

UC Merced

Biogeographia - The Journal of Integrative Biogeography

Title

Morfometria multivariata e variazione geografica

Permalink

<https://escholarship.org/uc/item/63m487vv>

Journal

Biogeographia - The Journal of Integrative Biogeography, 11(1)

ISSN

1594-7629

Authors

Corti, Marco

Loy, Anna

Publication Date

1987

DOI

10.21426/B611110273

Peer reviewed

Morfometria multivariata e variazione geografica

MARCO CORTI E ANNA LOY

Dipartimento di Biologia Animale e dell'Uomo, Università di Roma, «La Sapienza»

SUMMARY

Local populations have a unity which is maintained by the interbreeding of the individuals of which populations are composed. The study of the patterns of variations in space and time among these breeding units within a species is of paramount importance for the understanding of the mechanisms leading to evolutionary change. Multivariate morphometric techniques have been reviewed here because they represent undoubtedly one of the most powerful tools to investigate trends in variation and adaptation, phenetic distinction and similarity among populations, and possibly to infer systematic relationships. The term multivariate morphometrics includes a variety of techniques which all assume the simultaneous analysis of the greatest number of informative morphometric characters. Accuracy, rapidity and geometric criteria in collecting morphometric data have been reviewed too as they contribute mostly in successful analysis of geographic variation.

1. INTRODUZIONE

Lo studio della variazione interna delle specie, intesa come processo di diversificazione geografica, assume particolare rilevanza durante il periodo della piena affermazione della Modern Synthesis e si pone come presupposto fondamentale della comprensione di tutti i meccanismi, selettivi e non, che conducono alla speciazione. Huxley (1939), coniando il termine «cline», ha definito per la prima volta la variazione geografica delle popolazioni di una specie come prodotto della selezione naturale e del flusso genico. Questi concetti, esaurientemente discussi da Mayr (1963), sono stati poi ripresi negli anni '70, quando la scoperta dell'alto livello di polimorfismo genetico nelle popolazioni naturali (Selander, 1976) ha posto nuovamente lo studio della variazione geografica al centro dell'attenzione della biologia delle popolazioni (Endler, 1977).

Sebbene lo sviluppo della citogenetica e dell'analisi dei sistemi gene-enzima abbiano contribuito enormemente a descrivere il polimorfismo delle popolazioni naturali, queste tecniche non sono tuttavia in grado di rivelare in tutta la loro complessità quei meccanismi di selezione darwiniana operanti nei processi di diversificazione geografica, e che possono essere indagati attraverso analisi morfometriche.

Quando si tratta di affrontare lo studio della variazione geografica attra-

verso tecniche di morfometria multivariata o, più in generale, di tassonomia numerica, è inevitabile incorrere nella contraddizione di fondo fra l'adesione al concetto biologico di specie, così come è definito nella biologia evolutivistica (Mayr, 1942), e l'esigenza di definire la specie o altro taxon di qualsiasi rango in termini operazionali (Sneath e Sokal, 1973). La contraddizione è insita nell'obiettivo prioritario, in quanto propedeutico, dello studio della variazione geografica: individuare operativamente le categorie utili a descrivere la distribuzione della variabilità nello spazio; questo obiettivo primario è raggiunto unicamente considerando unità tassonomiche definite in termini operazionali. Questo approccio operativo è irrinunciabile al fine di individuare quei processi di isolamento e diversificazione delle popolazioni che vedono la specie biologica come unità evolutiva minima. La variazione di popolazioni o specie affini non è descrivibile in termini di adattamento di pochi caratteri in rapporto ad una variabile ambientale, ma è di per sé un processo multidimensionale che coinvolge l'adattamento di molti caratteri quantitativi ad una varietà di parametri ambientali interdipendenti. Quindi, lo studio della variazione geografica non può che essere basato sull'analisi e la comparazione di grandi campioni e di un alto numero di caratteri, e i metodi statistici multivariati rappresentano senza dubbio lo strumento di indagine più efficace. Queste metodologie assumono particolare importanza in quanto, operando una ridefinizione degli individui nello spazio, permettono di formulare ipotesi biogeografiche: dall'esame delle funzioni elaborate può essere individuato un modello di variazione che si presenta in forma di cline, stepcline o a cluster categorici (Endler, 1977).

Nella maggior parte delle analisi morfometriche è presente, implicitamente o esplicitamente, l'assunto che la variabilità fenotipica osservata possa essere considerata variabilità genetica (Atchley, 1983). Sebbene questa condizione non sia facile da verificare, a causa di una componente non-genetica difficilmente distinguibile, l'approccio morfometrico presenta notevoli vantaggi: 1) molti caratteri hanno un'origine poligenica e la loro variabilità può fornire un contributo maggiore all'analisi dei campioni e condurre, in alcuni casi, a risultati migliori rispetto all'analisi dei sistemi gene-enzima (Thorpe *et al.*, 1982). Ad esempio, si stima che in *Mus musculus* la crescita in peso corporeo veda coinvolti, in modo conservativo, più di 100 loci (Comstock 1973) e le metodologie morfometriche multivariate sono in genere utilizzate per accertare la purezza delle linee inbreed di topi da laboratorio (Festing, 1972); 2) i dati morfometrici consentono di indagare sulla variabilità sia su una dimensione spaziale che temporale (Williamson, 1981; Zink, 1983); 3) la reperibilità di materiale da collezione nei musei nonché la rapidità, accuratezza e economia del rilevamento dei dati consentono l'analisi di grandi campioni in breve tempo.

I risultati derivanti da analisi biometriche e da studi su frequenze geniche devono essere confrontati per verificare la congruenza nei pattern di variazione e di distanza delle popolazioni/specie (Lessios, 1981; Lambert e Paterson, 1982; Hausser, 1984), anche se questa concordanza è piuttosto rara, a causa

della diversa potenza dei test utilizzati (Lewontin, 1984) e/o del diverso significato evolutivo e biologico dei pattern di variazione (Turner, 1974; Lessios, 1981; Yoshiyama and Sassaman, 1983; Snyder and Linton, 1984; Lindenfesler, 1984).

2. Scelta dei caratteri

Nello studio delle relazioni fenetiche la fonte dei dati è generalmente rappresentata da frequenze geniche (Menozzi *et al.* 1978; Piazza *et al.*, 1981; Sokal e Menozzi, 1982), caratteri citogenetici (Nevo *et al.*, 1987) e morfometrici (Gould e Johnston, 1972; Thorpe, 1976; Oxnard, 1978; Reyment, 1981). Ci riferiamo a relazioni fenetiche come alla similitudine complessiva basata su tutti i caratteri disponibili, senza riguardo al diverso peso che questi hanno nel determinare le relazioni tra popolazione/specie (Sneath and Sokal, 1973). Un criterio generale su cui deve essere basata ogni analisi fenetica, è quello di considerare il più alto numero possibile di caratteri, indipendentemente dalla loro natura o tipo (Mc Neill, 1978): più alto è il numero di caratteri considerati, maggiori sono le informazioni che si possono trarre sullo stato di ogni OTU (si definisce OTU — Operational Taxonomic Unit — il taxon di più basso rango utilizzato in un dato studio, Sneath e Sokal, 1973). A livello intraspecifico, il numero di caratteri necessari per una descrizione esauriente è influenzato dal loro livello di correlazione. Più alto è questo livello (e di conseguenza la ridondanza dell'informazione) più alto deve essere il numero di caratteri necessari per raggiungere lo stesso livello di informazione (Thorpe, 1976).

Per quanto riguarda la morfometrica multivariata, la scelta di determinati complessi di caratteri morfometrici non restringe l'informazione ad una particolare classe di geni; questa ipotesi di non-specificità è stata discussa da Sneath e Sokal (1973). Sebbene a volte non vi sia congruenza tra classificazioni basate su insiemi distinti di caratteri morfometrici, nell'insieme i risultati spesso concordano; è comunque possibile stabilire di volta in volta il grado di congruenza tra i diversi risultati attraverso lo studio dei coefficienti di correlazione tra le matrici di somiglianza elaborate. È sconsigliabile invece l'uso di caratteri la cui risposta all'ambiente è altamente variabile, tanto da rendere difficile la distinzione tra determinazione genetica e ambientale.

2.1 Rilevamento dei caratteri

Gli studi morfometrici sono spesso basati sulla trattazione di poche misure o parametri, considerati singolarmente. È comune, ad esempio, l'uso di parametri distinti come lunghezza del corpo, altezza e peso, come anche di parametri standard di misura (ad esempio, vedi Thomas, 1905, per una descrizione delle misure craniali per i mammiferi). Sebbene questo approccio sia ancora molto diffuso, la sua utilità nel risolvere problemi biologici può essere molto limitata (Oxnard, 1978; Bookstein, 1982). Il potere descrittivo

dell'analisi viene invece notevolmente incrementato dal rilevamento di un alto numero di caratteri, che, considerati nel loro complesso, permettono di definire in modo esauriente la variabilità morfometrica delle OTU.

I metodi e gli strumenti di rilevamento dei caratteri morfometrici hanno subito in questi ultimi anni una notevole evoluzione, passando dal tradizionale calibro al calibro automatizzato (Marcus, 1983), all'oculare micrometrico (Hausser e Jammot, 1974), a metodi di rilevamento su coordinate (Festing, 1972), fino ai più sofisticati sistemi automatizzati di analisi dell'immagine (Rohlf and Ferson, 1983; Ferson *et al.*, 1985). Nel rilevamento dei dati su un sistema di coordinate (Fig. 1) le variabili rappresentano non più delle misure a sé stanti e indipendenti per ogni singolo esemplare, ma ogni oggetto è riferito allo stesso punto di origine del piano cartesiano. Questo metodo è estremamente rapido ed accurato ed è stato applicato da Thorpe *et al.* (1982) e da Corti *et al.* (1987).

Un altro metodo, utilizzato principalmente per evidenziare variazioni della forma in rapporto alla crescita iso- e allometrica, è stato descritto da Strauss e Bookstein (1984) (Fig. 2), e consiste nell'identificare un insieme di punti di riferimento omologhi lungo i contorni esterni di un oggetto, calcolando le distanze tra punti contigui e quindi le loro triangolazioni. Un rischio insito nel metodo consiste nella propagazione di errori anche piccoli che si risolvono in una distorsione della forma. Un approccio analogo è stato seguito da Dahl *et al.* (1984) che hanno utilizzato il metodo computerizzato SIMCA, sviluppato per studio non biologici, per individuare modelli di variazione della forma degli oggetti in esame.

È importante sottolineare che seguendo metodologie di rilevamento dei caratteri su un sistema di riferimento quale il piano cartesiano o altro, i caratteri non posseggono più un valore individuale ma questo viene trasferito all'intero complesso dei caratteri.

3. ANALISI DEI DATI

L'alto numero di caratteri e l'esigenza di analizzarli nel loro insieme, impongono l'uso di tecniche statistiche di analisi che, attraverso una trattazione contemporanea di tutti i dati, tengano conto delle relazioni che intercorrono fra i caratteri stessi. L'analisi si pone i seguenti obiettivi:

- 1) individuare nella struttura dei dati le categorie morfometriche (OTU);
- 2) individuare nello spazio multivariato il modello di variazione dell'OTU, in termini di clini, step clini o cluster categorici;
- 3) fornire una rappresentazione cartografica della variazione, con l'uso di metodi appropriati di mappatura;
- 4) quantificare le differenze tra le OTU attraverso l'elaborazione di matrici di distanze;

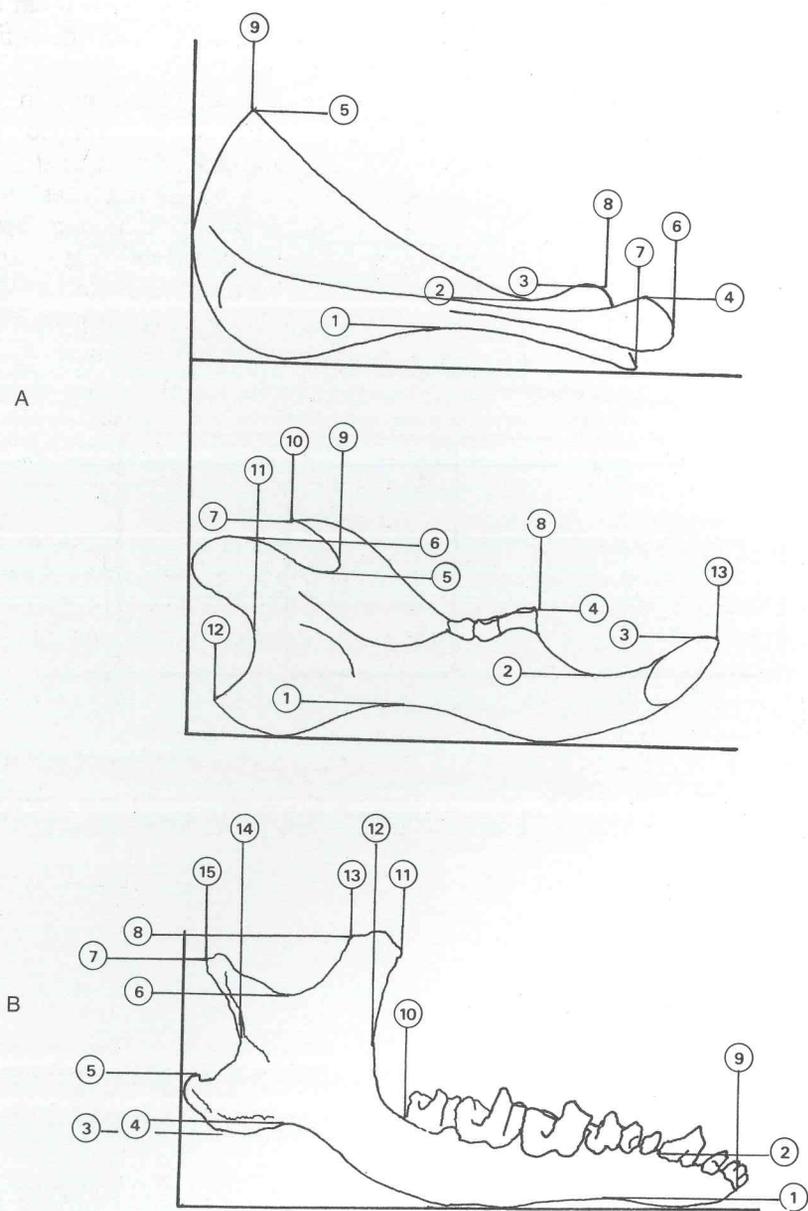


FIG. 1 - Metodo di rilevamento su coordinate. A) Caratteri rilevati su mandibola e scapola di *Mus domesticus* (Thorpe *et al.*, 1982); B) Caratteri rilevati su mandibola di *Talpa* (Corti *et al.*, 1985). I caratteri riguardano un solo lato (sinistro) del corpo, in modo da eliminare a priori qualsiasi influenza dovuta ad asimmetria laterale. Gli elementi scheletrici sono posti su carta millimetrata ridotta fotograficamente 9 volte e osservati con binoculare stereoscopico; la posizione costante è ottenuta poggiando gli elementi a due vetrini portaoggetto fissi e delimitanti due assi cartesiani. È importante sottolineare che durante l'elaborazione multivariata, i caratteri, o meglio, i punti di riferimento scelti sul piano cartesiano, perdono il loro valore individuale, che viene trasferito all'intero complesso dei caratteri.

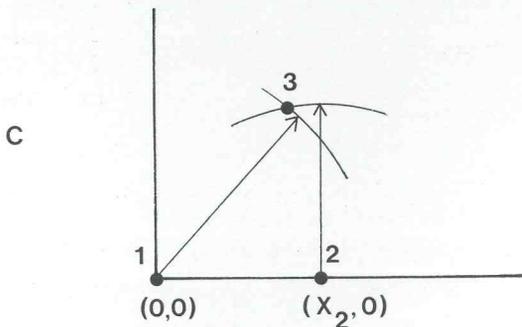
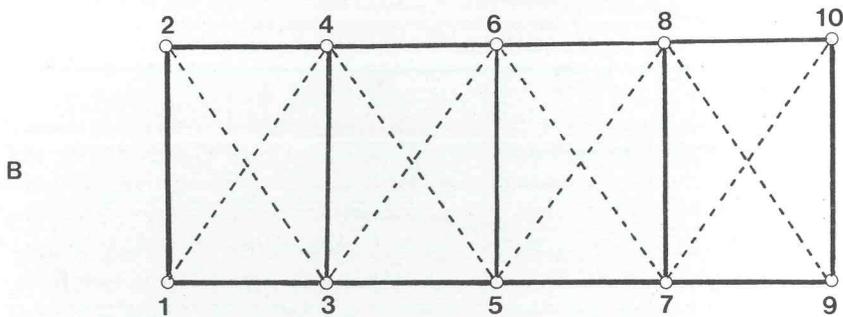
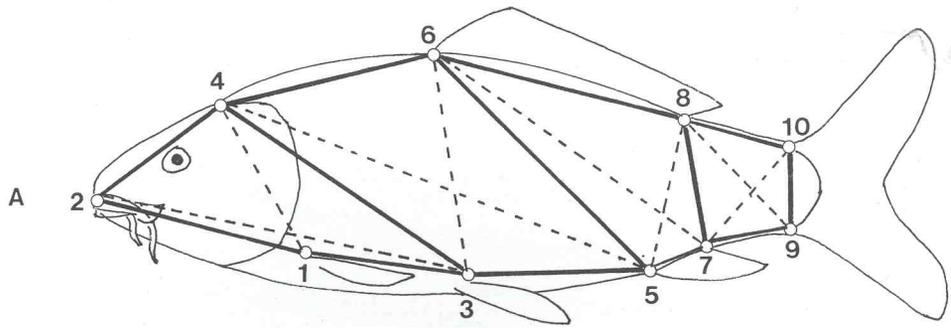


FIG. 2 - Sistema di rilevamento di caratteri per punti di riferimento omologhi. A) 21 caratteri rilevati in ceppi di carpa comune (Corti *et al.*, 1987); B) rappresentazione dello schema: le variabili rilevate riguardano sia i lati che le diagonali di ogni singolo quadrilatero. C) Il primo punto di riferimento (1) costituisce l'origine degli assi cartesiani e il segmento che unisce i primi due punti (2) rappresenta l'ascissa; le coordinate del terzo punto di riferimento sono determinate dall'intersezione delle due circonferenze che hanno come centro i primi due punti e come raggio le rispettive distanze da (3); il quarto punto viene individuato rispetto al secondo e al terzo fino a quando sono state individuate le coordinate per tutti i punti. Questa metodologia consente di: 1) individuare variazioni di forma in senso verticale, orizzontale ed obliquo; 2) ricostruire (mappare) la forma a partire da un insieme di punti di riferimento; 3) caratterizzare e visualizzare tendenze multivariate (forma e allometria) all'interno delle popolazioni; 4) standardizzare la forma per comparazioni fra gruppi/popolazioni diverse (vedi Straus e Bookstein, 1982, per una discussione esauriente).

- 5) individuare le correlazioni fra fattori storici e/o climatici e la variazione morfometrica.

3.1 Trasformazione dei dati

Spesso l'insieme dei dati morfometrici raccolti non è direttamente analizzabile: la trasformazione logaritmica dei dati costituisce in generale una procedura opportuna per rendere lineari e confrontabili i dati dei diversi campioni (Corruccini, 1975). La trasformazione logaritmica ha lo stesso vantaggio della standardizzazione nel produrre nuove variabili che sono indipendenti per quanto riguarda l'entità della varianza, ma, al contrario della standardizzazione, trattiene tutta l'informazione relativa all'ordine di grandezza (Schoenewald-Cox *et al.*, 1985).

Parallelamente ad una trasformazione matematica che la renda statisticamente appropriata, la struttura dei dati contiene elementi che hanno un significato prettamente biologico e che devono essere identificati prima di procedere in qualsiasi analisi: alla variazione geografica è associata una componente non-geografica (dimorfismo sessuale, variazione allometrica e ontogenetica, asimmetria laterale) che deve essere comunque considerata, sia che la si voglia eliminare o mantenere, per permettere una corretta interpretazione della variazione.

Il dimorfismo sessuale costituisce una forte sorgente di variabilità in diversi sistemi di caratteri (Fig. 3), che può essere consigliabile eliminare all'origine (Thorpe, 1976; Corti *et al.*, 1985), sebbene ad esso debba essere dedicata una particolare attenzione quale importante fattore selettivo nel determinare variazioni geografiche (Gould e Johnston, 1972), come è stato elegantemente descritto da Johnston e Selander (1973) per il passero domestico e da Thorpe (1975) per *Matrix natrix*.

L'età degli individui costituisce un'ulteriore fonte di variazione intraspecifica non-geografica. Il problema è di ovvia risoluzione per gli organismi che attraversano diversi stadi vitali (larve, pupe, ecc.); non lo è però nel caso di organismi che si accrescono in modo più o meno continuo fino allo stadio di adulto (variazione ontogenetica). L'influenza della crescita allometrica, uno degli aspetti più importanti della variazione ontogenetica, può essere eliminata attraverso l'adozione di appropriate procedure statistiche, dove per allometria si definisce generalmente lo studio della dimensione e delle conseguenze legate alla sua variazione in rapporto all'accrescimento (Atchley, 1983). Per comprendere il tipo di relazione esistente fra età e dimensioni si può calcolare l'equazione di regressione tra i rispettivi logaritmi (Sneath e Sokal, 1973; Thorpe, 1976).

Nel caso in cui sia individuabile una correlazione diretta, si possono scegliere due strategie: 1 - individuare classi di età ben definite e non ambigue, e svolgere analisi separate per campioni omogenei; solo in questo caso varia-

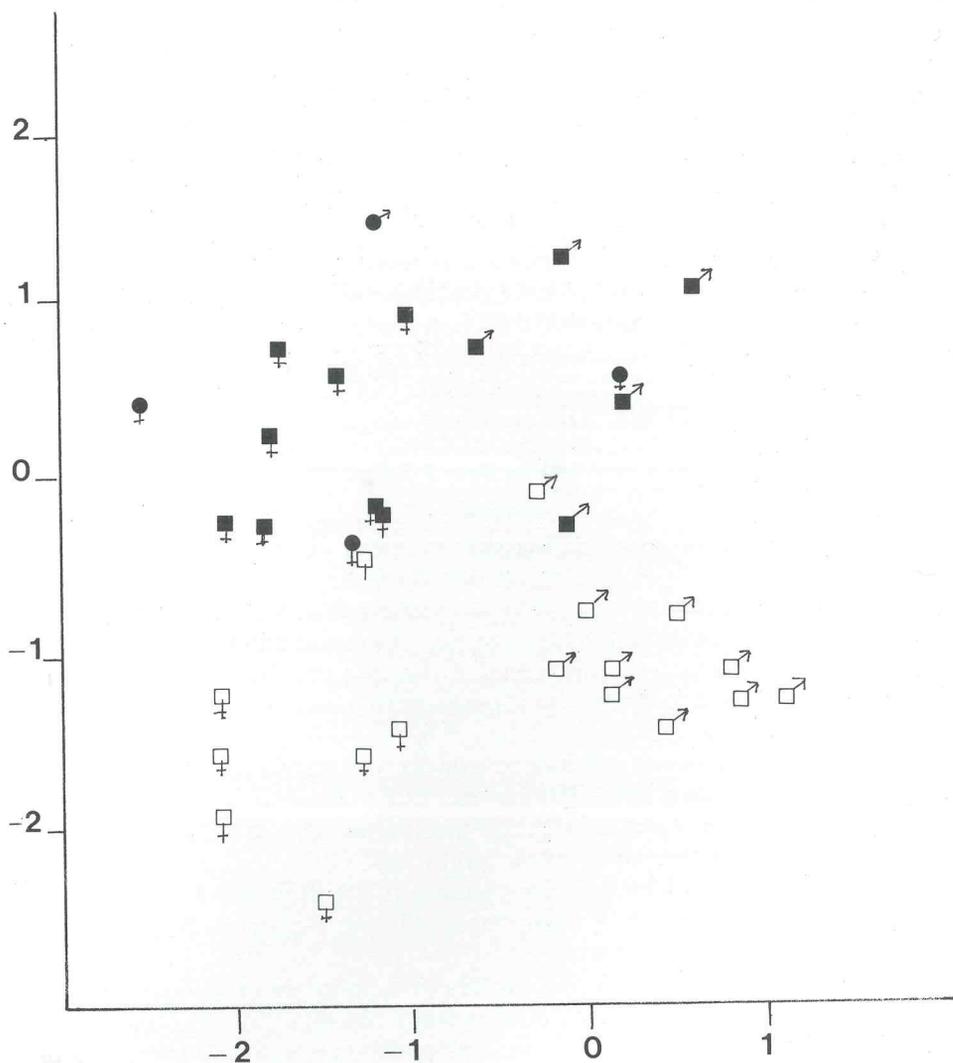


FIG. 3 - Analisi delle componenti principali condotta congiuntamente per i due sessi sulle tre sottospecie di *Ursus arctos* (Vigna Taglianti *et al.*, 1984). Appare evidente l'influenza del dimorfismo sessuale sulla distribuzione degli individui nello spazio multivariato.
U. arctos marsicanus □; *U. arctos alpinus* ●; *U. arctos pyrenaicus* ■.

zioni nella dimensione possono essere interpretate in termini di variazione geografica; 2 - eliminare a priori il fattore dimensione e rendere i caratteri «size-independent».

Nell'adottare questa seconda scelta, un metodo consiste nell'identificare un coefficiente che permetta di eliminare l'effetto delle dimensioni. Blondel *et al.* (1984), dividendo ogni carattere per un coefficiente direttamente correlato alle dimensioni (es. peso o volume), hanno ottenuto una nuova matrice

di dati in cui ogni variabile è considerata «size-independent». La trattazione di matrici di dati costituiti da rapporti rappresenta però un problema per quanto riguarda la statistica multivariata, a causa della strana distribuzione assunta dalla struttura dei dati (Atchley *et al.*, 1976).

Un metodo statistico più appropriato consiste nell'analisi della covarianza, dove un carattere *size dependent* è regredito contro un carattere *size independent* (Gould, 1966; Sokal and Rohlf, 1969; Thorpe, 1976) (Fig. 4). Poiché la maggior parte dei componenti ontogenetici sono multivariati in natura, un metodo più adeguato per rendere i caratteri size-independent si basa sull'uso delle componenti principali per individuare, ed eventualmente eliminare, l'influenza del fattore dimensione prima di definire qualsiasi affinità razziale e/o geografica tra i campioni. Se tutti i coefficienti del primo componente principale sono dello stesso segno ed hanno approssimativamente lo stesso valore,

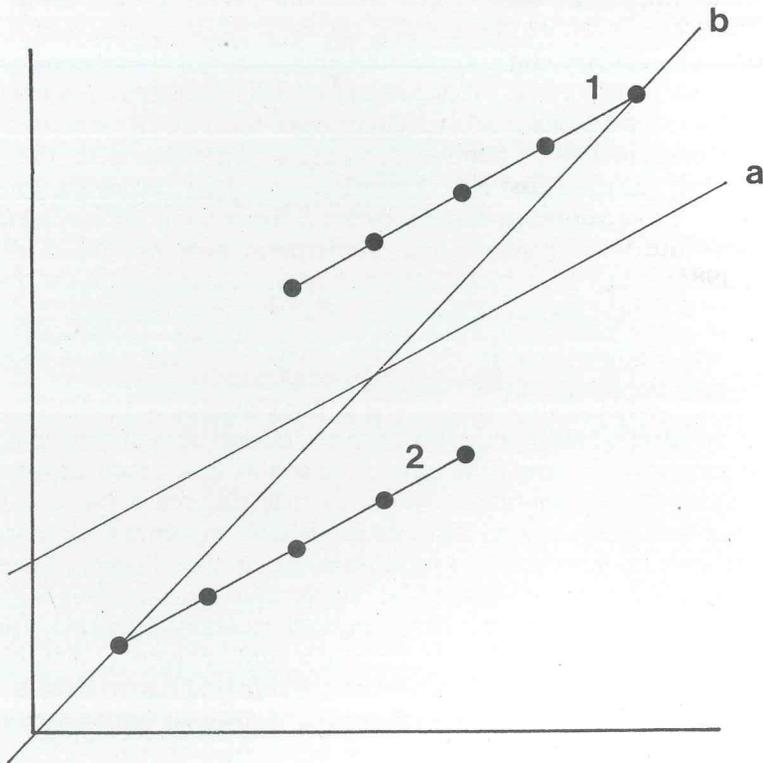


FIG. 4 - Un criterio per eliminare l'influenza allometrica consiste nell'individuare l'angolo definito dalla retta di regressione tra ogni carattere (ordinata) e le dimensioni dei singoli individui (ascissa). Le rette 1 e 2 sono due rette di regressione che rappresentano la variazione allometrica di un carattere all'interno di due popolazioni ipotetiche (i punti corrispondono ad individui di taglia diversa). Per individuare un coefficiente di correzione per i dati di tutte le popolazioni in esame, la procedura più corretta consiste nel calcolare la retta di regressione A sulla base delle rette di regressione delle singole popolazioni (1 e 2): l'angolo formato dalla retta A corrisponde all'angolo «medio» tra le rette 1 e 2. Al contrario, la retta di regressione B calcolata direttamente su tutti gli individui, è inappropriata, perché non tiene conto del rapporto esistente tra la variazione intra- ed inter- popolazione (Thorpe, 1976).

allora si può asserire che la direzione di maggiore variazione lungo l'asse del primo componente principale è correlata con il fattore dimensione (Neff e Marcus, 1980); in tal senso il primo componente è da considerarsi *size-component* e la sua eliminazione permette di considerare i restanti componenti principali come una nuova matrice di $n-1$ variabili *size-independent* (Jolicoeur e Mosiman, 1960). Thorpe (1983) e Thorpe *et al.* (1982) suggeriscono un diverso metodo che tenga conto delle differenze esistenti fra variazioni intra-OTU e variazioni inter-OTU: a partire da una matrice di varianza-covarianza intra-OTU, si calcolano le componenti principali (Multiple Group Principal Components) e, eliminando il primo componente, si considerano gli altri come nuovi dati *size independent*.

Spesso, e qualora possibile, diventa più importante conservare l'aspetto dimensionale piuttosto che considerarlo biologicamente non importante e scartarlo: infatti, l'alto numero di geni che controllano le dimensioni corporee suggerisce che il fattore dimensione sia da considerarsi in una relazione pleiotropica virtualmente con tutti gli altri loci presenti nel genoma (Atchley *et al.*, 1981). In questo caso, l'individuazione della relazione esistente tra la variazione della dimensione e la variazione della forma (Jolicoeur, 1986; Jolicoeur and Mosiman, 1960) permette di valutare l'influenza della dimensione e di distinguerla dalla forma nella determinazione della variazione geografica dei campioni. Ad esempio, in *Rattus rattus*, la variazione delle dimensioni è una componente fondamentale la cui ereditarietà raggiunge valori del 50% (Atchley, 1983).

3.2 Analisi multivariata

Le metodologie statistiche multivariate permettono di trasformare l'alto numero di variabili descrittive in un numero ridotto di funzioni che riassumono l'informazione in esse contenuta. Nell'analisi morfometrica questi metodi permettono di massimizzare o riassumere le differenze fra le OTU che vengono espresse, nello spazio multidimensionale, in termini di distanza o vicinanza relativa. È importante sottolineare che queste distanze sono un'espressione diretta del grado di similitudine morfometrica fra le OTU, e che non possono essere interpretate in termini di relazioni filetiche (Albrecht, 1980), le quali rappresentano invece un argomento successivo di speculazione. Gould e Johnston (1972), Thorpe (1976 e 1983), Oxnard (1978) e Reyment (1981) hanno esaurientemente discusso i metodi di morfometria multivariata applicati allo studio della variazione geografica, mentre Neff e Marcus (1980) hanno compiuto un'ampia revisione delle metodologie statistiche utilizzate in sistematica. Esistono diversi programmi per l'elaborazione dei diversi tipi di analisi multivariata attraverso elaboratori elettronici, la maggior parte dei quali è oggi disponibile anche per personal computers. Nonostante la facilità di accesso, resa possibile dai sistemi computerizzati, l'uso di queste metodologie presuppone la conoscenza e la verifica di alcuni assunti di base che, se non rispettati, possono seriamente compromettere l'attendibilità dei

risultati. Neff e Marcus (1980) hanno discusso e descritto i diversi assunti, definendo gli ambiti in cui i diversi metodi possono essere applicati; Andrews *et al.* (1973) hanno sviluppato in modo particolare il problema dei test di multinormalità; Gould (1984) ha discusso l'utilizzazione contemporanea di diversi tipi di caratteri, continui, meristici e binomiali.

I metodi di analisi multivariata più frequentemente utilizzati nel descrivere modelli di variazione geografica su caratteri continui sono l'analisi delle componenti principali (PCA) e l'analisi discriminante (DFA), ognuna delle quali ha presupposti e finalità diverse.

3.2.1 *Analisi delle componenti principali (PCA)*

La PCA costituisce un'utile tecnica esplorativa: i campioni sono disposti in un nuovo spazio, descritto da un insieme appropriato di assi generato attraverso l'analisi. Per generare le nuove variabili, dette componenti principali, la PCA opera trasformazioni lineari delle variabili originali, individuando degli assi indipendenti ed ortogonali lungo le direzioni di maggiore varianza. Nel nuovo spazio creato dalle componenti principali, gli individui sono distribuiti in base ai valori assunti rispetto ai singoli assi, ogni caso in modo indipendente dall'altro. L'estrazione delle componenti principali può essere operata sulla matrice di correlazione o sulla matrice di varianza-covarianza: la scelta tra i due metodi è dettata dalla struttura dei dati ed in particolare dalla distribuzione della varianza nelle variabili originali (vedi Neff e Marcus, 1980, per una trattazione approfondita).

La PCA è dunque un metodo di ordinazione, utilizzato principalmente per individuare gruppi omogenei e pattern di variazione sulla base di un elevato numero di caratteri.

Esiste un'ampia bibliografia relativa all'utilizzazione di questi modelli nell'ambito della morfometria multivariata: Sneath e Sokal (1973) hanno riportato ben 600 lavori di morfometria multivariata pubblicati fino al 1971. Le applicazioni della PCA riguardano sia il campo dell'analisi dei sistemi genenzima (Menozzi *et al.*, 1978; Sokal, 1983) sia l'ambito più propriamente morfometrico (Johnston e Selander, 1971; Jolicoeur e Mosiman 1960, Hausser e Jamnot, 1974; Thorpe, 1980; Hausser, 1984; Schonewald-Cox, 1985).

Alcuni esempi riportati qui di seguito riguardano l'utilizzazione della PCA come strumento per l'identificazione di modelli di variazione geografica o affinità razziali.

Corti *et al.* (1985) hanno utilizzato la PCA come primo strumento per definire lo stato tassonomico di *Talpa europea* e di *T. romana*, spesso poste in sinonimia: la netta separazione in cluster e l'assenza di gradienti geografici hanno permesso di formulare un'ipotesi di separazione a livello specifico. Un pattern di separazione analogo è stato verificato in *T. romana* e *T. stankovici* (Capanna *et al.*, 1985) (Fig. 5). Nello stesso modo è stato affrontato il problema delle due sottospecie di *Rupicapra rupicapra*, *R.r. ornata* e *R.r. pyrenaica* (Scala e Lovari, 1984). Gradienti di variazione clinale lungo gli assi delle

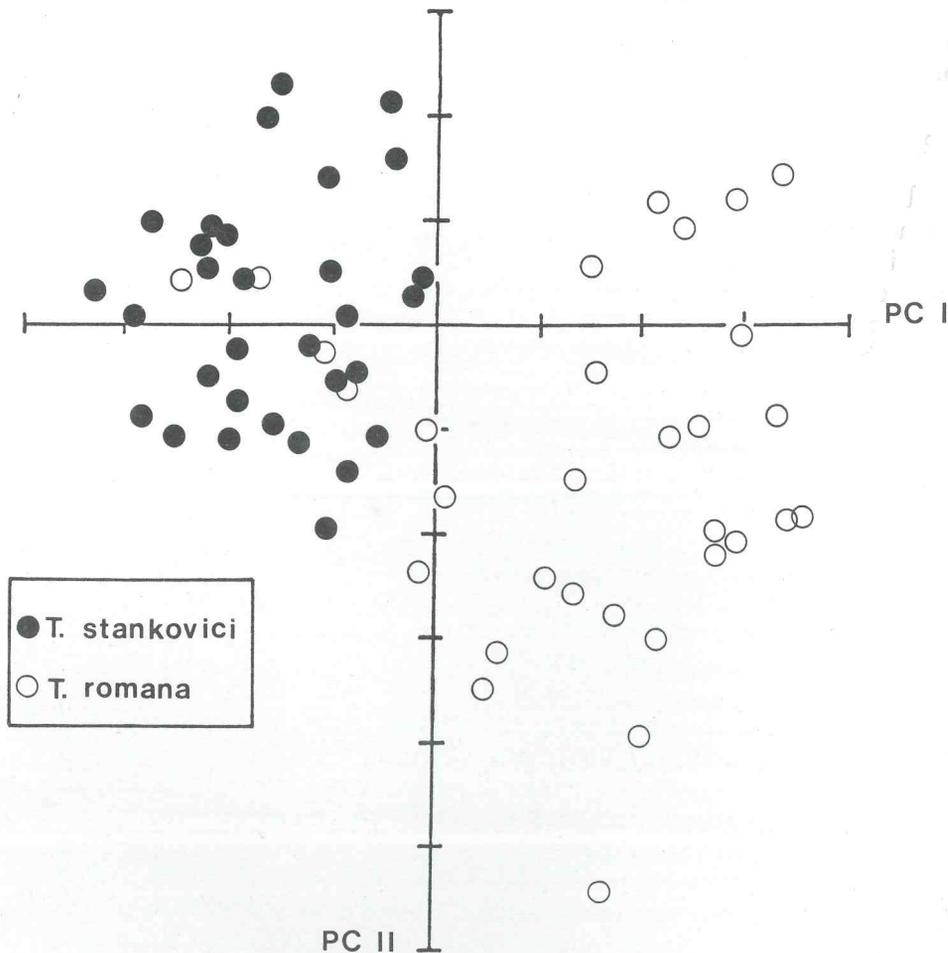


FIG. 5 - Analisi delle componenti principali condotta su coordinate di componenti scheletrici di *Talpa romana* (Italia centro-meridionale) e *T. stankovici* (Macedonia), (Capanna *et al.*, 1985). Gli individui appaiono distribuiti in due cluster, e ciò suggerisce una separazione dei due taxa a livello specifico.

componenti principali sono stati invece rilevati dal Blondel *et al.* (1984) nello studio sulla convergenza eco-morfologica delle comunità di uccelli mediterranei, e da Handford (1983) in uno studio sulla correlazione tra variabili geoclimatiche e caratteri morfologici del passero *Zonotrichia capensis*, distribuito nel continente sud americano. Johnston e Selander (1971) hanno utilizzato il primo componente principale per descrivere la variazione delle dimensioni del passero domestico (*Passer domesticus*) sull'intero continente nord americano.

I gruppi omogenei identificati attraverso la PCA possono in seguito essere analizzati attraverso differenti tecniche, tra cui l'analisi discriminante, che ne

illustrino la separazione e/o che richiedono come assunto la separazione a priori di gruppi (Corti *et al.*, 1985; Schonewald-Cox *et al.*, 1985).

3.2.2 *Analisi discriminante*

Il termine analisi discriminante (DFA) si riferisce ad alcune tecniche che hanno in comune una matrice di dati i cui campioni sono divisi in gruppi definiti a priori. Qui di seguito ci riferiremo all'analisi discriminante classica, che peraltro fornisce un modello generale anche per l'analisi canonica e l'analisi multivariata della varianza.

L'obiettivo della DFA è quello di massimizzare la separazione delle OTU, distribuendole in uno spazio di dimensioni ridotte in cui la varianza tra i gruppi è massimizzata rispetto alla varianza all'interno dei gruppi.

I gruppi sono definiti a priori e la DFA produce una funzione lineare di caratteri (funzione discriminante) che permette di quantificare le distanze fra i centroidi dei gruppi, verificare la corretta classificazione degli esemplari e assegnare esemplari sconosciuti ai gruppi definiti a priori. Inoltre, individua le variabili originali che contribuiscono in misura maggiore al potere discriminante delle singole funzioni.

Tuttavia, questo modello presenta delle restrizioni: sono descritte solo le relazioni lineari tra variabili continue e devono essere rispettati gli assunti di multinormalità e omogeneità delle matrici di varianza-covarianza dei gruppi (per una discussione del problema vedi Neff e Marcus, 1980). Nei dati di origine biologica la multinormalità è una condizione difficilmente rispettabile, anche nel caso in cui una condizione di normalità sia verificata per ogni singola variabile (Andrews *et al.*, 1973). Ciò nonostante, una trasformazione a priori dei dati originali (es., una trasformazione in logaritmi) può contribuire a rendere l'insieme dei dati più lineari e vicini ad una condizione di multinormalità (Gnanadesikan, 1977; Neff and Marcus, 1980).

Anche in questo caso la bibliografia è estremamente ampia e diversificata, e una rassegna esauriente sull'uso della DFA in morfometria è fornita da Albrecht (1980).

Nel lavoro di Corti *et al.* (1985) questa procedura ha consentito di quantificare la separazione di *Talpa europea*, *T. romana* (Fig. 6), fornendo così un modello di relazione, in termini di similitudine relativa, tra le popolazioni dei due taxa.

Similmente Thorpe *et al.* (1982) hanno utilizzato la DFA per discriminare in modo categorico popolazioni robertsoniane di *Mus domesticus* riproduttivamente isolate tra loro, individuando uno stepcline di variazione morfologica congruente con un'ipotesi di evoluzione cromosomica paraptrica (Fig. 7). In un elegante studio di Thorpe (1984a) la DFA ha permesso di identificare due diverse zone di contatto primario e secondario in *Natrix natrix*, nei Pirenei e nei Balcani. I risultati hanno fornito un modello chiaro e sintetico dell'espansione da sud a nord dell'areale di *Natrix natrix*.

La procedura di classificazione Jackknifed, contenuta in alcuni packages,

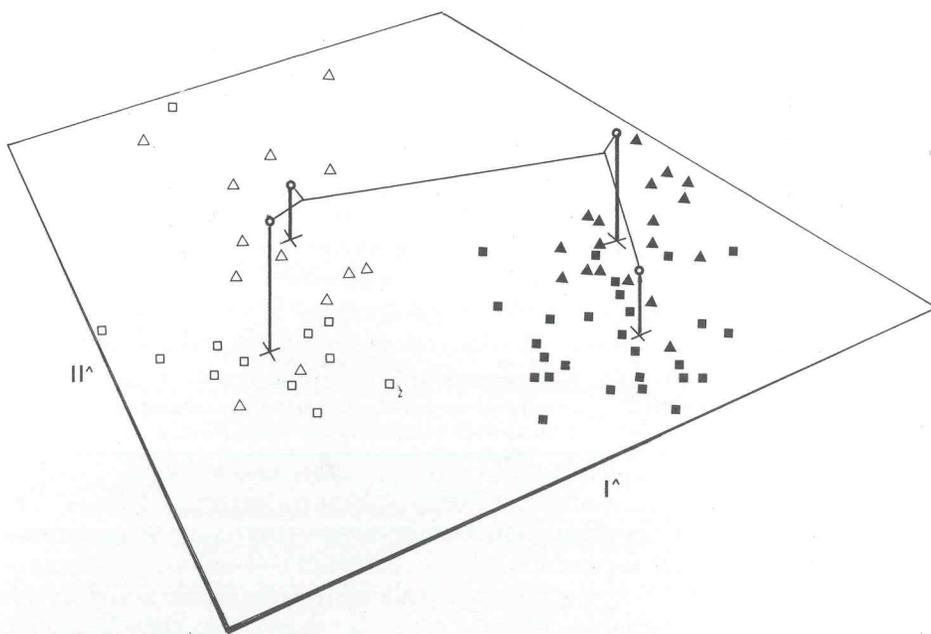


FIG. 6 - Distribuzione su un piano canonico di due popolazioni italiane centro-meridionali di *Talpa romana* e due popolazioni settentrionali di *T. europea*. Non si osserva un cline di variazione nord-sud determinato da possibile conspecificità delle popolazioni dei due taxa, ma piuttosto appare evidente una separazione categorica in cluster distinti. Un Wagner network, calcolato in base a una matrice di distanze di Manhattan fra i centroidi dei gruppi, è stato sovrainposto in una terza dimensione in modo da visualizzare le distanze fra le OTU. L'asse verticale è determinato dalla terza variabile canonica.

fornisce un indice generale di attribuzione corretta degli individui ai gruppi definiti a priori e rappresenta, assieme alla percentuale di classificazione corretta, una valida indicazione sulla robustezza complessiva dell'analisi condotta.

Nell'interpretazione dei risultati della DFA è importante non incorrere in un elemento di confusione che spesso tende ad accomunare il significato biologico e statistico dei termini classificazione, discriminazione e identificazione: in biologia, classificare significa ordinare gli individui in categorie, discriminare significa trovare differenze tra le categorie, identificare significa assegnare correttamente individui sconosciuti alla loro categoria di appartenenza (Sneath e Sokal, 1973); in statistica, il termine discriminazione assume il significato di assegnare individui sconosciuti alla loro corretta categoria di appartenenza. Quindi la discriminazione in statistica ha assunto il significato che in biologia ha l'identificazione (Kendall, 1973).

Per quanto riguarda la definizione sistematica dei taxa, il problema è che non esiste una chiara definizione di un taxon in termini di struttura multivariata (Neff e Marcus, 1980): il risultato delle analisi è oggetto di ipotesi che devono essere comunque verificate sulla base di altri modelli.

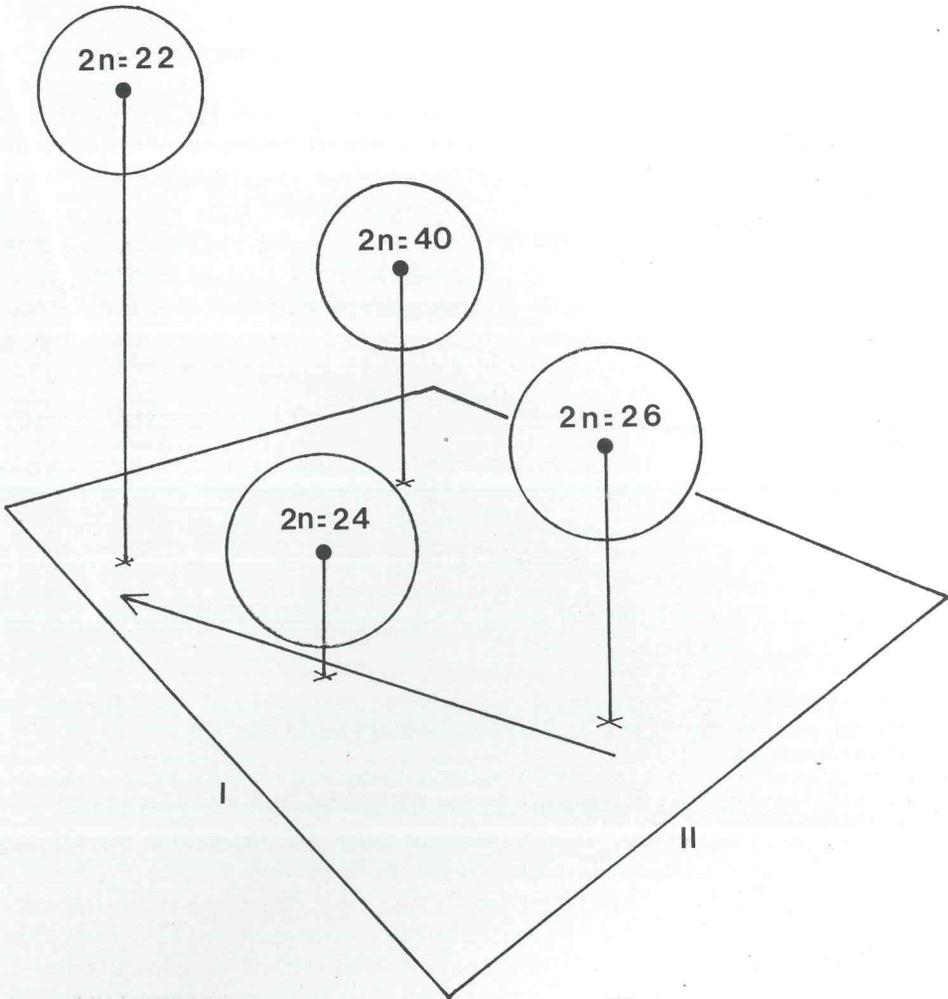


FIG. 7 - Stereogramma rappresentante i risultati dell'analisi discriminante condotta su quattro popolazioni di *Mus domesticus* (Thorpe *et al.*, 1982). I numeri diploidi rappresentano tre popolazioni robertsoniane parapatriche a numero cromosomico ridotto e $2n = 40$ una popolazione a cariotipo ancestrale. L'asse verticale è determinato dalla terza variabile canonica e il raggio delle circonferenze di probabilità attorno i centroidi del gruppo è pari al doppio della deviazione standard. Le popolazioni robertsoniane appaiono chiaramente discriminate fra loro ed è evidente uno step-cline di variazione morfometrica che coincide con l'ipotesi evolutiva di riduzione del numero diploide (rappresentato dalla freccia).

Una volta definito il modello di variazione delle OTU, devono essere individuati quei fattori casuali che hanno determinato il modello di variazione stesso; ciò significa esaminare la struttura dei dati in modo tale da poter distinguere l'influenza della selezione adattativa da tutti quei fattori stocastici (deriva genetica, effetto fondatore, ecc.) che hanno avuto un'influenza sulla diversificazione geografica delle OTU.

Se la variazione è determinata da selezione naturale, ci si deve aspettare una forte associazione fra i caratteri morfologici e parametri ambientali; le eventuali correlazioni rappresentano l'evidenza dell'azione della selezione nel determinare il modello di variazione. Gould e Johnston (1972) descrivono diversi tipi di analisi della regressione tra caratteri morfometrici e variabili geoclimatiche, seguendo un approccio sia univariato che multivariato.

Per identificare l'influenza della selezione adattativa sulla variazione geografica di quattro specie di *Sorex*, Hausser (1984) ha compiuto un'analisi della regressione multipla per ciascun carattere morfometrico e diversi parametri ambientali; questa analisi ha poi permesso di elaborare funzioni di correlazione canonica che, descrivendo le associazioni fra variabili ambientali e morfometriche, hanno posto in evidenza un cline transpecifico di variazione. L'analisi della regressione consente anche di identificare, sotto forma di residui, quella parte di variazione che non è direttamente correlata ai parametri ambientali e che può essere quindi spiegata in termini di deriva genetica e/o variazione ecologica (Hausser, 1984). Se la deriva genetica è anch'essa responsabile della variazione inter-OTU contenuta nei residui, i suoi effetti saranno distribuiti a caso sui diversi caratteri; se invece le variabili sono in qualche modo intercorrelate, i loro residui saranno anch'essi tra loro correlati e non potranno essere utilizzati per individuare l'effetto dovuto a deriva genetica (Hausser, 1984).

4. RAPPRESENTAZIONE DELLA VARIAZIONE GEOGRAFICA

I risultati delle analisi delle Componenti Principali e dell'Analisi Discriminante sono generalmente riassunti in grafici a due (grafici bivariati) o tre dimensioni (stereogrammi), che permettono una visualizzazione immediata della distribuzione delle OTU nello spazio multivariato.

Questa distribuzione può presentarsi come cline, stepcline o cluster categorici (Endler, 1977). Mentre una distribuzione clinale delle OTU indica un continuum di variazione, i cluster categorici mostrano il raggiungimento di un livello definitivo (attuale) di separazione. Gli stepcline assumono invece un significato particolare, in quanto sono spesso associati a zone di ibridazione, sia primarie che secondarie (Mayr, 1963), e dal loro esame possono essere rivelati processi di speciazione in atto (Huxley, 1938; Mayr, 1963; Endler, 1977; White, 1978). Modelli di distribuzione a step cline basati su analisi morfometriche sono stati messi in evidenza da Thorpe (1984a) e da Barnes *et al.* (1985).

Poiché nello stereogramma le relazioni fra OTU vicine possono essere soggette a effetti di distorsioni dovute alle tecniche di ordinazione, come spesso avviene durante l'elaborazione dei componenti principali, è utile sovrapporre a tali stereogrammi un diagramma di relazioni (*minimum spanning tree*, Gower e Ross, 1969; Rohlf, 1970) che ne illustri l'eventuale distorsione (Johnston e Selander, 1971; Sneath e Sokal, 1973).

La rappresentazione cartografica delle analisi condotte fornisce un'utile

la rappresentazione dei modelli di variazione geografica delle OTU in rapporto all'area esaminata. Gabriel e Sokal (1969) distinguono due modi diversi di illustrare la variazione su una mappa: per tendenze e per categorizzazione delle aree.

Per tendenze si intende l'individuazione di gradienti sfumati in rapporto a caratteri che variano in modo continuo; esempi illustrativi sono forniti da Johnston e Selander (1971) hanno costruito mappe di contorno sulla base delle singole componenti principali, usando il programma SYMAP; da Handford (1963), che ha prodotto mappe di contorno sulla base delle componenti principali e singole variabili usando l'algoritmo di interpolazione SURF (Van Horik e Goodchild, 1977); e da Menozzi *et al.* (1978) e Piazza *et al.* (1981) che hanno utilizzato le componenti principali ottenute dall'analisi di 39 loci elettroforetici per rappresentare sinteticamente su una mappa geografica tricromica la variazione delle frequenze geniche umane in Europa e nel mondo.

Per quanto riguarda la categorizzazione delle aree, una situazione ideale è rappresentata da caratteri che variano in modo discontinuo tra le diverse regioni e sono omogenei all'interno delle regioni stesse; questa procedura non ammette interpolazioni fra regioni. Gabriel e Sokal (1969) hanno elaborato il metodo STP (Symultaneous Test Procedure) per individuare insieme omogenei di località sulla base di tutte le ripartizioni possibili del territorio.

5. DISTANZE E RELAZIONI

Uno degli aspetti più rilevanti di molti studi sulla variazione geografica è costituito dall'interpretazione sistematica (sensu Simpson, 1961) dei risultati derivanti dall'analisi multivariata. Le distanze tra le OTU, calcolate su una base multivariata, quantificano la similitudine/dissimilitudine fra i taxa e permettono, poste determinate condizioni, la formulazione di ipotesi più propriamente filogenetiche.

Sneath e Sokal (1973) descrivono diversi modelli per quantificare distanze fra le OTU, e qui di seguito sono discussi i modelli più utilizzati in morfometria multivariata.

Il modello più semplice è rappresentato dalla distanza euclidea estesa ad uno spazio a n dimensioni (Sokal, 1961; Rohlf e Sokal, 1962):

Nell'analisi morfometrica ed allozimica di *Macrobrachium rosebergii*, Linderfelser (1984) ha costruito matrici di distanze euclidee sulla base dei punteggi discriminanti delle prime due variate canoniche.

Le distanze euclidee pongono però seri problemi quando applicate ad uno spazio multivariato, in quanto non sempre questo spazio è descritto da assi fra loro ortogonali (Rohlf e Sokal, 1965). Questo problema è superato dall'adozione di metriche diverse da quella euclidea, come la distanza di Mahalanobis, la distanza di Manhattan e le sue derivate.

Le distanze di Mahalanobis (1936), definite anche come distanze quadratiche generalizzate, sono calcolate massimizzando la differenza tra coppie di medie, in rapporto a particolari combinazioni lineari di caratteri, che presen-

tano una varianza tra gruppi maggiore della varianza entro gruppi. La distanza di Mahalanobis permette di eliminare la ridondanza dei caratteri tenendo conto della correlazione intra-OTU tra i caratteri stessi (Thorpe, 1976) e costituisce la procedura più appropriata nel caso di campioni sufficientemente grandi (Cherry *et al.*, 1982).

Hausser (1984) ha definito su questa base le distanze morfometriche in rapporto alla variazione geografica del genere *Sorex* in Europa, utilizzando i centroidi delle funzioni discriminanti; Blondel *et al.* (1984) hanno usato le distanze di Mahalanobis per quantificare le differenze tra regioni e substrati nello studio sulla convergenza morfologica di comunità ornitiche mediterranee.

Durante l'elaborazione delle funzioni discriminanti, la distanza di Mahalanobis viene utilizzata per assegnare ogni individuo al centroide più vicino (Sadocchi, 1981) e rappresenta la procedura statistica descrittiva più funzionale per classificare correttamente gli esemplari. Illuminante in questo senso è il lavoro di Andrews (1972) sulla collocazione di *Proconsul africanus*, che viene identificato come *outlier*, non appartenente cioè a nessuno dei gruppi definiti a priori e che assume quindi una posizione oscillante fra i gruppi «ominidi» e «scimmie».

Anche le distanze di Manhattan non sono distanze euclidee, in quanto non sono invarianti rispetto alla rotazione dello spazio definito dai caratteri. Queste distanze si basano sulla metrica di Minkowsky, e di questo modello sono state proposte diverse varianti, tra cui la metrica di Canberra (Sneath and Sokal 1973). Le distanze di Manhattan sono generalmente utilizzate in tassonomia numerica per le strette relazioni con il criterio della parsimonia (Farris, 1970) e con la teoria della sistematica filogenetica (Farris *et al.*, 1970); esse sono particolarmente raccomandate da Farris (1972) come matrice di base da cui partire per costruire diagrammi di relazione direzionali.

Un'ampia discussione e confronto tra i diversi metodi di elaborazione di matrici di distanze morfometriche è ancora in corso (Cherry *et al.*, 1972; Hafner *et al.*, 1984; Wilson *et al.*, 1984). Le matrici di distanze derivate dai diversi modelli sono utilizzate per elaborare diagrammi di relazione fra le OTU che ne evidenzino le interconnessioni e/o i rapporti filogenetici.

L'analisi dei cluster può essere considerata senza dubbio il metodo più oggettivo per rappresentare relazioni non direzionali fra le OTU. I metodi utilizzati più frequentemente nella analisi dei cluster sono compresi nel gruppo SAHN (Sequential Agglomerative Hierarchic Non-overlapping, Sneath e Sokal, 1973) e sono costituiti dal metodo del legame singolo, del legame completo e del legame medio. Il legame medio comprende a sua volta diversi metodi: metodo della media aritmetica (UPGMA, WPGMA; Sokal e Michener, 1958) e metodo dei centroidi (Lance e Williams, 1967). Rohlf (1970) ha definito le basi teoriche di applicazione della procedura di analisi dei cluster; l'importante aspetto comune ai diversi metodi consiste nel delimitare i gruppi tassonomici a diversi livelli e in modo obiettivo, a partire dalla matrice di distanze. Alcuni esempi sono contenuti nei lavori di Atchley (1981), Hausser (1984), Fujii (1969), Soulé (1967).

I cluster comunque non rappresentano diagrammi di relazioni filogenetiche e in questi ultimi anni si è assistito allo sviluppo di tecniche di analisi dei dati che permettano di proporre ipotesi più propriamente filogenetiche. Naturalmente queste tecniche fanno direttamente riferimento ai metodi della tassonomia numerica classica (Sneath e Sokal, 1973; Nelson e Platnick, 1981; Wiley, 1981). Per rappresentazione di ipotesi filogenetiche ci riferiamo ai «grafici direzionali» (*networks*, Busacker e Saaty, 1965), cioè a quella speciale famiglia di diagrammi di relazione fra le OTU che identificano una direzione tale che la relazione $a \rightarrow b$ è diversa da $b \rightarrow a$ (Fig. 6). Un diagramma di relazioni direzionali (*directed tree*) è costituito da un *network* che possiede una direzione definita a partire da una base (radice del diagramma), attraverso i punti di biforcazione (internodi) verso tutti gli altri vertici (nodi). In altri termini, il *directed tree* è un *network* ruotato su un vertice definito che viene identificato con la radice.

La costruzione di *network* e *trees* richiede che, per ogni carattere, vengano definiti stati discontinui e diversi (*character states*). In questo approccio sono compresi i modelli di parsimonia (Camin e Sokal, 1965; Hennig, 1966), di compatibilità (Estabrook *et al.*, 1977; Estabrook, 1983) e della massima probabilità (Felsestein, 1983).

Tutti questi metodi richiedono una definizione a priori dello stato di ogni carattere; nel caso di caratteri quantitativi, quali quelli trattati nella morfometria multivariata, il problema consiste nella definizione delle OTU in termini di *character states*.

La soluzione del problema può essere affrontata seguendo diverse strategie. Corti *et al.* (1985) hanno utilizzato i punteggi dei centroidi sui primi tre assi canonici per elaborare un Wagner network (Farris, 1979) che illustra i rapporti fra le tre specie di talpe italiane. Thorpe (1989) ha utilizzato i punteggi delle coordinate principali (basate su una matrice di distanze tassonomiche) nell'elaborazione di Wagner networks che illustrano i rapporti filogenetici delle popolazioni adriatiche di *Podarcis sicula* e *P. melisellensis*.

Nell'utilizzazione diretta dei risultati dell'analisi multivariata per costruire diagrammi di relazioni filogenetiche, si deve tenere presente che le distanze fra le OTU così identificate devono essere considerate in termini relativi: l'aggiunta o l'esclusione di un gruppo dall'analisi ne altera i rapporti di distanza relativa. La soluzione consiste nell'utilizzare una metodologia multivariata più appropriata: l'analisi delle componenti principali a gruppi multipli (Thorpe *et al.*, 1982; Thorpe, 1983) (Fig. 8), in cui le distanze fra i gruppi non sono influenzate dalla variazione del numero stesso dei gruppi.

Un'alternativa a questi metodi è costituita dalla trasformazione dei caratteri quantitativi in caratteri «qualitativi». Mickevich e Johnson (1976) hanno suggerito l'uso di *gap-coding*, in cui i diversi stati di un carattere sono determinati in base alla presenza di lacune (pari alla deviazione standard o a qualche multiplo di essa) nella curva di distribuzione dei valori di ogni variabile continua.

Un metodo più veloce è descritto da Almeida and Bisby (1983), che hanno

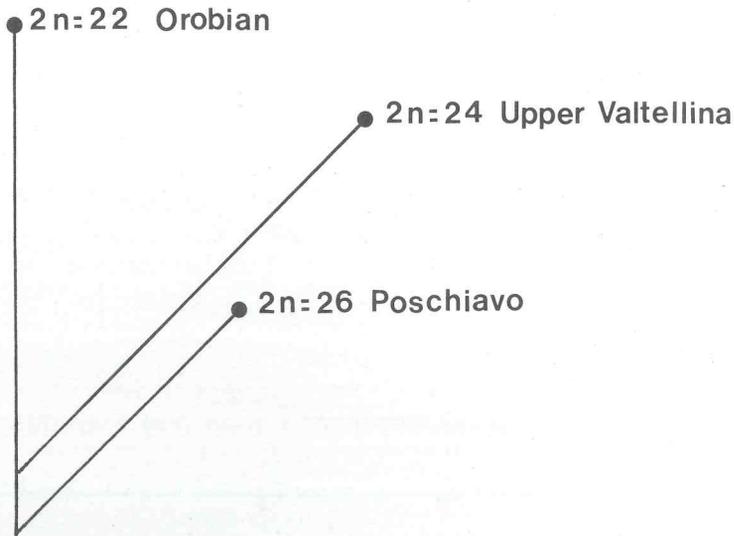


FIG. 8 - Cladogramma (Wagner tree) rappresentante le relazioni filetiche tra popolazioni robertsoniane a numero cromosomico ridotto di *Mus domesticus* (Thorpe *et al.*, 1982). Questo Wagner tree è stato calcolato sulla matrice di distanze di Manhattan fra le medie delle popolazioni (Multiple Group Principal Component) ed è ruotato su una popolazione a 40 cromosomi (*outgroup*), assunta come ancestrale. La lunghezza totale del Wagner tree è di 28.7 e la differenza fra la distanza patristica e la distanza di Manhattan è di 0.77. Questo modello di relazione cladistica fra le popolazioni, basato su caratteri *size-independent* della mandibola, è congruente con quello elaborato in base alla riduzione del numero di plodi per fusioni centriche.

identificato gli stati di ogni carattere in corrispondenza di eventuali deformazioni della curva di frequenza dei caratteri stessi. Simon (1983a) propone una procedura considerata più corretta (*homogeneous subset coding*): ogni carattere è analizzato singolarmente attraverso test di comparazione multipla a posteriori (Test simulato sulla somma dei quadrati, la procedura GT-2, il metodo T, tutti discussi in Sokal and Rohlf, 1981); il risultato di questi test è rappresentato in diagrammi a barre in cui sono raggruppati sottoinsiemi omogenei di OTU.

Rispetto alle due precedenti procedure, quest'ultima ha il vantaggio di utilizzare tutta l'informazione statisticamente significativa presente nei dati. Esempi di applicazione di questa procedura sono offerti da Thorpe (1984b) su *Natrix natrix* e da Simon (1983b) (*Cicades*).

BIBLIOGRAFIA

- ALBRECHT G.H., (1980) - *Multivariate analysis and the study of form, with special reference to canonical variate analysis* - Amer. Zool., **20**, 679-693.
 ALMEIDA M.T., BISBY F.E., (1983) - *A quick method for making multistate characters out of continuous measurements* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer-Verlag, 383.
 ANDREWS D.F., (1972) - *Plots of high-dimensional data* - Biometrics, **28**, 125-130.
 ANDREWS D.F., GNANADESIKAN R., WARNER R., (1973) - *Methods for assessing multivariate normality* - In: Multivariate analysis, vol. III, Krishnaiah P.R., (Ed.), Academic Press, 95-116.
 ATCHELEY W.R., GASKINS C.T., ANDERSON D., (1976) - *Statistical properties of ratios. I. Empirical results* - Syst. Zool., **25**, 137-148.
 ATCHLEY W.R., (1981) - *Chromosomal evolution and morphometric variability in the thelytokous insect*

- Warramaba virgo Key - In: Evolution and speciation, Atchley W.R. and Woodruff D.S. (Eds.), Cambridge Univ. Press, 371-397.
- ATCHLEY W.R., RUTLEDGE J.J., COWLEY D.E., (1981) - *Genetic components of size and shape. II. Multivariate covariance patterns in the rat and mouse skull* - Evolution, **35**, 1037-1055.
- ATCHLEY W.R., (1983) - *Some genetic aspects of morphometric variation* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer-Verlag, 346-363.
- BLONDEL J., VUILLEUMIER F., MARCUS L.F., TEROUANNE E., (1984) - *Is there an ecomorphological convergence among mediterranean birds communities of Chile, California and France?* - Evol. Biol., **18**, 141-123.
- BOOKSTEIN G.D., (1982) - *Foundations of morphometrics* - Ann. Rev. Ecol. Syst., **13**, 451-470.
- BUSACKER R.G., SAATY T.L., (1965) - *Finite graphs and networks* - Mac Grow Hill, New York.
- BURNS J.C., CHOATE J.R., ZIMMERMANN E.G., (1985) - *Systematic relationships of pocket gophers (genus Geomys) on the central great plains* - J. Mamm., **66**, 102-118.
- CAMIN J.H., SOKAL R.R., (1965) - *A method for reducing branching sequences in phylogeny* - Evolution, **19**, 311-326.
- CAPANNA E., CIVITELLI M.V., CORTI M., LOY A., FILIPPUCCI M.G., BULLINI L., NASCETTI G., PETROV B., (1985) - *Genetic and morphological studies on the european moles (genus Talpa)* - Proc. IV ITC, Edmonton (Canada), 1985.
- CHERRY L.M., CASE S.M., KUNKEL J.G., WYLES J.S., WILSON A.C., (1982) - *Body shape metrics and organismal evolution* - Evolution, **36**, 914-933.
- COMSTOCK R.E., (1973) - *Growth in mice* - Genetics, **74**, 51.
- CORRUCCINI R.S., (1975) - *Multivariate analysis in biological anthropology: some considerations* - J. of Hum. Evol., **4**, 1-19.
- CORTI M., LOY A., AZZAROLI M.L., CAPANNA E., (1985) - *Multivariate analysis of osteometric traits in Italian moles (genus Talpa)* - Z. Säugetierkunde, **50**, 12-17.
- CORTI M., THORPE R.S., SOLA L., SBORDONI V. and CATAUDELLA S., (1987) - *Multivariate morphometrics in aquaculture: a case study of six stocks of the common carp Cyprinus carpio from Italy*. Can. J. Fish Aquat. Sci., in press.
- DAHL C., WOLD S., NIELSEN L.T., NILSSON C., (1984) - *A SIMCA pattern recognition study in taxonomy: claw shape in mosquitoes (Culicidae, Insecta)* - Syst. Zool., **33**, 355-369.
- ENDLER J.A., (1977) - *Geographic variation, speciation and clines* - Princeton Univ. Press.
- ESTABROOK J.F., STARCK J.G., FIALA K.L., (1977) - *An application of compatibility analysis to the Bkackiths' data on Orthopteroid insects* - Syst. Zool., **26**, 269-276.
- ESTABROOK J.F., (1983) - *The causes of incompatibility* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer-Verlag, 279-295.
- FARRIS J.S., (1970) - *Methods for computing Wagner trees* - Syst. Zool., **19**, 83-92.
- FARRIS J.S., KLUGE A.G., ECKARDT M.J., (1970) - *A numerical approach to phylogenetic systematics* - Syst. Zool., **19**, 172-189.
- FARRIS J.S., (1972) - *Estimating phylogenetic trees from distance matrices* - Amer. Nat., **106**, 645-668.
- FELSESTEIN J., (1983) - *Methods for inferring phylogenies: a statistical view* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer-Verlag, 215-234.
- FERSON S., ROHLF F.J., KOHEN R.K., (1985) - *Measuring shape variation of two dimensional outlines* - Syst. Zool., **34**, 59-68.
- FESTING M., (1972) - *Mouse strain identification* - Nature, **238**, 351-352.
- FUJII K., (1969) - *Numerical taxonomy of ecological characteristics and niche concept* - Syst. Zool., **18**, 151-153.
- GABRIEL K.R., SOKAL R.R., (1969) - *A new statistical approach to geographic variation analysis* - Syst. Zool., **18**, 259-278.
- GNANADESIKAN R., (1977) - *Methods for statistical data analysis of multivariate observations* - John Wiley and sons, N.Y.
- GOULD S.J., (1966) - *Allometry and size in ontogeny and philogeny* - Biol. Rev., **41**, 587-640.
- GOULD S.J., JOHNSTON R.F., (1972) - *Geographic variation* - Ann. Rev. of Ecol. and Syst., **3**, 457-498.
- GOULD S.J., (1984) - *Covariance sets and ordered geographic variation in Cerion from Amba, Bonarie and Curasao: a way of studying non adaptation* - Syst. Zool., **33**, 217-237.
- GOWER J.C., ROSS G.J.S., (1969) - *Minimum spanning tree and single linkage cluster analysis* - App. Statist., **18**, 54-64.
- HAFNER M.S., REMSEN J.V. JR, LANYON S.M., (1984) - *Bird versus mammal morphological diversity* - Evolution, **38**, 1154-1156.
- HANDFORD P., (1983) - *Continental patterns of morphological variation in a South American sparrow* - Evolution, **37**, 920-930.
- HAUSSER J., JAMMOT D., (1974) - *Étude biometrique des machoires chez les Sorex du groupe araneus en Europe continentale (Mammalia, Insectivora)* - Mammalia, **38**, 324-343.
- HAUSSER J., (1984) - *Genetic drift and selection: their respective weights in the morphological and genetic*

- differentiation of four species of shrews in Southern Europe (Insectivora, Soricidae)* - Z. Zool. Syst. Evolut-forsch., **22**, 302-320.
- HENNIG W., (1966) - *Phylogenetic systematics* - University of Illinois Press, Urbana.
- HUXLEY J.S., (1983) - *Clines: an auxiliary taxonomic principle* - Nature, **142**, 219-220.
- JOHNSTON R.F., SELANDER R.K., (1971) - *Evolution in the house sparrow. II. Adaptive differentiation in North American populations* - Evolution, **25**, 1-28.
- JOHNSTON R.F., SELANDER R.K., (1973) - *Evolution in the house sparrow. III. Variation in size and sexual dimorphism in Europe and North and South America* - Amer. Nat., **107**, 373-390.
- JOLICOEUR P., MOSIMAN J.E., (1960) - *Size and shape variation in the painted turtle. A principal component analysis* - Growth, **24**, 339-354.
- JOLICOEUR P., (1963) - *The multivariate generalization of the allometry equation* - Biometrics, **19**, 497-499.
- KENDALL M.J., (1973) - *The basic problems of cluster analysis* - In: Discriminant analysis and applications, Cacoullos T. (Ed), Academic Press, N.Y., 179-198.
- LAMBERT D.M., PATERSON H.E., (1982) - *Morphological resemblance and its relationships to genetic distance measures* - Evol. Theory, **5**, 291-300.
- LANCE G.N., WILLIAMS W.T., (1967) - *A general theory of classificatory sorting strategies. I. Hierarchical Systems* - Computers, **9**, 373-380.
- LESSIOS H.A., (1981) - *Divergence in allopatry: molecular and morphological differentiation between sea urchins separated by the isthmus of Panama* - Evolution, **35**, 618-634.
- LEWONTIN R.C., (1984) - *Detecting population differences in quantitative characters as opposed to gene frequencies* - Amer Nat., **123**, 115-124.
- LINDENFELSER M.E., (1984) - *Morphometric and allozymic congruence: evolution in the prawn Macrobrachium rosenbergii (Decapoda: Palaemonidae)* - Syst. Zool., **32**, 195-204.
- MAHALANOBIS P.C., (1963) - *On the generalized distance in statistics* - Proc. Nat. Inst. Sci. India, **2**, 49-55.
- MARCUS L.F., (1983) - *Automated measurement with portable microcomputers* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer Verlag, 620-624.
- MAYR E., (1942) - *Systematics and the origin of species from a viewpoint of a zoologist* - Columbia Univ. Press, N.Y.
- MAYR E., (1963) - *Animal species and evolution* - Princeton Univ. Press.
- MC NEILL I., (1978) - *Purposeful phenetics* - Syst. Zool., **28**, 465-482.
- MENOZZI P., PIAZZA A., CAVALLI SFORZA L., (1978) - *Synthetic maps of human gene frequencies in Europeans* - Science, **201**, 786-792.
- MICKEVICH M.F., JOHNSON M.S., (1976) - *Congruence between morphological and allozyme data in evolutionary inference and character evolution* - Syst. Zool., **25**, 260-270.
- NEFF N.A., MARCUS L.E., (1980) - *A survey of multivariate methods for systematics* - New York: privately published.
- NELSON G., PLATNICK N., (1981) - *Systematic and biogeography: cladistics and vicariance* - Columbia Univ. Press, N.Y.
- NEVO E., CORTI M., HETH G., BEILES A., SIMSON S., (1978) - *Chromosomal polymorphisms in subterranean mole rats: origins and their evolutionary significance*. Biol. J. Linn. Soc., in press.
- OXNARD C.E., (1978) - *One biologist's view of morphometrics* - Ann. Rev. of Ecol. and Syst., **9**, 219-241.
- PIAZZA A., MENOZZI P., CAVALLI SFORZA L., (1981) - *Synthetic gene frequencies maps of man and selective effects of climate* - Proc. Nat. Acad. Sci. USA, **78**, 2638-2642.
- REYMENT R.A., (1981) - *Multivariate morphometrics* - Academic Press, N.Y.
- ROHLF F.J., SOKAL R.R., (1962) - *The description of taxonomic relationships by factor analysis* - Syst. Zool., **11**, 1-16.
- ROHLF F.J., SOKAL R.R., (1965) - *Coefficients of correlation and distances in numerical taxonomy* - Univ. Kansas Sci. Bull., **45**, 3-27.
- ROHLF F.J., (1970) - *Adaptive hierarchical clustering schemes* - Syst. Zool., **19**, 58-82.
- ROHLF F.J., FERSON S., (1983) - *Image analysis* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer Verlag, 583-599.
- SADOCCHI S., (1981) - *Manuale di statistica multivariata per le scienze sociali* - F. Angeli, Milano.
- SCALA C. and LOVARI S., (1984) - *Revision of Rupicapra genus. II. Skull and horn statistical comparison of Rupicapra rupicapra ornata and R. rupicapra pyrenaica chamois* - Boll. Zool., **51**, 285-294.
- SCHNELL G.D., (1970) - *A phenetic study of the suborder Lari (Aves). II. Phenograms, discussion, and conclusions* - Syst. Zool., **19**, 264-302.
- SCHONEWALL-COX C.M., BAYLESS J.W., SCHONEWALL J., (1985) - *Cranial morphometry of Pacific coast elk Cervus elaphus* - J. Mamm., **66**, 63-74.
- SELANDER R.K., (1976) - *Genetic variation in natural populations* - In: Molecular evolution, Ayala F.J. (Ed.), Sinauer Ass. Inc., 21-45.
- SIMON C., (1983a) - *A new coding procedure for morphometric data with an example from periodical cicada wing veins* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer Verlag, 378-382.

- SIMON C., (1983b) - *Morphological differentiation among broods of 13- and 17- ears periodical Cicadas* - Evolution, **37**, 104-115.
- SIMPSON G.G., (1961) - *Principles of animal taxonomy* - Columbia Univ. Press, N.Y.
- SNEATH P.H.A., SOKAL R.R., (1973) - *Numerical Taxonomy* - W.H. Freeman and Co., San Francisco.
- SNYDER P.T., LINTON M.C., (1984) - *Population structure in black flies: allozymic and morphological estimates from Prosimulium miztum and P. fuscum (Diptera: Simuliidae)* - Evolution, **38**, 942-956.
- SOKAL R.R., MICHENER C.D., (1958) - *A statistical method for evaluating systematic relationships* - Univ. Kansas Sci. Bull., **38**, 1490-1438.
- SOKAL R.R., (1961) - *Distance as a measure on quantitative systematics* - Evolution, **13**, 420-423.
- SOKAL R.R., ROHLF F.J., (1969) - *Biometry, the principles and practice of statistics in biological researches* - W.H. Freeman and Co., San Francisco.
- SOKAL AND ROHLF, (1981) - *Biometry. 2nd edition* - W.H. Freeman and Co., S. Francisco.
- SOKAL R.R., MENOZZI P., (1982) - *Spatial autocorrelation of HLA frequencies in Europe support demic diffusion of early farmers* - Amer. Nat., **119**, 1-17.
- SOKAL R.R., (1983) - *Analysing character variation in geographic space* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer Verlag, 385-403.
- SOULÉ M., (1967) - *Phenetic of natural populations. I. Phenetic relationship on insular populations of the side-blocked lizard* - Evolution, **21**, 584-591.
- STRAUSS R.E., BOOKSTEIN F.L., (1982) - *The truss: body form reconstruction in morphometrics* - Syst. Zool., **31**, 113-135.
- THOMAS O., (1905) - *Suggestions for the nomenclature of the cranial length of Mammals* - Proc. Biol. Soc. Wash., **18**, 191-196.
- THORPE R.S., (1975) - *Biometric analysis of incipient speciation in the ringed snake Natrix natrix (L.)* - Experientia, **31**, 180-182.
- THORPE R.S., (1976) - *Biometric analysis of geographic variation and racial affinities* - Biol. Rev., **51**, 407-452.
- THORPE R.S., (1980) - *Microevolution and taxonomy of European reptiles with particular reference to the grass snake Natrix natrix and the wall lizard Podarcis sicula and P. melisellensis* - Biol. J. of Linn. Soc., **14**, 215-233.
- THORPE R.S., CORTI M., CAPANNA E., (1982) - *Morphometric divergence of Robertsonian population/species of Mus: a multivariate analysis of size and shape* - Experientia, **38**, 920-923.
- THORPE R.S., (1983) - *A review of the numerical methods for recognising and analysing racial differentiation* - In: Numerical Taxonomy, Felsestein J. (Ed.), Springer Verlag, 404-423.
- THORPE R.S., (1984a) - *Primary and secondary transition zones in speciation and population differentiation: a philogenetical analysis of range-expansion* - Evolution, **38**, 233-243.
- THORPE R.S., (1984b) - *Coding morphometric characters for constructing distance Wagner networks* - Evolution, **38**, 244-255.
- TURNER B.J., (1974) - *Genetic divergence of Death Valley pupfish species: biochemical versus morphological evidence* - Evolution, **28**, 281-294.
- VAN HORIK M., GOODCHILD G., (1977) - SURF. General surface interpolation package. Version 3 - Dept. of Geography, Univ. West Ontario, London, Canada.
- VIGNA TAGLIANTI A., IACOBONE M.G., LOY A., (1984) - *Osservazioni sistematiche e zoogeografiche sull'orso bruno dell'Appennino centrale* - Boll. Zool., **51**, (suppl.), 113.
- WHITE M.J.D., (1978) - *Modes of speciation* - W.H. Freeman, San Francisco.
- WILEY E.D., (1981) - *Phylogenetics: the theory and practice of phylogenetic systematic* - J. Wiley and sons, N.Y.
- WILLIAMSON P.G., (1981) - *Paleontological documentation of speciation in Cenozoic molluscs from Turkana basin* - Nature, **293**, 437-443.
- WILSON A.C., KUNKEL J.G., WYLES J.S., (1984) - *Morphological distance: an encounter between two perspectives in evolutionary biology* - Evolution, **38**, 1156-1159.
- YOSHIYAMA R.M., SASSAMAN C., (1983) - *Morphological and allozymic variation in the stichaeid fish Anoplarchus purpureus* - Syst. Zool., **32**, 52-71.
- ZINK R.M., (1983) - *Evolutionary and systematic significance of temporal variation in the fox sparrow* - Syst. Zool., **32**, 223-238.