Z. Säugetierkunde **61** (1996) 304–319
© 1996 Gustav Fischer, Jena

ZEITSCHRIFT FÜR
SÄUGETIERKUNDE
INTERNATIONAL JOURNAL
OF MAMMALIAN BIOLOGY

Multivariate morphometric analysis of European species of the genus *Mus* (Mammalia, Muridae)

By M. Macholán

Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Brno, Czech Republic

Receipt of Ms. 24. 10. 1995
Acceptance of Ms. 10. 04. 1996

Abstract

Systematic relationships among 12 population groups of five European species of house mice were studied using multivariate morphometric methods. Multiple group principal component analysis (MGPCA) was used to assess the contribution of the size component to the total variation. It was shown that part of the 'shape' information may be resident in the first principal component and, likewise, subsequent components may contain residual 'size' information. Hence, removing the 'size-vector' should be done with caution and after an appropriate examination of the data. Canonical variate analysis (CVA) revealed similar results both on 'size-in' and 'size-out' MGPCA scores. The first canonical variate discriminated between the aboriginal and commensal mice lineages, while the second axis identified species clusters. The third canonical variate separated groups of populations within commensal species. Both CVA and cluster analysis demonstrated that (i) *M. macedonicus* and *M. spretus* are morphologically more similar to each other than either species is to *M. spicilegus*; (ii) the distance between *M. musculus* and *M. domesticus* is similar to distances among aboriginal (= outdoor) species; (iii) interpopulation distance is relatively high compared to interspecific relationships.

Introduction

For many years, the systematics of house mice of the genus *Mus* has been an intricate puzzle. In 1943, Schwarz and Schwarz tried to simplify the taxonomy by condensing more then 130 known scientific names of *Mus* into a single species, *Mus musculus*. They recognized 15 subspecies and proposed the evolutionary scenario of a multiple origin of commensal mouse taxa from exoanthropic forms. This approach has been later followed by other authors (Ellerman and Morrison-Scott 1951; Serafiński 1965; Corbet 1978; Reichstein 1978). However, the concept oversimplified hierarchical relationships among house mice, ignoring such phenomena as absence of interbreeding between some taxa.

An advent of biochemical and molecular studies on free-living small mammals some 15 years ago shed light onto the systematic interrelationships and evolutionary history of the house mouse complex (see Boursot et al. 1993; Sage et al. 1993, for recent reviews), falsifying Schwarz and Schwarz's (1943) concept. It has been shown that there are five taxa of house mice in Europe representing two major lineages (Marshall and Sage 1981; Thaler et al. 1981; Bonhomme et al. 1984): one lineage consists of three aboriginal (Sage 1981), 'outdoor', species (*M. spretus*, *M. spicilegus*, *M. macedonicus*), while the other includes commensal, 'indoor', taxa (*M. domesticus*, *M. musculus*). Note that the latter two may be regarded as subspecies of a single species, *M. musculus*, by some authors (e.g. Bonhomme and Guénet 1989; Auffray et al. 1990 a). Following Marshall, (1981), Ferris et al. (1983), and SAGE et al. (1993) and objectives given therein, however, all the European house mouse taxa are treated as distinct species throughout this study.

Unequivocal genetic identification of investigated animals was used to establish morphological discrimination criteria between some of the mouse taxa (DARVICHE and ORSINI 1982; ORSINI et al. 1983; KRAFT 1985; KRATOCHVÍL 1986 a, b; LYALYUKHINA et al. 1991). An array of studies dealing with the morphometrics of house mice has been published in last two decades; however, these studies were limited in the number of variables involved (i. e. uni- or bivariate analyses) and/or in the number of taxa studied (SANS-COMA et al. 1979; DARVICHE and ORSINI 1982; ORSINI et al. 1983; PALOMO et al. 1983; LYALYUKHINA et al. 1991). Several studies employed multivariate methods, yet they were either focused on one or a few species (THORPE et al. 1982; DAVIS 1983; SCRIVEN and BAUCHAU 1992) or did not take into account the relative contribution of size and shape to the total variation observed (ENGELS 1980, 1983; GERASIMOV et al. 1990; LAVRENCHENKO 1994).

In a previous study (MACHOLÁN 1996), morphometric and morphological relationships among populations of all five mouse species were evaluated and the taxonomic status of central European mouse populations was documented. This analysis indicated that the assessment was obscured by the 'size' component, which was, to some extent, independent of the age structure of a population (see also THORPE and LEAMY 1983); therefore, size-adjusting of the data was suggested. Nevertheless, as size and shape are assumed to be essentially multivariate concepts (HUMPHRIES et al. 1981; THORPE and LEAMY 1983) because one measurement cannot encompass the various facets of length, width, etc., an appropriate multivariate statistics including neglecting the size influence should be employed.

This study focuses on the multivariate analysis of morphometric relationships among European house mouse populations, including the relative importance of 'size' and 'shape' components in their morphological differentiation.

Material and methods

Mouse skulls used in this study are deposited in collections of the Institute of Landscape Ecology in Brno, National Museum in Prague, Museum of Natural History in Vienna, Institute of Zoology in Kiev, Charles University in Prague, University of Lausanne and University of Montpellier.

A total of 297 skulls of five house mouse species was analysed. The material was pooled into 13 groups (populations): 'GR' (*Mus macedonicus*, Greece, n = 25); 'AUT' (*M. spicilegus*, Austria, n = 20); 'UKR' (*M. spicilegus*, Ukraine and Moldavia, n = 23); 'SPR' (*M. spretus*, France, Spain, Morocco, n = 24); 'DA' (*M. domesticus*, Albania, n = 10); 'DCH' (*M. d.*, Switzerland, n = 24); 'DWM' (*M. d.*, western Mediterranean islands, n = 20); 'MC' (*M. musculus*, Bohemia, n = 25); 'MM' (*M. m.*, Moravia, n = 25); 'MS' (*M. m.*, Slovakia, n = 25); 'MH' (*M. m*, Hungary, n = 25); 'MU' (*M. m.*, Ukraine, n = 25); 'MSP' (*M. domesticus/musculus* hybrids, W Bohemia, n = 25). A detailed description of specific localities is given elsewhere (MACHOLÁN 1996) with the only exception of the *M. spicilegus* population from Ukraine and Moldavia which was added to the original material. This sample consisted of mice from the Chernomorskiy zapovednik Reserve (n = 4), Tyaginka, Cherson (n = 6), Golo-Pristan'sk (n = 3), Kirovograd (n = 3), Melovsk (n = 3), Kishinev (n = 4), and Nikolaev (n = 2).

This study is based on 18 cranial and dental variables (MACHOLÁN 1996); namely, width of the upper ramus of the zygomatic process of maxilla (A); width of the zygomatic process of maxilla (B); condylobasal length (LCb); basal length (LB); rostral width (LaR); width of the skull per bullae (LaC); zygomatic width (LaZ); height of the braincase (hC); length of the diastema (LD); length of the first lower molar (LM1i); M_1 width (LaM1i); M_2 length (LM2i); M_3 length (LM3i); M_3 width (LaM3i); length of the lower toothrow (LM13i); length of the first upper molar (LM1s); M^2 length (LM2s); length of the upper toothrow (LM13s).

Only adult individuals were measured. In order to estimate age of the animals under study, the condition of their reproductive organs, their weight (LAURIE 1946; PELIKÁN 1981), and the level of abrassion of their molars (KELLER 1974) were assessed; where possible, a combination of the three approaches was taken into account (see MACHOLÁN 1996, for details). All individuals of doubtful age were exluded from subsequent investigations.

Multiple group principal component analysis (MGPCA, THORPE 1983 a), based on a pooled within-group variance-covariance matrix, was used in order to assess the contribution of within-group components to the between-group discrimination. The 'size' vector was searched in order to be eventually extracted from the data. This 'size' component is generally the eigenvector corresponding to the first principal component; however, three conditions have to be met for such an assumption: an eigenvector expressing general size should have coefficients of the same sign (1) and 'similar' magnitude (2); and the first principal components within localities should have the same orientation (3). The latter condition can be tested by comparing the first eigenvectors across localities.

Since substantial between-character differences due to different scales in individual variables was expected, the contribution of each character to a component was compared by computing the pooled within-population correlation between the character and the component score according to the formula (THORPE 1983 a):

$$r_{ij} = \frac{a_{ij} \cdot \lambda_j^{1/2}}{s_i},$$

where r_{ij} is the pooled within-group (within-population) correlation between the ith character and the jth eigenvector, a_{ij} is the coefficient for ith character for the normalized jth eigenvector, λ_j is the latent root (eigenvalue) of the jth eigenvector and s_i is the pooled within-group standard deviation of the ith character. Before computing the correlations, the eigenvectors were normalized so that each component coefficient was divided by $\sqrt{\Sigma a_{ij}^2}$, where a_{ij} is as defined above and k is the number of characters.

Correlation coefficients were then compared and their significance was tested (SOKAL and ROHLF 1981; THORPE and LEAMY 1983). Since for ρ close to ±1.0, the distribution of sample values of r is markedly asymmetrical, we have to transform r to a function z; standard normal deviate value t_s is then defined as z/σ_z, where $\sigma_z = \sqrt{\frac{1}{\Sigma(n_i - 3)}}$, and $z = {}^1/_2 \ln \frac{1+r}{1-r}$; r is the correlation coefficient as defined above, k is the number of populations, and n_i is the sample size of the ith population. Since z is approximately normally distributed and we are using a parametric standard deviation, t_s is compared with $t_{\alpha[\infty]}$ (where $\alpha = 0.01$).

Two techniques were employed in order to extract the size vector: one produces new data as principal component scores with the first eigenvector removed (THORPE 1983 b), whereas the other is based on BURNABY'S (1966) adjustment as suggested in ROHLF and BOOKSTEIN (1988).

Population interrelationships were assessed by subjecting the component scores to canonical variate analysis (CVA, FISHER 1936). This multivariate ordination method separates groups so that between-group variation is maximized while within-group variation is minimized (CAMPBELL and ATCHLEY 1981). As multiple-group PCA uses pooled within-group covariances, CVA performed on all of the MGPCA component scores ('size-in' analysis) gives the same results as CVA on the original data (THORPE et al. 1982; THORPE 1983 a). CVA computed on MGPCA component scores with the 'size' vector extracted ('size-out' analysis) revealed the same results as CVA performed on BURNABY-adjusted data.

Matrices of Mahalanobis generalized distances D^2, computed as a part of canonical variate analysis, were employed both in the Mantel test comparing the results of 'size-in' and 'size-out' multivariate analyses, and subjected to cluster analysis.

The System for Statistics (SYSTAT, Release 5.02, WILKINSON 1990) and Numerical Taxonomy System (NTSYS-pc, Version 1.60, ROHLF 1990) packages were used for all the statistical analyses.

Results

As stated above there are three assumptions for the first principal component, extracted from the pooled within-population covariance matrix, to be treated as the 'size' vector. The first principal component within localities appeared to be of the same orientation as substantiated by checking the signs of the first eigenvectors for each locality. Furthermore, the coefficients corresponding to the first principal axis were all of the same sign (see the first column in Tab. 1). However, differences in their magnitude were strikingly high. Since the (pooled) variance-covariance matrix was used, the relative magnitude of the coefficients was dependent on the variances of the original data, i. e., on the scale of the respective characters. Therefore, two transformation techniques were used in order to decrease the differences in the variances: firstly, the variates were converted to logarithms, and secondly, the data were normalized by subtracting the mean and dividing by the standard deviation as provided by the standard SYSTAT routine.

Either transformation of data is only feasible under the expectation of a general improvement in linearity. Since non-linear relations between variables would result in a lower inter-character correlation, an improvement in linearity should generally be apparent by a larger first eigenvalue in the correlation matrix. As shown in table 1, both the transformations yielded a slight (although insignificant) decrease in curvilinearity between the variables. If we compare the total variance explained by the first principal component for the three data sets we can see a steady decrease in the percentage from the original to the normalized data sets.

Although logging the variates reduced the differences among individual component coefficients of the first PC their magnitude still remained highly heterogeneous. Moreover, whereas six characters showed insignificant character-component correlations in the raw data, this number was increased to as many as nine in log transformed characters (all of them being the tooth measures). Hence, it is obvious that log-transforming data may

Table 1. Principal component coefficients of the first normalized eigenvector of MGPCA (left columns) and character-component correlation coefficients (right columns) for raw (RAW), log-transformed (LOG) and normalized (NORM) data. Nonsignificant correlations are in parentheses. Below, the percentage of the total variance explained by the first principal component, the variance explained by the first three components (all the eigenvalues being extracted from the covariance matrix), and the proportion of the first eigenvalue computed from the correlation matrix, respectively, are given.

Character	Coefficients/Correlations					
	RAW		LOG		NORM	
A	0.033	0.47	0.774	0.89	0.194	0.50
B	0.038	0.50	0.392	0.66	0.169	0.54
LCb	0.652	1.00	0.136	0.60	0.397	0.90
LB	0.632	1.00	0.150	0.62	0.411	0.90
LaR	0.097	0.59	0.134	0.48	0.312	0.70
LaC	0.128	0.57	0.062	0.39	0.313	0.70
LaZ	0.296	0.80	0.130	0.56	0.327	0.80
hC	0.084	0.38	0.056	0.25	0.224	0.52
LD	0.229	0.88	0.161	0.50	0.341	0.78
LM1i	0.004	(0.10)	0.013	(0.07)	0.119	0.32
LaM1i	0.007	0.23	0.025	(0.11)	0.149	0.42
LM2i	0.005	(0.13)	0.016	(0.05)	0.137	0.35
LM3i	0.002	(0.06)	0.013	(0.03)	0.082	0.25
LaM3i	0.002	(0.06)	0.013	(0.04)	0.128	0.30
LM13i	0.013	0.16	0.014	(0.07)	0.132	0.43
LM1s	0.000	(0.01)	0.012	(0.04)	0.084	0.22
LM2s	0.004	(0.09)	0.027	(0.08)	0.117	0.27
LM13s	0.017	0.18	0.021	(0.11)	0.151	0.43
1st V eigenvalue	81.09%		42.24%		33.80%	
3 V eigenvalues	91.45%		68.02%		61.21%	
1st C eigenvalue	31.32%		32.73%		31.61%	

not adequately standardize the variance of the characters. In addition, it is apparent from table 1 that this transformation substantially changed the relative contribution of individual characters to the total variation explained by the component (cf. the coefficients and correlations in the RAW and LOG columns of the Tab. 1).

On the contrary, MGPCA computed on the normalized data revealed all the correlations to be signifcant although the magnitude of the coefficients was neither the same nor 'similar'. These results suggest that although the differences in the magnitude of the coefficients of the first principal component were partly due to differences in character variances, there was still some portion of persistent variance which could not be associated with the 'size' component: this especially concerned the dental measures. Therefore, the size-adjustment must be used with caution as these results indicate that the first principal component is also very likely to contain some 'shape' information which would be lost on

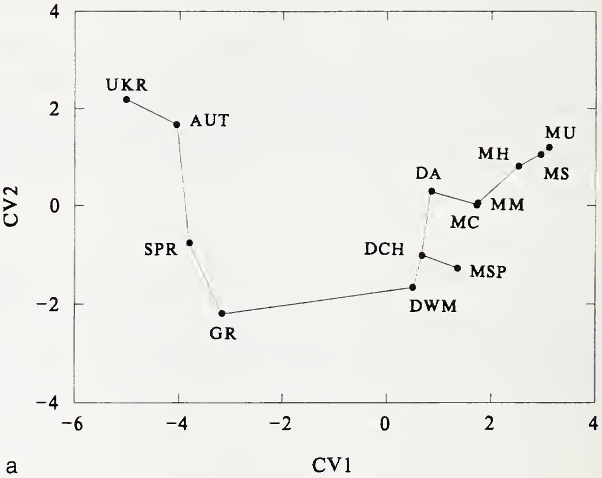

Fig. 1 a

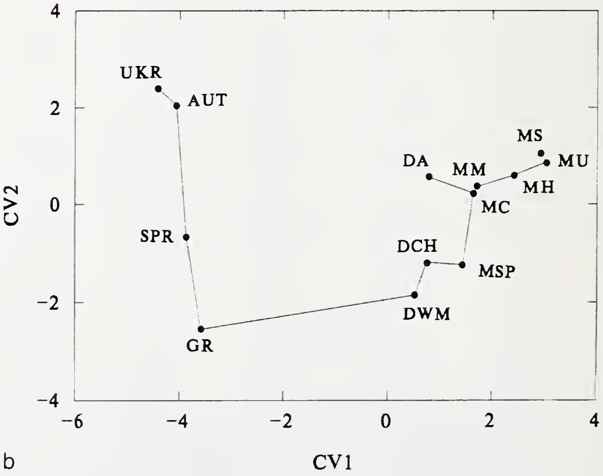

Fig. 1 b

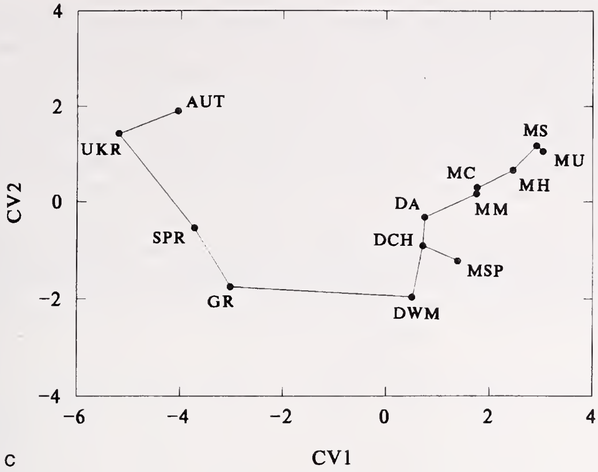

Fig. 1 c

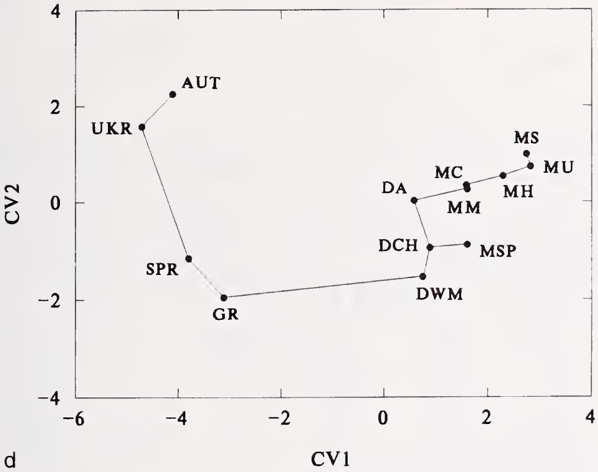

Fig. 1 d

Fig. 1. Plots of the first two canonical variate scores for group centroids; (a), (b)–'size-in' analysis; (c), (d) – 'size-out' analysis; plots (a) and (c) show results of the log-transformed data, whereas (b) and (d) concern the normalized data. Minimum spanning trees were superimposed on the plots.

extraction of the 'size' vector; on the other hand, we may expect a small part of residual size information to be contained in following (intentionally 'size-free') components. Thus results of both the 'size-in' and 'size-out' analyses should be taken into account (accordingly, these terms are rather loose and henceforth will only be used for convenience).

Figure 1 shows plots of scores for the first two canonical variates. For log-transformed and normalized data (Fig. 1 a, c vs. b, d), similar results were revealed. The first variate (CV1) apparently separated the two major lineages, i. e. aboriginal and commensal species groups, while the second one (CV2) identified individual species (or groupings of populations) within the lineages. According to relative values of discriminant coefficients, the contrast of A and B (the so-called zygomatic index), and, to a lesser degree, also LM13i and LaZ were the variables contributing the most to the first discriminant function. In constrast, in the second canonical variate, the relative contribution was not so clear, with the highest coeffcients being those for LB (contrasted by LCb), LM1s, LD and B.

As displayed by the minimum spanning tree, the two *M. spicilegus* populations formed the most remote group within the aboriginal lineage in the two-dimensional discriminant space, whilst *M. spretus* was closer to *M. macedonicus*. When 'size-in' and 'size-out' analyses were compared, the patterns were similar except for the changed relative position of the Ukrainian and Austrian *spicilegus* samples, and the Albanian *M. domesticus* population which tended to be closer to *M. musculus* populations in both the 'size-in' analyses contrary to the 'size-out' ones.

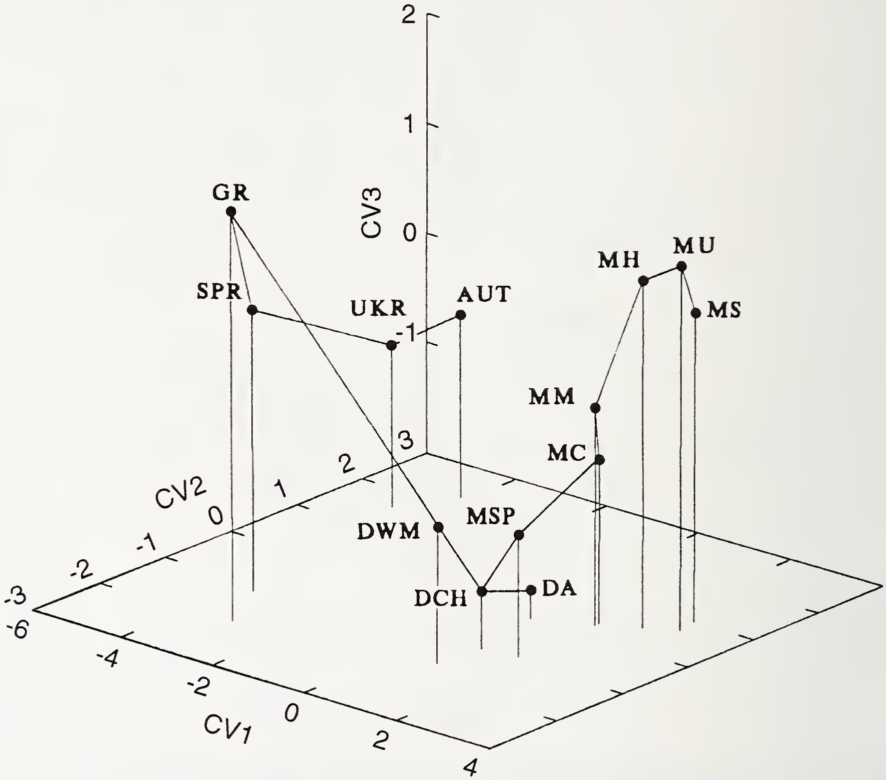

Fig. 2. A three-dimensional plot of the first three canonical variate scores for the 'size-out' CVA on the normalized data. Group centroids are connected by the minimum spanning tree.

A three-dimensional plot of the first three canonical variates is presented in figure 2 for the 'size-out' CVA on normalized data. The third canonical axis, based mostly on the relative rostral width (LaR as compared to LCb), placed the Albanian mice into the *domesticus* cluster and separated Bohemian and Moravian *M. musculus* populations from the Slovakian, Hungarian, and Ukrainian ones. Within the aboriginal group the species were not distantly separated by CV3 with, again, *M. spretus* being between the two eastern short-tailed indoor species. There were no substantial differences between the 'size-in' and 'size-out' and between logged and normalized data analyses on the third canonical axis.

The correctness of canonical discriminations were assessed by a posteriori classification tests. In table 2, actual memberships (rows) were tabulated against predicted ones (columns) from ‘size-out’ CVA on log-transformed data, where populations were pooled within species except the hybrid (MSP) and DA samples. In this analysis, 84.12% cases were classified correctly. When the MSP sample was excluded, the correctness increased to 91.51–92.25% (Cohen’s Kappa ranging between 0.884 and 0.895) depending on the type of the analysis used (see Tab. 3).

Table 2. A posteriori classification testing the correctness of the assignment of each individual to a particular group for the ‘size-in’ CVA performed on log-transformed data. Here, actual group membership (rows) is tabulated against predicted (columns). The populations are pooled within species except for DA and MSP samples.

	MAC	SPI	SPR	DA	DOM	MUS	MSP	*T*
MAC	24	0	0	0	1	0	0	25
SPI	0	42	0	0	1	0	0	43
SPR	1	1	22	0	0	0	0	24
DA	0	0	0	10	0	0	0	10
DOM	0	0	0	4	37	1	2	44
MUS	0	0	0	5	7	94	18	124
MSP	0	0	0	0	4	3	19	26
T	25	43	22	19	50	98	39	296

Table 3. A comparison of the correctness of canonical discriminations on different data sets; LOGIN, ‘size-in’ CVA on logged data; LOGOUT, ‘size-out’ CVA on logged data; NORMIN, ‘size-in’ CVA on normalized data; NORMOUT, ‘size-out’ CVA on normalized data. In columns, percentages of erroneous assignment for each species (plus DA sample, MSP excluded), total classification error (in %), and Cohen’s Kappa are given, respectively. Cohen’s Kappa is an association measure testing if counts along the diagonal in Table 2 are significantly greater than those expected by chance alone; values greater than 0.75 are usually said to indicate strong agreement (Wilkinson 1990).

	MAC	SPI	SPR	DA	DOM	MUS	Total	Kappa
LOGIN	4.00	2.33	8.33	0.00	11.36	9.68	7.78	0.894
LOGOUT	4.00	0.00	8.33	0.00	11.36	10.48	7.78	0.894
NORMIN	4.00	0.00	8.33	0.00	9.09	11.20	7.75	0.895
NORMOUT	8.00	0.00	8.33	0.00	13.64	10.40	8.49	0.884

The ‘size-in’ and ‘size-out’ analyses were compared so that matrices were plotted against each other using the Mantel procedure (NTSYS-pc) separately for each data set (Fig. 3). In both cases, CVA revealed similar results when distances were close to the diagonal (product-moment correlations $r = 0.998$ and $r = 0.991$, respectively). However, there were some differences between size-adjusted and non-adjusted data, especially in the normalized variables, mainly due to *M. macedonicus* which tended to show higher distances in the ‘size-in’ CVA compared to the ‘size-out’ analyses.

Results of a UPGMA cluster analysis based on the Mahalanobis distances are shown in figure 4. Because of the hybrid nature of the MSP population this sample was excluded from the clustering. Interestingly, the ‘size-in’ and ‘size-out’ procedures gave the same trees, while there was a difference between log-transformed and normalized data sets: the Swiss *M. domesticus* (DCH) appeared in the *M. musculus* MC + MM cluster with the logged data (Fig. 4 a), whereas both the species made separate clusters

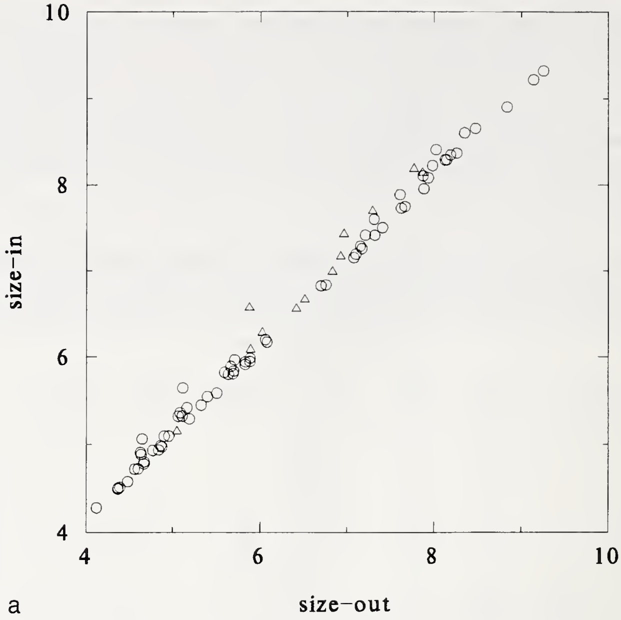

Fig. 3 a

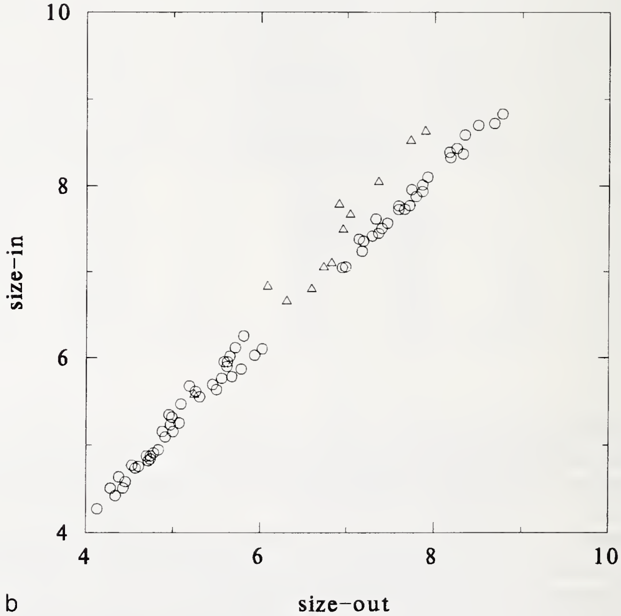

Fig. 3 b

Fig. 3. Mantel plots of 'size-in' (ordinate) and 'size-out' (abscissa) Mahalanobis distances betweeen group centroids; (a) log-transformed data, r = 0.998; (b) normalized data, r = 0.991. In both the cases, *M. macedonicus* is marked by triangles, while all other populations are indicated by circles.

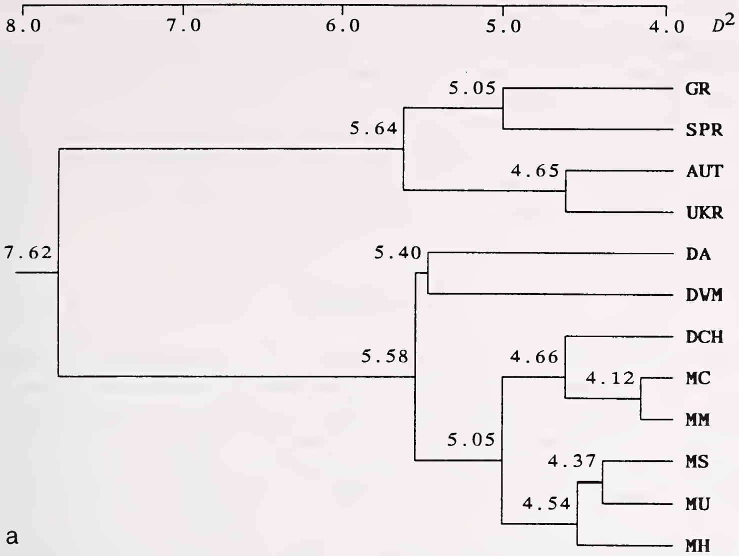

Fig. 4 a

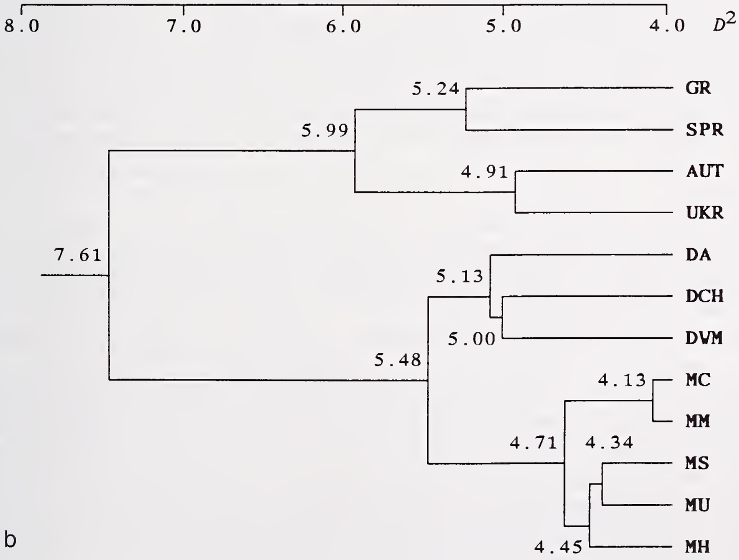

Fig. 4 b

Fig. 4. UPGMA dendrograms based on the Mahalanobis distances; (a) log-transformed data; (b) normalized data. Both the 'size-in' and 'size-out' analyses gave the same results for each type of the data transformation; MSP hybrid sample was excluded from the clustering.

({[MC + MM) + [[MS + MU] + MH]} and {DA + [DCH + DWM]}, respectively) with the normalized data (Fig. 4 b). Albanian mice, although being quite distantly related, clustered with *M. domesticus* in both cases. What should be noted in the dendrograms is the small distance between *M. macedonicus* and *M. spretus* in comparison to the distance between both the *M. spicilegus* populations and among populations within the *domesticus* and *musculus* clusters. The mound-building mouse (*M. spicilegus*) thus appeared to be morphologically the most distant species within aboriginal mice, yet interspecific morphological distances within this group were, in general, strikingly small.

Discussion

Although size can provide significant information on morphometric differences among taxa, it is sometimes desirable to avoid size variation as it may cause a substantial bias in an assessment of group interrelationships due to the growth allometry (Röhrs 1961; Thorpe 1976, 1983 a). This especially concerns organisms with indeterminate growth; however, nutritional, seasonal, sexual, ecological and other factors are also likely to affect morphological characters (Leamy 1981) and thus the size-adjusting of data may be necessary.

Several techniques have been employed to remove the size component from this analysis. In morphometrics, the most familiar methods are those using ratios; however, it has been argued repeatedly that because of some undesirable statistical properties and conceptual difficulties ratios should be avoided (Atchley et al. 1976; Corruccini 1977; Atchley 1978; Atchley and Anderson 1978). Nor does taking logarithms of the ratios (Blackith and Reyment 1971; Dodson 1978; Hills 1978) entirely remove size from data as stated by Humphries et al. (1981) and evidenced by Reist (1985). Another possibility is a univariate regression analysis of variables on a standard size measurement such as snout-vent length in reptiles and amphibians, standard length in fish, wing length in birds, or condylobasal length in mammal skulls (Thorpe 1975; Corruccini 1977; Kuhry and Marcus 1977). However, since size is designated a single variable in these techniques, only one particular variable is partialed out. As pointed out by Humphries et al. (1981) and Thorpe and Leamy (1983) size is not equal to any single measurement and using a multivariate ordination method for comparing size and shape differences among groups is more appropriate.

Multiple group principal component analysis (MGPCA) is now widely used in various types of studies (Thorpe 1983 b; Corti and Thorpe 1989; Allegrucci et al. 1992; Bekele et al. 1993) as a method of evaluating the relative contribution of size and shape to the between-group variation and to extract the size component from data. Nevertheless, some criticisms have appeared concerning the biological meaning of the general size aspect of the first principal component (Shea 1985), and/or pointing out that the first component may contain shape information and remaining vectors may retain size information (reviewed in Humphries et al. 1981; Reist 1985). While the first criticism is irrelevant to this study, the second may pose a problem, because even though the coefficients related to the first principal component appeared all to be of the same sign, their magnitude was very different even after transformation. This may indicate some residual size component could be resident in the second and following axes, whereas a proportion of shape information is likely to be lost with an extraction of the first component.

Comparison of two transformation techniques employed for standardization of the measurements with different scales (log-transformation and normalization) proved to be of interest. While both approaches resulted in a slight improvement in between-variable linearity, logging the data (probably the most widely used method in morphological studies) tended to change the relative contribution of variables to the variation explained by principal components. Moreover, the influence of non-equality of the variances of the original characters, although being lowered relative to the raw data, was not entirely restrained, and in turn the number of variables with insignificant character-component correlations was even increased by 50% in the log-transformed vs. original data. Normalization, on the other hand, resulted in better results both in the standardization of the variances and in the relative contribution of individual variables.

The results of CVA computed on the MGPCA scores revealed similar results for both the normalized and log-transformed variates and only slight differences between the 'size-in' and 'size-out' analyses. In all the cases, the two separate evolutionary and ecological lineages were clearly discriminated. Within the aboriginal lineage, the most distant species were *M. macedonicus* and *M. spicilegus* with *M. spretus* being morphologically intermediate between them. This result is rather surprising given the close genetic relationships between the former two species (Bonhomme et al. 1983, 1984). Likewise, Gerasimov et al. (1990), using the stepwise discriminant analysis, found these forms to be morphologically very similar. On the contrary, in a previous study focused mostly on uni- and bivariate analyses (Macholán 1996) *M. spicilegus* appeared to be closer to *M. spretus* when original untransformed data were analysed, whereas the latter showed greater similarity with *M. macedonicus* when the variables were size-adjusted using Thorpe's (1975) allometric formula. A comparison of the univariate (Macholán 1996) and multivariate (this study) study carried out on the same material shows the former to be more affected by the growth/size influence than the latter. It is not clear, however, to what extent the close similarity between *M. macedonicus* and *M. spretus* reflects the circum-Mediterranean ecological vicariance of the two species (Auffray et al. 1990 b).

Within the commensal lineage, a somewhat peculiar position was displayed by the *M. domesticus* sample from Albania (considered as *M. d. brevirostris*, Reichstein 1978; Marshall 1981, but see the discussion about validity of subspecific categories in *M. domesticus* in Ferris et al. 1983; Sage et al. 1986; Macholán 1996) which was rather distinct both from other *M. domesticus* populations and from *M. musculus*, mainly due to its relatively narrow rostrum (Macholán 1996). However, since 10 Albanian individuals were only studied, more animals should be investigated and perhaps other measurements should be included before the systematic relationships of Albanian and other commensal house mice can be established.

In his multivariate morphometric analyses of house mice from eastern Europe and central Asia, Lavrenchenko (1990, 1994) found the variation within *M. musculus* to be categorical rather than clinal and this led him to distinguish three subspecies: *musculus* from the European part of the former USSR, southern Siberia and eastwards to the Far East; *wagneri* from lowlands north of the Caspian Sea, Kazakhstan, and ex-Soviet Central Asia; and *raddei* from eastern Kazakhstan, Altai, most of Mongolia, and eastern Transbaikalia.

In this study, the investigation of *M. musculus* populations from western parts of its range showed quite different patterns and the variation changed rather continually in the east-west direction. This conclusion is corroborated by the results of the univariate analysis (Macholán 1996) where some measurements were shown to be similar to *M. domesticus* in western localities, especially when the raw variates were taken. This suggests a possibility of introgressing *domesticus* alleles into the *musculus* range across the hybrid zone in western Bohemia and south-eastern Germany (Sage et al. 1986; Tucker et al. 1992; Macholán and Zima 1994). The introgression of polygenic traits is similar to that of biochemical markers (Macholán and Zima 1994 and unpubl. results) but the gene-flow distance might be much higher as indicated by the results of the present study.

Acknowledgements

I wish to express my gratitude to the following persons for kindly providing material (listed in alphabetical order): Dr. M. Anděra (National Museum, Prague), Dr. J.-C. Auffray (University of Montpellier), Dr. K. Bauer (Museum of Natural History, Vienna), Mrs. S. Beránková (Institute of Landscape Ecology, Brno), Prof. J. Hausser (University of Lausanne), Dr. S. V. Mezhzherin (Institute of Zoology, Kiev), Dr. F. Spitzenberger (Museum of Natural History, Vienna), Dr. V. Vohralík (Charles University, Prague), Dr. I. V. Zagorodnyuk (Institute of Zoology, Kiev).

I am indebted to Dr. Heidi Hauffe for helpful comments and a language revision of the manuscript. This work was partly supported by the Ministry of Environment of the Czech Republic (GA/846/93), and by the Grant Agency of the Czech Republic (204/93/0531 and 206/95/1 596).

Zusammenfassung

Morphometrische multivariate Analyse europäischer Arten der Gattung Mus (Mammalia, Muridae)

Die systematischen Beziehungen zwischen 12 Populationsgruppen von fünf europäischen Hausmausarten wurden mit Hilfe verschiedener Methoden zur morphometrischen Multivarianzanalyse untersucht. Die ‚Multiple group principal component analysis' (MGPCA) wurde dazu genutzt, den Beitrag der Größe zur Gesamtvariation zu beurteilen. Es wurde gezeigt, daß ein Teil der ‚Form'-Information in der ersten Hauptkomponente enthalten sein könnte, und weitere Komponenten ähnlich dazu eine residuale ‚Größen'-Information beinhalten könnten. Deshalb sollte das Eliminieren der ‚Größen'-Information mit Vorsicht und erst nach einer angemessenen Überprüfung der Daten vorgenommen werden. Die kanonische Diskriminanzanalyse brachte ähnliche Resultate wie die ‚size-in' und ‚size-out' MGPCA-Untersuchungen. Die erste kanonische Achse grenzte die Freiland- und die kommensalen Artgruppen der Mäuse voneinander ab, währenddessen die zweite Achse Artgruppen identifizierte. Die dritte kanonische Achse teilte Populationsgruppen innerhalb der kommensalen Arten ab. Sowohl die CVA als auch die Clusteranalyse zeigten, daß (1) *M. macedonicus* und *M. spretus* morphologisch miteinander mehr Ähnlichkeit haben als eine der beiden Arten mit *M. spicilegus*; daß (2) die Distanz zwischen *M. musculus* und *M. domesticus* ähnlich groß ist wie die Distanz zwischen den Freilandarten; und daß (3) die Distanz zwischen den Populationen verglichen mit den zwischenartlichen Beziehungen relativ hoch ist.

Reference

Allegrucci, G.; Baldari, F.; Cesaroni, D.; Thorpe, R. S.; Sbordoni, V. (1992): Morphometric analysis of interspecific and microgeographic variation of crayfish from a Mexican cave. Biol. J. Linn. Soc. **47**, 455–468.

Atchley, W. R. (1978): Ratios, regression intercepts, and the scaling of data. Syst. Zool. **27**, 78–83.

Atchley, W. R.; Anderson, D. (1978): Ratios and the statistical analysis of biological data. Syst. Zool. **27**, 71–78.

Atchley, W. R.; Gaskins, C. T.; Anderson, D. (1976): Statistical properties of ratios. I. Empirical results. Syst. Zool. **25**, 137–148.

Auffray, J.-C.; Marshall, J. T.; Thaler, L.; Bonhomme, F. (1990 a): Focus on the nomenclature of European species of *Mus*. Mouse Genome **88**, 7–8.

Auffray, J.-C.; Tchernov, E.; Bonhomme, F.; Heth, G.; Simson, S.; Nevo, E. (1990 b): Presence and ecological distribution of *Mus* "*spretoides*" and *Mus musculus domesticus* in Israel. Circum-Mediterranean vicariance in the genus *Mus*. Z. Säugetierkunde **55**, 1–10.

Bekele, A.; Capanna, E.; Corti, M.; Marcus, L. F.; Schlitter, D. A. (1993): Systematics and geographic variation of Ethiopian *Arvicanthis* (Rodentia, Muridae), J. Zool. (London) **230**, 117–134.

Blackith, R. E.; Reyment, R. A. (1971): Multivariate morphometrics. London, New York: Acad. Press.

Bonhomme, F.; Catalan, J.; Britton-Davidian, J.; Chapman, V. M.; Moriwaki, K.; Nevo, E.; Thaler, L. (1984): Biochemical diversity and evolution in the genus *Mus*. Biochem. Genet. **22**, 275–303.

Bonhomme, F.; Catalan, J.; Gerasimov, S.; Orsini, P.; Thaler, L. (1983): Le complexe d'espèces du genre *Mus* en Europe Centrale et Orientale. Z. Säugetierkunde **48**, 78–85.

Bonhomme, F.; Guénet, J.-L. (1989): The wild house mouse and its relatives. In: Genetical Variants and Strains of the Laboratory Mouse. Ed. by M. F. Lyon and A. G. Searle. Oxford: Oxford Univ. Press. Pp. 649–662.

Boursot, P; Auffray, J.-C.; Britton-Davidian, J.; Bonhomme, F. (1993): The evolution of house mice. Ann. Rev. Ecol. Syst. **24**, 119–152.

Burnaby, T. P. (1966): Growth-invariant discriminant functions and generalized distance. Biometrics **22**, 96–110.

Campbell, N. A.; Atchley, W. R. (1981): The geometry of canonical variate analysis. Syst. Zool. **30**, 268–280.

Corbet, G. B. (1978): The mammals of the Palaearctic region: a taxonomic review. London, Ithaca: Cornell Univ. Press.

Corruccini, R. S. (1977): Correlation properties of morphometric ratios. Syst. Zool. **26**, 211–214.

Corti, M.; Thorpe, R. S. (1989): Morphological clines across a karyotypic zone of house mice in Central Italy. J. evol. Biol. **2**, 253–264.

Darviche, D.; Orsini, P. (1982): Critères de différentiation morphologique et biométrique de deux espèces de souris sympatriques: *Mus spretus* et *Mus musculus domesticus*. Mammalia **46**, 205–217.

Davis, S. J. M. (1983): Morphometric variation of populations of house mice *Mus domesticus* in Britain and Faroe. J. Zool. (London) **199**, 521–534.

Dodson, P. (1978): On the use of ratios in growth studies. Syst. Zool. **27**, 62–67.

Ellerman, J. R.; Morrison-Scott, T. C. S. (1951): Checklist of Palaearctic and Indian mammals, 1758 to 1946. London: Trustees of the British Museum.

Engels, H. (1980): Zur Biometrie und Taxonomie von Hausmäusen (Genus *Mus* L.) aus dem Mittelmeergebiet. Z. Säugetierkunde **45**, 366–375.

Engels, H. (1983): Zur Phylogenie und Ausbreitungsgeschichte mediterraner Hausmäuse (Genus *Mus* L.) mit Hilfe von „Compatibility Analysis". Z. Säugetierkunde **48**, 9–19.

Ferris, S. D.; Sage, R. D.; Prager, E. M.; Ritte, U.; Wilson, A. C. (1983): Mitochondrial DNA evolution in mice. Genetics **105**, 681–721.

Fisher, R. A. (1936): The use of multiple measurements in taxonomic problems. Ann. Eugen. **7**, 179–188.

Gerasimov, S.; Nikolov, H.; Mihailova, V.; Auffray, J.-C., Bonhomme, F. (1990): Morphometric stepwise discriminant analysis of the five genetically determined European taxa of the genus *Mus*. Biol. J. Linn. Soc. **41**, 47–64.

Hills, M. (1978): On ratios – A response to Atchley, Gaskins, and Anderson. Syst. Zool. **27**, 61–62.

Humphries, J. M.; Bookstein, F. L.; Chernoff, B.; Smith, G. R.; Elder, R. L.; Poss, S. G. (1981): Multivariate discrimination by shape in relation to size. Syst. Zool. **30**, 291–308.

Keller, A. (1974): Détermination de l'âge de *Mus musculus* Linné par l'usure de la dentition. Rev. suisse Zool. **81**, 839–844.

Kraft, R. (1985): Merkmale und Verbreitung der Hausmäuse *Mus musculus musculus* L., 1758, und *Mus musculus domesticus* Rutty, 1772 (Rodentia, Muridae) in Bayern. Säugetierkdl. Mitt. **32**, 1–12.

Kratochvíl, J. (1986 a): Die intraspezifische Evolution der Art *Mus domesticus*. Acta Sc. Nat. Brno **20**, 1–49.

Kratochvíl, J. (1986 b): *Mus abbotti* – Eine Kleinasiatisch-Balkanische Art (Muridae – Mammalia). Folia Zool. **35**, 3–20.

Kuhry, B.; Marcus, L. F. (1977): Bivariate linear models in biometry. Syst. Zool. **26**, 201–209.

Laurie, E. M. O. (1946): The reproduction of the house mouse (*Mus musculus*) living in different environments. Proc. R. Soc. Lond. B **133**, 248–281.

Lavrenchenko, L. A. (1990): Systematic analysis of the superspecies complex *Mus musculus* s. lat. PhD Thesis, IEMEA AS USSR, Moscow. [In Russian]

Lavrenchenko, L. A. (1994): Analysis of craniometric characters of house mice *Mus musculus* sensu lato (Rodentia, Muridae): A multivariate approach. Zool. zh. **73**, 169–178. [In Russian with Engl. summary]

Leamy, L. (1981): Heritability of morphometric ratios in random bred house mice. In: Mammalian Population Genetics. Ed. by M. H. Smith and J. Joule. Athens, Georgia: Univ. Georgia Press. Pp. 254–271.

Lyalyukhina, S.; Kotenkova, E.; Walkova, W.; Adamczyk, K. (1991): Comparison of craniological parameters in *Mus musculus musculus* Linnaeus, 1758 and *Mus hortulanus* Nordmann, 1840. Acta theriol. **36**, 95–107.

Macholán, M. (1996): Morphometric analysis of European house mice. Acta theriol. **41** (in press).

Macholán, M.; Zima, J. (1994): *Mus domesticus* in western Bohemia: a new mammal for the Czech Republic. Folia Zool. **43**, 39–41.

Marshall, J. T. (1981): Taxonomy. In: The Mouse in Biomedical Research, Vol. 1. Ed. by L. Foster, J. D. Small, and J. G. Fox. New York: Academic Press. Pp. 17–26.

Marshall, J. T.; Sage, R. D. (1981): Taxonomy of the house mouse. Symp. zool. Soc. Lond. **47**, 15–25.

Orsini, P.; Bonhomme, F.; Britton-Davidian, J.; Croset, H.; Gerasimov, S.; Thaler, L. (1983): Le complexe d'espèces du genre *Mus* en Europe Centrale et Orientale. II. Critères d'identification, répartition et caractéristiques écologiques. Z. Säugetierkunde **48**, 86–95.

Palomo, L. J.; España, M.; López-Fuster, M.-J.; Gosàlbez, J.; Sans-Coma, V. (1983): Sobre la variabilidad fenética y morfométrica de *Mus spretus* Lataste, 1883 en la Península Ibérica. Misc. Zool. **7**, 171–192.

Pelikán, J. (1981): Patterns of reproduction in the house mouse. Symp. zool. Soc. Lond. **47**, 205–229.

Reichstein, H. (1978): *Mus musculus* Linnaeus, 1758. – Hausmaus. In: Handbuch der Säugetiere Europas. Bd. 1. Ed. by J. Niethammer and F. Krapp. Wiesbaden: Akad. Verlagsges. Pp. 421–451.

Reist, J. D. (1985): An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. Can. J. Zool. **63**, 1429–1439.

Rohlf, F. J. (1990): NTSYS-pc. Numerical Texonomy and Multivariate Analysis System, Version 1.60. Setauket, New York: Exeter Publ.

Rohlf, F. J.; Bookstein, F. L. (1988): A comment on shearing as a method for "size correction". Syst. Zool. **36**, 356–367.

Röhrs, M. (1961): Allometrie und Systematik. Z. Säugetierkunde **26**, 130–137.

Sage, R. D. (1981): Wild mice. In: The Mouse in Biomedical Research, Vol. 1. Ed. by L. Foster, J. D. Small, and J. G. Fox. New York: Acad. Press. Pp. 39–90.

Sage, R. D.; Whitney, J. B. III.; Wilson, A. C. (1986): Genetic analysis of a hybrid zone between *domesticus* and *musculus* mice (*Mus musculus* complex): Hemoglobin polymorphisms. Curr. Top. Microbiol. Immunol. **127**, 75–85.

Sage, R. D.; Atchley, W. R.; Capanna, E. (1993): House mice as models in systematic biology. Syst. Biol. **42**, 523–561.

Sans-Coma, V.; López-Fuster, M.-J.; Gosàlbez, J. (1979): Über die Hausmaus, *Mus musculus* Linné, 1758, auf der Insel Meda Crossa, Katalonien, Spanien. Säugetierkdl. Mitt. **27**, 107–113.

Schwarz, E.; Schwarz, H. K. (1943): The wild and commensal stocks of the house mouse, *Mus musculus* Linnaeus. J. Mammalogy **25**, 59–72.

Scriven, P. N.; Bauchau, V. (1992): The effect of hybridization on mandible morphology in an island population of the house mouse. J. Zool. (London) **226**, 573–583.

Serafiński, W. (1965): The subspecific differentiation of the central European house mouse (*Mus musculus* L.) in the light of their ecology and morphology. Ecol. Polska, Ser. A **13**, 305–348.

Shea, B. T. (1985): Bivariate and multivariate growth allometry: statistical and biological considerations. J. Zool. (London) **206**, 367–390.

Sokal, R. R.; Rohlf, F. J. (1981): Biometry. 2nd ed. New York: Freeman and Co.

Thaler, L.; Bonhomme, F.; Britton-Davidian, J. (1981): Processes of speciation and semi-speciation in the house mouse. Symp. zool. Soc. Lond. **47**, 27–41.

Thorpe, R. S. (1975): Quantitative handling of characters useful in snake systematics with a particular reference to intraspecific variation in the Ringed snake *Natrix natrix* (L.). Biol. J. Linn. Soc. **7**, 27–43.

Thorpe, R. S. (1976): Biometric analysis of geographic variation and racial affinities. Biol. Rev. **51**, 407–452.

Thorpe, R. S. (1983 a): A biometric study of the effects of growth on the analysis of geographic variation: Tooth number in green geckos (Reptilia: *Phelsuma*). J. Zool. (London) **201**, 13–26.

Thorpe, R. S. (1983 b): A review of the numerical methods for recognising and analysing racial differentiation. In: Numerical Taxonomy. Ed. by J. Felsenstein. Berlin, Heidelberg, New York, Tokyo: Springer-Verlag. Pp. 404–423.

Thorpe, R. S.; Corti, M.; Capanna, E. (1982): Morphometric divergence of Robertsonian populations/species of *Mus*: a multivariate analysis of size and shape. Experientia **38**, 920–923.

THORPE, R. S.; LEAMY, L. (1983): Morphometric studies in inbred and hybrid House mice (*Mus* sp.): Multivariate analysis of size and shape. J. Zool. (London) **199**, 421–432.
TUCKER, P. K.; SAGE, R. D.; WARNER, J. H.; WILSON, A. C.; EICHER, E. M. (1992): Abrupt cline for sex chromosomes in a hybrid zone between two species of mice. Evolution **46**, 1146–1163.
WILKINSON, L. (1990): SYSTAT: The System for Statistics, Release 5.02. Evanson, IL: SYSTAT, Inc.

Author's address: Dr. MILOŠ MACHOLÁN, Laboratory of Genetics and Embryology, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Veveří 97, CZ-60200 Brno, Czech Republic.

Zeitschrift/Journal: Mammalian Biology (früher Zeitschrift für Säugetierkunde)

Jahr/Year: 1996

Band/Volume: 61

Autor(en)/Author(s): Macholán Milos

Artikel/Article: Multivariate morphometric analysis of European species of the genus Mus (Mammalia, Muridae) 304-319