.9999	48.0000	66.0000	1.0000	2.0000	2.0000	3.0000	0.0000	2.000
	0.0000	2.0000	1.0000	1.0000	3.0000	1.0000	1.0000	332.5000	318.2498	166.250
	85.5000	57.0000	408.4999	413.2499	223.2499	95.0000	71.2500	403.7498	403.7498	208.999
	118.7500	71.2500	5.0000	0.0000	1.1106	1.9116	1.4280	1.1428	3.8142	.606
	.8820	.6111								
OTU 58	10.0000	.0680	.0500	.0530	.0560	.0570	.0560	.0610	.0550	.087
	.0890	.0990	.1140	.1520	4.0000	1.0000	1.0000	2.0000	2.0000	3.000
	90.4800	73.0800	32.4800	38.2800	32.4800	256.4998	199.4999	118.7500	156.7500	17.000
	16.0000	140.3999	60.0000	90.0000	1.0000	2.0000	2.0000	2.0000	0.0000	2.000
	0.0000	0.0000	1.0000	1.0000	2.0000	1.0000	1.0000	413.2499	427.4999	213.749
	118.7500	76.0000	531.9997	522.4996	265.9999	118.7500	80.7500	531.9997	522.4996	270.750
	137.7499	80.7500	5.0000	0.0000	1.1790	2.2500	1.2860	1.2023	3.9050	.757
	.9411	.6410								
OTU 59	7.0000	.0650	.0500	.0540	.0590	.0590	.0590	.0590	.0590	.085
	.0870	.1020	.1150	.1444	3.0000	1.0000	1.0000	1.0000	2.0000	3.000
	92.8000	81.1999	23.2000	32.4800	20.8800	218.4999	180.4999	95.0000	123.5000	16.000
	14.0000	140.3999	57.6000	87.5999	1.0000	4.0000	3.0000	4.0000	5.0000	3.000
	3.0000	3.0000	2.0000	2.0000	4.0000	3.0000	2.0000	389.4999	389.4999	204.249
	109.2500	71.2500	475.0000	498.7499	227.9998	109.2500	76.0000	475.0000	498.7499	237.500
	151.9999	95.0000	6.0000	0.0000	1.1474	3.8880	1.2105	1.3026	3.5460	.769
	.8750	.6239								
OTU 60	10.0000	.0680	.0480	.0490	.0610	.0590	.0590	.0570	.0590	.085
	.0870	.1007	.1064	.1615	3.0000	1.0000	1.0000	1.0000	2.0000	3.000
	71.9200	63.8000	25.5200	34.8000	20.8800	232.7499	175.7500	95.0000	118.7500	16.000
	14.0000	131.9999	56.4000	81.5999	3.0000	3.0000	3.0000	4.0000	4.0000	3.000
	3.0000	3.0000	1.0000	2.0000	4.0000	3.0000	2.0000	370.4998	365.7499	180.499
	104.5000	71.2500	475.0000	451.2498	227.9998	118.7500	76.0000	475.0000	484.4999	213.749
	142.4999	85.5000	6.0000	0.0000	1.1825	3.0555	1.3243	1.1418	3.5540	.800
	.8750	.6181								
OTU 61	7.0000	.0696	.0649	.0618	.0587	.0649	.0587	.0572	.0541	.078
(30)	.0866	.0958	.1005	.1484	2.0000	1.0000	1.0000	2.0000	3.0000	3.000
	109.0399	76.5600	29.0000	39.4400	32.4800	284.9999	246.9999	118.7500	171.0000	17.000
	15.0000	166.7999	48.0000	100.7999	1.0000	2.0000	2.0000	1.0000	3.0000	1.000
	0.0000	2.0000	1.0000	2.0000	1.0000	0.0000	1.0000	465.4998	460.7498	246.999
	128.2499	90.2500	593.7500	574.7495	332.5000	142.4999	90.2500	593.7500	569.9996	194.749
	171.0000	90.2500	6.0000	0.0000	1.0410	2.3571	1.1538	1.0529	3.1106	.694
	1.1333	.6043								
OTU 62	8.0000	.0730	.0598	.0584	.0598	.0598	.0584	.0598	.0584	.081
(31)	.0823	.0996	.1062	.1434	3.0000	2.0000	2.0000	2.0000	4.0000	5.000
	121.7999	83.5200	41.7600	41.7600	32.4800	322.9999	256.4998	132.9999	171.0000	22.000
	17.0000	188.3999	78.0000	113.9999	1.0000	3.0000	3.0000	2.0000	3.0000	0.000
	2.0000	2.0000	1.0000	1.0000	2.0000	1.0000	2.0000	484.4999	546.2497	284.999
	118.7500	71.2500	645.9995	688.7496	379.9999	142.4999	95.0000	683.9996	688.7496	365.749
	208.9999	118.7500	6.0000	0.0000	1.0518	2.5714	1.2593	1.1528	3.4861	.777
	1.2941	.6051								
OTU 63	9.0000	.0699	.0497	.0606	.0606	.0606	.0606	.0621	.0606	.085
(32)	.0854	.0963	.0963	.1522	2.0000	1.0000	2.0000	1.0000	3.0000	5.000
	134.3999	81.1999	34.8000	41.7600	29.0000	275.4999	237.5000	113.9999	151.9999	17.000
	18.0000	158.3999	72.0000	90.0000	1.0000	1.0000	2.0000	1.0000	2.0000	1.000
	1.0000	2.0000	1.0000	1.0000	1.0000	1.0000	1.0000	451.2498	489.2498	218.499
	118.7500	80.7500	579.4996	593.7500	318.2498	151.9999	80.7500	560.4995	603.2497	308.749
	151.9999	95.0000	6.0000	0.0000	1.0641	2.8000	1.1600	1.1300	3.2200	.750
	.9444	.5682								
OTU 64	7.0000	.0670	.0542	.0574	.0606	.0622	.0590	.0574	.0574	.086
(33)	.0797	.0957	.1037	.1595	3.0000	2.0000	1.0000	2.0000	3.0000	4.000
	104.3999	81.1999	35.9600	40.6000	29.0000	275.4999	237.5000	104.5000	161.4999	19.000
	17.0000	139.1999	70.8000	98.4000	2.0000	3.0000	3.0000	4.0000	3.0000	1.000
	2.0000	2.0000	2.0000	2.0000	2.0000	3.0000	2.0000	455.9997	465.4998	246.999
	128.2499	85.5000	569.9996	555.7498	318.2498	142.4999	85.5000	560.4995	593.7500	284.999
	175.7500	90.2500	6.0000	0.0000	1.1040	2.8000	1.1600	1.1300	3.1350	.647
	1.1176	.7069								
OTU 65	7.0000	.0732	.0498	.0586	.0630	.0659	.0615	.0615	.0600	.080
(34)	.0820	.0908	.1025	.1508	2.0000	1.0000	1.0000	2.0000	4.0000	3.000
	104.3999	78.8800	38.2800	42.9200	27.8400	284.9999	256.4998	19.0000	161.4999	21.000
	16.0000	175.1998	79.2000	105.6000	1.0000	3.0000	3.0000	3.0000	3.0000	1.000
	1.0000	2.0000	1.0000	1.0000	2.0000	2.0000	1.0000	475.0000	484.4999	256.499
	142.4999	95.0000	617.4997	593.7500	332.5000	147.2499	95.0000	617.4997	626.9998	308.749
	147.2499	99.7500	5.0000	0.0000	1.0267	2.8333	1.1111	1.0556	3.1620	.117
	1.3125	.6027								
OTU 66	9.0000	.0691	.0553	.0568	.0584	.0614	.0614	.0614	.0614	.082
(35)	.0860	.0968	.1029	.1459	3.0000	1.0000	1.0000	1.0000	3.0000	4.000
	104.3999	76.5600	37.1200	40.6000	30.1600	275.4999	227.9998	109.2500	156.7500	17.000
	15.0000	161.9999	73.2000	101.9999	2.0000	3.0000	3.0000	3.0000	3.0000	2.000
	2.0000	3.0000	1.0000	2.0000	3.0000	2.0000	2.0000	465.4998	451.2498	237.500
	128.2499	85.5000	593.7500	579.4996	351.5000	142.4999	90.2500	593.7500	593.7500	299.249
	180.4999	99.7500	5.0000	0.0000	1.0601	2.5385	1.2083	1.1615	3.3906	.697
	1.1333	.6296								

	[illegible]	[illegible]								
OTU 67	9.0000	.0700	.0537	.0566	.0596	.0596	.0566	.0581	.0611	.0790
(36)	.0989	.1013	.1073	.1461	3.0000	1.0000	1.0000	2.0000	2.0000	4.0000
	104.3999	80.0400	35.9600	37.1200	30.1600	294.4999	237.5000	261.2500	299.2499	17.0000
	15.0000	173.9999	74.4000	101.9999	1.0000	2.0000	3.0000	3.0000	3.0000	2.0000
	1.0000	2.0000	1.0000	1.0000	2.0000	1.0000	1.0000	475.0000	489.2498	242.2499
	123.5000	85.5000	622.2496	603.2497	313.5000	156.7500	95.0000	617.4997	622.2496	313.5000
	180.4999	104.5000	6.0000	0.0000	1.1034	2.6538	1.2400	1.1100	3.3550	.8739
	1.1333	.5862								
OTU 68	6.0000	.0751	.0591	.0607	.0639	.0607	.0623	.0575	.0671	.0847
(37)	.0879	.0879	.1022	.1310	3.0000	1.0000	1.0000	1.0000	4.0000	4.0000
	104.3999	87.0000	31.3200	40.6000	29.0000	270.7500	227.9998	109.2500	151.9999	13.0000
	14.0000	161.9999	69.6000	98.4000	1.0000	2.0000	2.0000	2.0000	3.0000	2.0000
	2.0000	3.0000	1.0000	1.0000	2.0000	2.0000	1.0000	427.4999	422.7499	227.9998
	104.5000	76.0000	603.2497	522.4996	289.7499	128.2499	90.2500	569.9996	579.4996	318.2498
	175.7500	95.0000	7.0000	0.0000	.9748	3.0000	1.1875	1.1823	3.2604	.7188
	.9286	.6074								
OTU 69	9.0000	.0697	.0530	.0545	.0576	.0621	.0591	.0606	.0606	.0833
(38)	.0879	.1015	.1061	.1439	3.0000	1.0000	1.0000	1.0000	2.0000	4.0000
	96.2799	81.1999	34.8000	41.7600	32.4800	261.2500	232.7499	118.7500	142.4999	16.0000
	17.0000	155.9999	73.2000	95.9999	2.0000	2.0000	3.0000	3.0000	3.0000	2.0000
	2.0000	3.0000	1.0000	1.0000	2.0000	1.0000	1.0000	451.2498	460.7498	237.5000
	123.5000	76.0000	560.4995	451.2498	322.9999	142.4999	80.7500	569.9996	584.2496	318.2491
	185.2499	95.0000	5.0000	0.0000	1.0952	2.5000	1.1224	1.1173	3.3673	.8333
	.9412	.6154								
OTU 70	9.0000	.0695	.0579	.0592	.0644	.0656	.0618	.0644	.0669	.0849
(39)	.0862	.0952	.0952	.1287	3.0000	1.0000	1.0000	1.0000	1.0000	2.0000
	104.3999	84.6799	38.2800	44.0800	30.1600	289.7499	251.7499	123.5000	194.7499	25.0000
	22.0000	161.9999	67.2000	101.9999	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000
	2.0000	3.0000	1.0000	1.0000	2.0000	2.0000	1.0000	512.9995	546.2497	275.4999
	147.2499	95.0000	641.2496	664.9999	360.9998	161.4999	99.7500	593.7500	641.2496	327.7499
	189.9999	104.5000	5.0000	0.0000	.9621	2.8077	1.1509	1.1038	3.6651	.6341
	1.1364	.6296								
OTU 71	9.0000	.0794	.0556	.0595	.0595	.0615	.0595	.0595	.0615	.0754
(40)	.0833	.0893	.1071	.1488	3.0000	1.0000	0.0000	1.0000	2.0000	3.0000
	76.5600	59.1600	31.3200	37.1200	20.8800	218.4999	175.7500	85.5000	128.2499	12.0000
	14.0000	129.5999	60.0000	81.5999	3.0000	3.0000	3.0000	5.0000	5.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	5.0000	2.0000	365.7499	384.7499	204.2499
	104.5000	76.0000	465.4998	484.4999	246.9999	118.7500	76.0000	475.0000	498.7499	227.9998
	142.4999	85.5000	6.0000	0.0000	1.0160	2.8333	1.2432	1.1892	3.4054	.6667
	.8571	.6296								
OTU 72	9.0000	.0745	.0577	.0577	.0577	.0596	.0540	.0596	.0559	.0782
(41)	.0857	.0931	.1099	.1564	2.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	81.1999	63.8000	32.4800	38.2800	23.2000	223.2499	185.2499	80.7500	132.9999	14.0000
	15.0000	131.9999	60.0000	74.4000	3.0000	4.0000	3.0000	4.0000	5.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	4.0000	3.0000	384.7499	379.9999	189.9999
	104.5000	71.2500	460.7498	475.0000	237.5000	118.7500	71.2500	455.9997	479.7498	237.5000
	142.4999	85.5000	6.0000	0.0000	1.0977	2.7500	1.2051	1.1923	3.4423	.6071
	.9333	.5636								
OTU 73	9.0000	.0809	.0570	.0607	.0588	.0607	.0551	.0570	.0551	.0754
(42)	.0827	.0993	.1085	.1489	2.0000	0.0000	0.0000	1.0000	2.0000	2.0000
	87.0000	64.9600	30.1600	41.7600	23.2000	237.5000	199.4999	85.5000	151.9999	16.0000
	14.0000	143.9999	64.8000	93.6000	2.0000	3.0000	3.0000	3.0000	4.0000	1.0000
	3.0000	3.0000	0.0000	1.0000	2.0000	3.0000	2.0000	403.7498	403.7498	218.4999
	99.7500	76.0000	508.2496	508.2496	261.2500	137.7499	80.7500	498.7499	522.4996	261.2500
	137.7499	76.0000	6.0000	0.0000	1.0606	2.8000	1.1905	1.1488	3.2381	.5625
	1.1429	.6500								
OTU 74	9.0000	.0811	.0570	.0592	.0636	.0570	.0614	.0592	.0570	.0768
(43)	.0746	.0921	.1009	.1601	2.0000	1.0000	0.0000	1.0000	1.0000	3.0000
	60.3200	58.0000	27.8400	33.6400	18.5600	208.9999	156.7500	85.5000	85.5000	14.0000
	14.0000	118.7999	55.2000	72.0000	3.0000	3.0000	4.0000	4.0000	4.0000	2.0000
	3.0000	3.0000	1.0000	2.0000	2.0000	4.0000	2.0000	322.9999	342.0000	171.0000
	85.5000	52.2500	403.7498	403.7498	199.4999	80.7500	61.7500	398.9998	441.7497	204.2499
	123.5000	71.2500	6.0000	0.0000	1.0177	3.1250	1.3333	1.1742	3.4545	1.0000
	1.0000	.6061								
OTU 75	8.0000	.0784	.0510	.0588	.0608	.0588	.0549	.0588	.0569	.0804
(44)	.0804	.0980	.1118	.1510	2.0000	1.0000	0.0000	1.0000	2.0000	3.0000
	81.1999	67.2800	32.4800	39.4400	22.0400	223.2499	171.0000	85.5000	118.7500	13.0000
	13.0000	125.9999	60.0000	83.9999	3.0000	3.0000	4.0000	4.0000	4.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	4.0000	2.0000	370.4998	379.9999	199.4999
	109.2500	71.2500	465.4998	484.4999	246.9999	118.7500	71.2500	38.0000	498.7499	237.5000
	142.4999	85.5000	6.0000	0.0000	1.0902	3.0526	1.3056	1.3194	3.5417	.7200
	1.0000	.6667								
OTU 76	9.0000	.0751	.0558	.0601	.0601	.0601	.0579	.0579	.0601	.0751
(45)	.0837	.1073	.1073	.1395	2.0000	0.0000	0.0000	2.0000	1.0000	2.0000
	67.2800	52.2000	25.5200	34.8000	25.5200	204.2499	151.9999	85.5000	104.5000	13.0000
	13.0000	119.9999	54.0000	74.4000	3.0000	3.0000	4.0000	4.0000	5.0000	2.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	4.0000	2.0000	332.5000	342.0000	171.0000
	80.7500	61.7500	417.9998	417.9998	218.4999	109.2500	66.5000	403.7498	427.4999	189.9999
	118.7500	66.5000	6.0000	0.0000	1.0529	2.8455	1.3437	1.2109	3.6406	.8182
	1.0000	.6200								
OTU 77	9.0000	.0640	.0548	.0585	.0622	.0622	.0567	.0585	.0548	.0800
(46)	.0823	.1042	.1133	.1481	2.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	87.0000	68.4400	34.8000	33.6400	20.8800	223.2499	189.9999	99.7500	142.4999	12.0000
	13.0000	137.9999	63.6000	86.3999	3.0000	3.0000	4.0000	3.0000	3.0000	1.0000
	2.0000	3.0000	0.0000	2.0000	2.0000	4.0000	2.0000	394.2499	408.4999	204.2499
	104.5000	76.0000	503.4996	512.9995	261.2500	118.7500	80.7500	503.4996	527.2497	265.9999
	147.2499	80.7500	6.0000	0.0000	1.1202	3.2778	1.1750	1.2063	3.4187	.7000
	.9231	.6261								

OTU 78	11.0000	.0778	.0599	.0599	.0599	.0619	.0579	.0579	.0539	.0739
(47)	.0798	.1016	.1098	.1457	3.0000	1.0000	1.0000	1.0000	2.0000	3.0000
	75.4000	61.4800	32.4800	34.8000	22.0400	223.2499	171.0000	132.9999	104.5000	14.0000
	13.0000	137.9999	58.8000	83.9999	3.0000	2.0000	4.0000	2.0000	5.0000	3.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	3.0000	2.0000	360.9998	389.4999	194.7499
	104.5000	66.5000	451.2498	475.0000	251.7499	113.9999	71.2500	451.2498	484.4999	227.9998
	137.7499	80.7500	6.0000	0.0000	1.0449	2.7895	1.3056	1.2222	3.4792	1.2727
	1.0769	.6087								
OTU 79	11.0000	.0822	.0600	.0600	.0622	.0622	.0600	.0578	.0556	.0689
(48)	.0800	.0978	.1000	.1533	2.0000	0.0000	0.0000	1.0000	2.0000	2.0000
	67.2800	52.2000	29.0000	33.6400	18.5600	204.2499	151.9999	66.5000	99.7500	16.0000
	15.0000	109.2000	48.0000	78.0000	3.0000	4.0000	4.0000	5.0000	5.0000	3.0000
	3.0000	3.0000	0.0000	2.0000	3.0000	3.0000	2.0000	322.9999	327.7499	151.9999
	90.2500	61.7500	389.4999	394.2499	189.9999	90.2500	66.5000	403.7498	403.7498	189.9999
	118.7500	71.2500	6.0000	0.0000	1.0000	2.8125	1.3437	1.2266	3.5156	.6667
	1.0667	.7143								
OTU 80	10.0000	.0801	.0584	.0563	.0606	.0628	.0628	.0584	.0563	.0736
(49)	.0779	.0909	.1082	.1537	2.0000	1.0000	1.0000	1.0000	2.0000	2.0000
	58.0000	49.8800	30.1600	30.1600	19.7200	204.2499	147.2499	76.0000	123.5000	15.0000
	13.0000	119.9999	51.6000	78.0000	3.0000	4.0000	3.0000	4.0000	5.0000	3.0000
	3.0000	3.0000	0.0000	2.0000	2.0000	3.0000	2.0000	318.2498	327.7499	171.0000
	80.7500	61.7500	398.9998	394.2499	204.2499	99.7500	61.7500	398.9998	413.2499	208.9999
	118.7500	71.2500	5.0000	0.0000	1.0175	2.5294	1.3871	1.1694	3.7258	.6154
	1.1538	.6500								
OTU 81	8.0000	.0661	.0573	.0587	.0587	.0602	.0587	.0587	.0587	.0822
(50)	.0896	.0999	.1043	.1468	3.0000	2.0000	2.0000	2.0000	3.0000	5.0000
	115.9999	88.1599	30.1600	40.6000	32.4800	294.4999	261.2500	123.5000	166.2500	16.0000
	15.0000	177.5999	75.6000	101.9999	2.0000	4.0000	3.0000	3.0000	3.0000	1.0000
	2.0000	3.0000	0.0000	2.0000	3.0000	4.0000	3.0000	479.7498	479.7498	261.2500
	123.5000	95.0000	593.7500	593.7500	332.5000	151.9999	99.7500	593.7500	593.7500	322.9999
	199.4999	109.2500	5.0000	0.0000	1.0954	2.7143	1.1273	1.0773	3.0955	.7429
	1.0667	.5743								
OTU 82	9.0000	.0745	.0542	.0569	.0569	.0610	.0569	.0569	.0596	.0881
(51)	.0867	.1016	.1043	.1423	2.0000	1.0000	1.0000	1.0000	4.0000	5.0000
	121.7999	98.6000	37.1200	52.2000	30.1600	308.7499	308.7499	118.7500	175.7500	17.0000
	17.0000	170.3999	78.0000	107.9999	3.0000	4.0000	3.0000	4.0000	5.0000	2.0000
	3.0000	3.0000	2.0000	2.0000	3.0000	4.0000	3.0000	531.9997	522.4996	356.2500
	147.2499	95.0000	664.9999	641.2496	356.2500	166.2500	104.5000	641.2496	626.9998	356.2500
	189.9999	113.9999	5.0000	0.0000	1.0966	3.2692	1.0000	1.0269	2.8385	.6757
	1.0000	.6338								
OTU 83	7.0000	.0701	.0472	.0580	.0593	.0606	.0566	.0566	.0580	.0809
(52)	.0863	.0997	.1119	.1550	3.0000	2.0000	1.0000	1.0000	3.0000	4.0000
	136.8799	106.7200	38.2800	54.5200	30.1600	313.5000	308.7499	128.2499	156.7500	20.0000
	19.0000	189.5998	81.5999	117.5999	3.0000	3.0000	4.0000	4.0000	5.0000	2.0000
	3.0000	3.0000	2.0000	2.0000	3.0000	4.0000	2.0000	546.2497	536.7498	289.7499
	156.7500	99.7500	698.2498	674.4995	379.9999	166.2500	104.5000	683.9996	693.4995	379.9999
	194.7499	118.7500	7.0000	0.0000	1.1445	3.5385	1.0154	1.1154	2.8538	.8182
	1.0526	.6203								
OTU 84	8.0000	.0662	.0611	.0561	.0599	.0599	.0586	.0573	.0573	.0854
(53)	.0917	.0981	.0994	.1490	3.0000	1.0000	1.0000	1.0000	3.0000	3.0000
	141.5199	110.1999	41.7600	55.6800	34.8000	327.7499	308.7499	142.4999	204.2499	19.0000
	18.0000	185.9999	86.3999	119.9999	3.0000	3.0000	2.0000	2.0000	3.0000	2.0000
	2.0000	3.0000	1.0000	1.0000	1.0000	3.0000	2.0000	555.7498	560.4995	308.7499
	151.9999	104.5000	712.4998	683.9996	403.7498	166.2500	104.5000	712.4998	721.9995	389.4999
	213.7499	118.7500	6.0000	0.0000	1.0989	3.1667	1.0615	1.1577	3.0192	.6977
	1.0556	.6452								

****

SUMMARY OF DATA USED IN CANONICAL VARIATE ANALYSIS
SEVEN TAXA BASED ON LARGER SAMPLES

DATA FOR C. NEWSTEADI

4.0000	.0768	1.0000	1.0000	4.0000	2.0000	3.0000	1.0000	2.5833	.5700
4.0000	.0720	1.0000	1.0000	4.0000	2.0000	3.0000	1.0376	2.5385	.5439
3.0000	.0803	1.0000	1.0000	3.0000	2.0000	4.0000	1.0837	1.6667	.5275
3.0000	.0624	1.0000	1.0000	2.0000	1.0000	3.0000	1.1025	2.1429	.5636
5.0000	.0734	0.0000	1.0000	3.0000	2.0000	4.0000	1.0639	2.2837	.6296
5.0000	.0778	1.0000	1.0000	3.0000	1.0000	4.0000	1.0156	2.1428	.5555
7.0000	.0773	1.0000	1.0000	3.0000	1.0000	4.0000	1.0842	2.4285	.6240
7.0000	.0716	1.0000	1.0000	3.0000	1.0000	4.0000	1.0380	2.2812	.5600
5.0000	.0728	1.0000	1.0000	4.0000	1.0000	5.0000	1.0497	2.2222	.5636
5.0000	.0772	1.0000	1.0000	3.0000	1.0000	5.0000	1.0155	2.2142	.5803
5.0000	.0771	1.0000	1.0000	4.0000	1.0000	5.0000	1.0896	2.6000	.5600

SUMMARY FOR C. NEWSTEADI

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
4.8182	.0744	.9091	1.0000	3.2727	1.3636	4.0000	1.0528	2.2822	.5707
3.0000	.0624	0.0000	1.0000	2.0000	1.0000	3.0000	1.0000	1.6667	.5275
7.0000	.0803	1.0000	1.0000	4.0000	2.0000	5.0000	1.1025	2.6000	.6296
1.3280	.0049	.3015	0.0000	.6467	.5045	.7746	.0345	.2655	.0309
.4004	.0015	.0909	0.0000	.1950	.1521	.2335	.0104	.0800	.0093

CORRELATION MATRIX

```
 1.00
  .19 1.00
 -.05  .07 1.00
 0.00 0.00 0.00 0.00
  .06  .41  .14 0.00 1.00
 -.49  .20 -.42 0.00  .28 1.00
  .39  .42  .00 0.00  .20 -.51 1.00
 -.13 -.30 -.11 0.00 -.38 -.15  .05 1.00
  .35 -.09 -.00 0.00  .56 -.04 -.11 -.23 1.00
  .52 -.02 -.63 0.00 -.19 -.08  .11  .10  .33 1.00
```

DATA FOR C. GRISESCENS

4.0000	.0705	1.0000	1.0000	4.0000	0.0000	5.0000	1.1228	4.0000	.6533
4.0000	.0711	1.0000	1.0000	4.0000	0.0000	4.0000	1.0965	4.1304	.6452
4.0000	.0709	1.0000	1.0000	4.0000	0.0000	4.0000	1.1585	3.7692	.6429
4.0000	.0618	1.0000	1.0000	4.0000	0.0000	5.0000	1.2342	4.5238	.6536
6.0000	.0675	2.0000	2.0000	4.0000	0.0000	5.0000	1.1629	3.8846	.6774
7.0000	.0779	1.0000	1.0000	4.0000	0.0000	5.0000	1.0778	4.4545	.6571
7.0000	.0728	1.0000	1.0000	4.0000	0.0000	5.0000	1.1054	4.3181	.6285
7.0000	.0695	1.0000	1.0000	2.0000	0.0000	5.0000	1.1750	3.5185	.6250
5.0000	.0780	1.0000	1.0000	3.0000	0.0000	5.0000	1.1168	3.2857	.6500
5.0000	.0715	1.0000	1.0000	4.0000	0.0000	5.0000	1.1250	4.3500	.6456
5.0000	.0711	2.0000	1.0000	3.0000	0.0000	3.0000	1.1346	3.7200	.6500
5.0000	.0609	1.0000	1.0000	3.0000	1.0000	5.0000	1.1859	2.0571	.5862
7.0000	.0671	1.0000	2.0000	4.0000	0.0000	3.0000	1.1299	3.3600	.6207
7.0000	.0703	2.0000	1.0000	4.0000	0.0000	3.0000	1.1038	4.1667	.5743

SUMMARY FOR C. GRISESCENS

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
5.5000	.0701	1.2143	1.1429	3.6429	.0714	4.4286	1.1378	3.8242	.6376
4.0000	.0609	1.0000	1.0000	2.0000	0.0000	3.0000	1.0778	2.0571	.5743
7.0000	.0780	2.0000	2.0000	4.0000	1.0000	5.0000	1.2342	4.5238	.6774
1.2860	.0048	.4258	.3631	.6333	.2673	.8516	.0416	.6438	.0286
.3437	.0013	.1138	.0971	.1693	.0714	.2276	.0111	.1721	.0076

CORRELATION MATRIX

```
 1.00
  .22 1.00
  .21 -.05 1.00
  .33 -.24  .28 1.00
 -.14  .03  .02  .24 1.00
 -.11 -.54 -.14 -.11 -.29 1.00
 -.14  .03 -.48 -.21 -.12  .19 1.00
 -.33 -.80 -.85  .09 -.29  .33  .24 1.00
  .03  .37  .08 -.13  .57 -.79  .03 -.29 1.00
 -.35  .34 -.02  .17  .15 -.52  .34 -.00  .41 1.00
```

DATA FOR C. DELTA

8.0000	.0673	1.0000	2.0000	2.0000	1.0000	1.0000	1.1108	2.8276	.5938
8.0000	.0685	1.0000	3.0000	2.0000	0.0000	2.0000	1.0294	3.3333	.5938
7.0000	.0719	1.0000	2.0000	2.0000	0.0000	3.0000	1.0565	3.5000	.6939
8.0000	.0720	1.0000	1.0000	2.0000	0.0000	2.0000	1.0472	3.1600	.7143
9.0000	.0693	1.0000	2.0000	2.0000	1.0000	2.0000	1.0868	2.5555	.6027
9.0000	.0716	1.0000	1.0000	2.0000	1.0000	2.0000	1.0644	3.3333	.5885
9.0000	.0759	1.0000	2.0000	2.0000	1.0000	1.0000	1.2018	2.8333	.5800
9.0000	.0640	1.0000	1.0000	4.0000	0.0000	4.0000	1.1202	3.2778	.6261

SUMMARY FOR C. DELTA

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
8.3750	.0701	1.0000	1.7500	2.2500	.5000	2.1250	1.0896	3.1026	.6241
7.0000	.0640	1.0000	1.0000	2.0000	0.0000	1.0000	1.0294	2.5555	.5800
9.0000	.0759	1.0000	3.0000	4.0000	1.0000	4.0000	1.2018	3.5000	.7143
.7440	.0036	0.0000	.7071	.7071	.5345	.9910	.0549	.3264	.0514
.2631	.0013	0.0000	.2500	.2500	.1890	.3504	.0194	.1154	.0182

CORRELATION MATRIX

```
 1.00
-.07 1.00
0.00 0.00 0.00
-.34  .06 0.00 1.00
 .34 -.68 0.00 -.43 1.00
 .54  .29 0.00 -.00 -.38 1.00
-.07 -.53 0.00 -.36  .76 -.67 1.00
 .52  .23 0.00 -.09  .22  .51 -.28 1.00
-.48 -.08 0.00 -.14  .22 -.70  .57 -.51 1.00
-.63  .10 0.00 -.34  .02 -.68  .42 -.45  .43 1.00
```

DATA FOR C. PUNCTATUS

6.0000	.0724	1.0000	1.0000	3.0000	1.0000	0.0000	1.1111	2.6154	.5000
7.0000	.0714	1.0000	1.0000	2.0000	1.0000	1.0000	1.0725	2.6786	.6000
7.0000	.0688	1.0000	1.0000	2.0000	1.0000	1.0000	1.1165	3.1200	.5926
7.0000	.0731	1.0000	2.0000	2.0000	1.0000	1.0000	1.0297	3.0000	.5797
8.0000	.0746	2.0000	1.0000	1.0000	1.0000	1.0000	1.1300	3.0435	.5785
9.0000	.0722	1.0000	1.0000	2.0000	2.0000	4.0000	1.1263	3.0455	.6452
7.0000	.0719	1.0000	1.0000	3.0000	2.0000	4.0000	1.1955	2.8571	.6557
8.0000	.0681	1.0000	1.0000	4.0000	1.0000	3.0000	1.1661	2.8571	.5917
8.0000	.0730	1.0000	1.0000	3.0000	1.0000	2.0000	1.0699	2.0450	.6560
8.0000	.0698	1.0000	1.0000	3.0000	2.0000	4.0000	1.1618	2.5454	.6160
8.0000	.0713	1.0000	1.0000	3.0000	2.0000	5.0000	1.1111	2.4000	.6307
8.0000	.0726	0.0000	1.0000	3.0000	1.0000	4.0000	1.1486	2.8888	.8355
8.0000	.0734	1.0000	1.0000	2.0000	1.0000	2.0000	1.1734	2.8888	.6328
6.0000	.0766	1.0000	1.0000	2.0000	1.0000	1.0000	1.1180	2.6207	.6194
8.0000	.0671	1.0000	1.0000	2.0000	1.0000	2.0000	1.1912	2.9545	.6296
9.0000	.0666	1.0000	1.0000	2.0000	1.0000	3.0000	1.0582	2.8261	.5913
9.0000	.0679	1.0000	1.0000	3.0000	1.0000	4.0000	1.1341	2.9524	.6019
5.0000	.0763	1.0000	2.0000	2.0000	1.0000	.5000	1.0681	3.0500	.5906
7.0000	.0733	1.0000	1.0000	2.0000	1.0000	2.0000	1.1773	2.9565	.5770
5.0000	.0720	1.0000	1.0000	3.0000	1.0000	2.0000	1.1120	2.7273	.5784
7.0000	.0684	1.0000	1.0000	2.0000	1.0000	1.5000	1.1296	2.5000	.6000
7.0000	.0714	1.0000	1.0000	3.0000	1.0000	3.5000	1.1374	2.6957	.5810
8.0000	.0722	1.0000	1.0000	2.0000	1.0000	3.0000	1.1434	2.3913	.6087
9.0000	.0756	1.0000	1.0000	2.5000	1.0000	1.5000	1.0951	2.6000	.5702
9.0000	.0755	0.0000	1.0000	2.5000	1.0000	2.0000	1.0366	2.7083	.5826
9.0000	.0713	1.0000	1.0000	3.0000	1.5000	3.0000	1.0116	2.8571	.5714
9.0000	.0692	1.0000	1.0000	3.0000	1.0000	1.5000	1.0852	2.9286	.5797
7.0000	.0726	1.0000	1.0000	3.5000	1.0000	2.0000	1.1274	2.4800	.5556
8.0000	.0735	1.0000	1.0000	2.0000	1.0000	1.0000	1.1103	2.6667	.5556
7.0000	.0716	1.0000	1.0000	3.0000	1.0000	1.5000	1.0742	2.9545	.6087
7.0000	.0675	1.0000	1.0000	2.5000	1.0000	1.5000	1.1579	2.6957	.5575
7.0000	.0760	1.0000	1.0000	2.5000	1.0000	1.0000	1.0627	2.8800	.6261
7.0000	.0761	1.0000	1.0000	2.5000	1.0000	1.5000	1.0849	2.7083	.5652
8.0000	.0733	1.0000	1.0000	3.0000	1.0000	2.0000	1.2140	3.2727	.6000
9.0000	.0743	1.0000	1.0000	3.0000	1.0000	2.5000	1.1489	2.6923	.6000
8.0000	.0729	1.0000	1.0000	2.0000	2.0000	4.0000	1.1429	2.6154	.6299
9.0000	.0712	2.0000	1.0000	2.5000	1.0000	2.0000	1.1418	3.1304	.6161
8.0000	.0721	1.0000	2.0000	1.0000	1.0000	1.0000	1.1360	3.9000	.6143
11.0000	.0779	1.0000	1.0000	1.0000	.5000	0.0000	.9578	2.0000	.5827
11.0000	.1142	1.0000	2.0000	1.0000	.5000	0.0000	.8585	2.3333	.5893

SUMMARY FOR C. PUNCTATUS

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
7.8250	.0732	1.0000	1.1000	2.4125	1.1125	2.0625	1.1082	2.7771	.5974
5.0000	.0666	0.0000	1.0000	1.0000	.5000	0.0000	.8585	2.0000	.5000
11.0000	.1142	2.0000	2.0000	4.0000	2.0000	5.0000	1.2140	3.9000	.6560
1.2788	.0072	.3203	.3038	.6969	.3667	1.2719	.0663	.3294	.0309
.2022	.0011	.0506	.0480	.1102	.0580	.2011	.0105	.0521	.0049

CORRELATION MATRIX

1.00									
.36	1.00								
.00	-.03	1.00							
-.02	.50	.00	1.00						
-.25	-.40	-.23	-.44	1.00					
-.09	-.31	.00	-.22	.27	1.00				
.11	-.38	-.19	-.38	.48	.73	1.00			
-.40	-.67	.10	-.43	.37	.37	.45	1.00		
-.19	-.27	.14	.30	-.10	.05	-.01	.38	1.00	
.14	-.06	-.06	-.04	-.04	.43	.50	.23	.07	1.00

DATA FOR SP. A

11.0000	.0750	1.0000	1.0000	3.0000	1.0000	2.0000	1.0075	2.1428	.6000
9.0000	.0660	1.0000	0.0000	3.0000	1.0000	2.0000	1.3300	2.1150	.5980
9.0000	.0730	1.0000	1.0000	2.0000	0.0000	0.0000	1.0990	2.1153	.6120
9.0000	.0743	1.0000	1.0000	3.0000	1.0000	1.0000	1.1189	2.6818	.5904
9.0000	.0740	1.0000	1.0000	2.0000	1.0000	1.0000	1.1330	2.1379	.5762
9.0000	.0711	1.0000	1.0000	2.0000	1.0000	1.0000	1.1106	1.9116	.6111
10.0000	.0680	1.0000	2.0000	2.0000	1.0000	1.0000	1.1790	2.2500	.6410

SUMMARY FOR SP. A

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
9.4286	.0716	1.0000	1.0000	2.4286	.8571	1.1429	1.1397	2.1935	.6041
9.0000	.0660	1.0000	0.0000	2.0000	0.0000	0.0000	1.0075	1.9116	.5762
11.0000	.0750	1.0000	2.0000	3.0000	1.0000	2.0000	1.3300	2.6818	.6410
.7868	.0034	0.0000	.5774	.5345	.3780	.6901	.0985	.2377	.0204
.2974	.0013	0.0000	.2182	.2020	.1429	.2608	.0372	.0899	.0077

CORRELATION MATRIX

1.00									
.19	1.00								
0.00	0.00	0.00							
.37	.17	0.00	1.00						
.28	.04	0.00	-.54	1.00					
.24	-.18	0.00	.00	.35	1.00				
.48	-.25	0.00	-.42	.71	.73	1.00			
-.48	-.87	0.00	-.44	.12	.18	.24	1.00		
-.04	.26	0.00	.16	.47	.15	-.05	-.03	1.00	
.30	-.50	0.00	.61	-.37	-.17	-.21	.06	-.17	1.00

DATA FOR C. PULICARIS

7.0000	.0680	1.0000	1.0000	3.0000	2.0000	3.0000	1.1474	3.8880	.6239
10.0000	.0680	1.0000	1.0000	3.0000	1.0000	3.0000	1.1825	3.0555	.6181
7.0000	.0696	1.0000	2.0000	2.0000	1.0000	0.0000	1.0410	2.3571	.6043
8.0000	.0730	2.0000	2.0000	3.0000	1.0000	1.0000	1.0518	2.5714	.6051
9.0000	.0699	2.0000	1.0000	2.0000	1.0000	1.0000	1.0641	2.8000	.5682
7.0000	.0670	1.0000	2.0000	3.0000	2.0000	3.0000	1.1040	2.8000	.7069
7.0000	.0732	1.0000	2.0000	3.0000	1.0000	2.0000	1.0267	2.8333	.6027
9.0000	.0691	1.0000	1.0000	3.0000	1.0000	2.0000	1.0601	2.5385	.6296
9.0000	.0700	1.0000	2.0000	3.0000	1.0000	1.0000	1.1034	2.6538	.5862
6.0000	.0751	1.0000	1.0000	2.0000	1.0000	2.0000	.9748	3.0000	.6074
9.0000	.0697	1.0000	1.0000	3.0000	1.0000	1.0000	1.0952	2.5000	.6154
9.0000	.0695	1.0000	1.0000	2.0000	1.0000	2.0000	.9621	2.8077	.6296
9.0000	.0712	1.0000	1.0000	1.0000	1.0000	.5000	.9859	2.9167	.6071
7.0000	.0604	1.0000	1.0000	2.0000	1.0000	1.0000	1.1609	3.0476	.5763
5.0000	.0657	1.0000	1.0000	2.0000	1.0000	.5000	1.1443	2.8846	.5715
8.0000	.0747	1.0000	1.0000	2.0000	1.0000	.5000	.9929	2.6400	.6000
9.0000	.0709	1.0000	1.0000	2.0000	1.0000	1.0000	.9859	2.0312	.6000
8.0000	.0793	1.0000	1.0000	2.0000	1.0000	1.0000	1.0065	2.8333	.6000
8.0000	.0779	2.0000	2.0000	1.0000	1.0000	1.0000	1.0197	2.8800	.5600
8.0000	.0719	1.0000	2.0000	1.0000	.5000	1.0000	1.0816	2.5385	.5833
9.0000	.0748	1.0000	2.0000	1.0000	1.0000	1.0000	1.0338	3.2609	.5500
11.0000	.0779	2.0000	1.0000	1.0000	1.0000	.5000	.9727	2.3226	.5760
7.0000	.0799	1.0000	1.0000	3.0000	1.0000	4.0000	1.0458	2.3333	.6000
7.0000	.0682	1.0000	1.0000	3.0000	1.0000	3.0000	1.1504	2.6316	.6083
9.0000	.0749	1.0000	1.0000	3.5000	1.5000	4.0000	1.0855	2.5909	.6296
10.0000	.0726	1.0000	1.0000	4.0000	1.0000	4.0000	.9558	1.9062	.5714
7.0000	.0641	1.0000	2.0000	2.5000	2.0000	2.0000	1.1118	2.0833	.6302
9.0000	.0721	1.0000	1.0000	2.0000	1.0000	2.0000	1.0764	2.6250	.6232
5.0000	.0795	2.0000	2.0000	2.0000	1.0000	1.5000	1.0278	2.8125	.6125
5.0000	.0706	1.0000	1.0000	4.0000	1.5000	4.0000	1.0417	3.3182	.6015
7.0000	.0761	1.0000	1.0000	2.0000	1.0000	1.5000	1.0531	2.7200	.5926
9.0000	.0785	1.0000	1.0000	2.5000	1.0000	1.0000	1.0740	2.7586	.6296
8.0000	.0737	2.0000	2.0000	3.0000	2.0000	4.0000	1.0543	2.5000	.6061

SUMMARY FOR C. PULICARIS

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
7.9394	.0719	1.1818	1.3333	2.3788	1.1364	1.8182	1.0568	2.7103	.6038
5.0000	.0604	1.0000	1.0000	1.0000	.5000	0.0000	.9558	1.9062	.5500
11.0000	.0799	2.0000	2.0000	4.0000	2.0000	4.0000	1.1825	3.8880	.7069
1.4564	.0047	.3917	.4787	.8294	.3595	1.2172	.0603	.3819	.0284
.2535	.0008	.0682	.0833	.1444	.0626	.2119	.0105	.0665	.0049

CORRELATION MATRIX

1.00									
.13	1.00								
.07	.34	1.00							
-.19	.05	.33	1.00						
-.12	-.21	-.22	-.13	1.00					
-.22	-.31	.04	.18	.42	1.00				
-.10	-.02	-.13	-.13	.74	.52	1.00			
-.21	-.60	-.20	.03	.27	.27	.15	1.00		
-.33	-.17	-.08	-.09	-.07	.17	.03	.32	1.00	
-.10	-.20	-.27	.01	.40	.53	.34	.21	.01	1.00

DATA FOR C. IMPUNCTATUS

9.0000	.0794	0.0000	1.0000	3.0000	0.0000	5.0000	1.0160	2.8333	.6296
9.0000	.0745	1.0000	1.0000	3.0000	0.0000	4.0000	1.0977	2.7500	.5636
9.0000	.0809	0.0000	1.0000	3.0000	0.0000	3.0000	1.0606	2.8000	.6500
9.0000	.0811	0.0000	1.0000	4.0000	1.0000	4.0000	1.0177	3.1250	.6061
8.0000	.0784	0.0000	1.0000	4.0000	0.0000	4.0000	1.0902	3.0526	.6667
9.0000	.0751	0.0000	2.0000	4.0000	0.0000	4.0000	1.0529	2.0455	.6200
11.0000	.0778	1.0000	1.0000	4.0000	0.0000	3.0000	1.0449	2.7895	.6087
11.0000	.0822	0.0000	1.0000	4.0000	0.0000	3.0000	1.0000	2.8125	.7143
10.0000	.0801	1.0000	1.0000	3.0000	0.0000	3.0000	1.0175	2.5294	.6500
10.0000	.0857	0.0000	1.0000	4.0000	0.0000	0.0000	1.0038	2.5000	.6190
10.0000	.0769	0.0000	1.0000	4.0000	0.0000	0.0000	1.0233	4.1176	.6087
9.0000	.0781	0.0000	1.0000	4.0000	0.0000	.5000	1.0000	2.2500	.6190
9.0000	.0857	1.0000	1.0000	4.0000	0.0000	1.0000	.8943	3.0556	.6195
10.0000	.0736	0.0000	1.0000	4.0000	0.0000	.5000	1.0375	2.8333	.6481
11.0000	.0838	0.0000	1.0000	2.0000	0.0000	1.5000	.9811	2.6000	.6333
10.0000	.0822	0.0000	1.0000	4.0000	0.0000	3.0000	1.0429	2.7368	.6102
10.0000	.0782	1.0000	1.0000	4.0000	0.0000	2.0000	1.0112	2.6842	.6182
10.0000	.0832	0.0000	1.0000	4.0000	0.0000	1.5000	.9633	2.8421	.5981
10.0000	.0827	0.0000	1.0000	4.0000	0.0000	1.5000	1.0000	2.8889	.6182
10.0000	.0867	1.0000	1.0000	3.0000	0.0000	3.5000	1.0175	3.0000	.6154
10.0000	.0843	0.0000	1.0000	4.0000	0.0000	3.0000	.9841	2.7778	.6095
9.0000	.0823	0.0000	1.0000	4.0000	0.0000	2.0000	1.0335	2.8889	.6100
10.0000	.0799	0.0000	1.0000	4.0000	0.0000	3.0000	1.0944	2.8235	.6000

SUMMARY FOR C. IMPUNCTATUS

MEAN/LOW/HIGH/SD/SE

1	2	3	4	5	6	7	8	9	10
9.6957	.0806	.2609	1.0435	3.6957	.0435	2.4348	1.0211	2.8146	.6233
8.0000	.0736	0.0000	1.0000	2.0000	0.0000	0.0000	.8943	2.0455	.5636
11.0000	.0867	1.0000	2.0000	4.0000	1.0000	5.0000	1.0977	4.1176	.7143
.7648	.0036	.4490	.2085	.5588	.2085	1.4167	.0448	.3767	.0290
.1595	.0008	.0936	.0435	.1165	.0435	.2954	.0094	.0785	.0061

CORRELATION MATRIX

1.00									
.23	1.00								
.11	-.01	1.00							
-.20	-.33	-.13	1.00						
-.12	-.12	-.21	.12	1.00					
-.20	.03	-.13	-.05	.12	1.00				
-.29	-.13	.14	.24	-.26	.24	1.00			
-.24	-.60	-.10	.15	-.04	-.02	.46	1.00		
.01	.05	-.02	-.45	.13	.18	-.17	-.05	1.00	
.14	.06	-.22	-.02	-.03	-.13	.04	-.08	-.07	1.00