

A Multivariate Study of Variation in Two Species of Rock Crab of the Genus *Leptograpsus*

N. A. Campbell^A and R. J. Mahon^B

^A Division of Mathematical Statistics, CSIRO, Floreat Park, W.A. 6014.

^B Zoology Department, University of Western Australia, Nedlands, W.A. 6009.

Abstract

A multivariate approach has been used to study morphological variation in the blue and orange-form species of rock crab of the genus *Leptograpsus*. Objective criteria for the identification of the two species are established, based on the following characters: width of the frontal region of the carapace; width of the posterior region of the carapace (rear width); length of the carapace along the midline; maximum width of the carapace; and the depth of the body.

The first canonical variate, which differentiates between the two species, represents a contrast between the carapace width relative to the width of the front lip and the depth of the body; the blue-form species has a greater relative carapace width than has the orange form.

The second canonical variate, which presents a contrast between the rear width and the carapace length, identifies males and females within each species; males have a greater relative carapace length than have females.

All individuals, including 20 not used in the study, were correctly identified for colour form.

The clear separation of the blue and orange forms achieved by canonical analysis supports the previously determined specific status of the two forms.

Introduction

The rock crab *Leptograpsus variegatus* (Fabricius, 1793), is recorded as occurring on a number of southern Pacific islands, the western coast of South America, and the coasts of Australia south of the Tropic of Capricorn (Balss 1935; Griffin 1969).

Mahon (1973), using ecological studies which extended those of Shield (1959), and a genetical analysis based on an electrophoretic study, established the specific distinctness of rock crabs of the blue and orange forms of the genus *Leptograpsus* which occur on the coasts of Australia. These colour forms were previously regarded as morphs of *L. variegatus* (Shield 1959; Griffin 1973).

The conclusion that the two colour forms are indeed two separate species has made it necessary to examine the extant type material of all the taxa previously placed in the synonymy of *L. variegatus*. This is necessary to determine whether any of these taxa are available for the two Western Australian species; the types of *L. variegatus* itself are apparently lost (Rathbun 1918). However, leptograpsid crabs which have been chemically preserved lose their distinguishing colour.

In an attempt to resolve this problem of identification, a morphological study of the Western Australian species was undertaken. This paper reports the use of multivariate analyses to examine morphological data on the two species (here referred to as the colour forms), and to develop objective criteria for identification, independently of colour.

Materials and Methods

In this study, 200 animals (50 males and 50 females of each colour form) were collected at Fremantle, W.A. (32°S., 117°E.). All samples were collected and measured by one of the authors (R.J.M.). Samples were collected at night, animals being captured on sight, as the collector moved along the sample area. Measurements taken were: (1) the width of the frontal region (tip) of the carapace, just anterior to the front tubercles (FL); (2) the width of the posterior region of the carapace (rear width RW); (3) the length of the carapace, along the midline (CL); (4) the maximum width of the carapace (CW); (5) the depth of the body measured after displacement of the abdomen (only necessary for females) (BD). Measurements were made with vernier calipers, to the nearest 0.1 mm.

The four carapace measurements recorded are shown in Fig. 1, and the data for each group summarized in Table 1. A sample of the specimens was deposited with the Western Australian Museum.

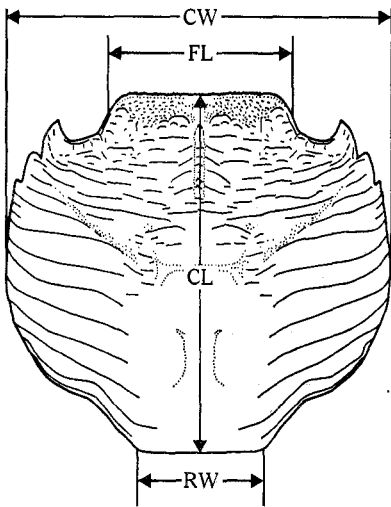


Fig. 1. Dorsal view of carapace of *Leptograpsus*, showing measurements taken. *FL*, width of frontal region just anterior to frontal tubercles. *RW*, width of posterior region. *CL*, length along midline. *CW*, maximum width. The body depth was also measured; in females but not in males the abdomen was first displaced.

Results

The multivariate statistical procedures adopted are discussed in detail by Phillips *et al.* (1973).

Normality of the Data

The use of multivariate techniques assumes a multivariate normal distribution of the variables for tests of significance. Examination of the distribution for each variable, by consideration of Fisher's g_1 and g_2 statistics (Fisher 1958) and frequency histograms showed that none of the distributions differed from the normal, so the untransformed data were used.

Multivariate Analysis of Variance

A multivariate analysis of variance of the between-group variation showed Wilk's lambda (Λ) to equal 0.0234, which was highly significant. Subdivision of the between-groups variation into components for 'colour form', 'sex' and 'colour form \times sex' showed the interaction to be significant, though the variations accounted for by 'colour form' and 'sex' were of a greater order. This is considered in further detail below.

Canonical Analysis

Canonical analysis may be used to determine linear combinations of the original variables that best discriminate between the four groups. These linear combinations

are determined so that the variation between the group means, relative to the variation within groups, is maximized. The first canonical axis effectively maximizes the ratio of the between-groups to within-groups sums of squares for the first canonical variate; this ratio is the canonical root.

A canonical analysis was carried out to determine groupings. Bartlett's multi-variate test of homogeneity of the covariance matrices showed a significant departure from homogeneity (significant results for both the covariance and correlation matrices indicated significant differences in both the size and orientation of the confidence ellipsoids). However, examination of the effect of this lack of homogeneity (see below) on the conclusions showed that the results of the canonical analysis are valid.

Table 1. Summary of data for each species and sex of *Leptograpsus*
Measurements in millimetres. For abbreviations of measurements, see text p. 418

Group	Value	FL	RW	CL	CW	BD
Blue male	Mean	14.8	11.7	32.0	36.8	13.4
	Range	8.1-21.3	6.7-15.8	16.1-47.1	19.0-54.6	7.0-20.0
	Variance	10.256	4.459	53.417	69.781	10.239
	S.E.	0.453	0.299	1.034	1.181	0.453
Blue female	Mean	13.3	12.1	28.1	32.6	11.8
	Range	7.2-19.2	6.5-16.9	14.7-40.9	17.1-47.9	6.1-18.1
	Variance	6.905	5.947	35.042	46.291	7.576
	S.E.	0.372	0.345	0.837	0.962	0.389
Orange male	Mean	16.6	12.3	33.7	37.2	15.3
	Range	9.1-23.1	6.9-15.8	16.7-47.6	18.6-52.8	7.4-21.6
	Variance	12.355	4.821	57.930	70.348	12.441
	S.E.	0.497	0.311	1.076	1.186	0.499
Orange female	Mean	17.6	14.8	34.6	39.0	15.6
	Range	10.7-23.1	9.2-20.2	21.4-46.2	24.0-52.5	9.7-21.1
	Variance	8.845	5.520	34.072	42.801	7.575
	S.E.	0.421	0.332	0.825	0.925	0.389

Five animals were randomly excluded from each group to provide 'independent' data for examination of the value of the resulting canonical variates. Fig. 2 shows the transformed group means resulting from the canonical analysis of the body measurements (given in Table 2), and the canonical variates for the animals included in the analysis, and for the five animals excluded from each group.

The first canonical variate accounted for 69% of the variation between groups. The coefficients of the first canonical variate, expressed in terms of the within-group standard deviations of the corresponding variates (Table 2), show that the first canonical variate results from a contrast between the maximum carapace width and, on the other hand, the width of the front lip and the body depth of the animal; the blue form has a greater carapace width relative to the front lip and to the depth than has the orange form. The second canonical variate (which accounted for 29% of the variation between groups) results from a contrast between the maximum length of the carapace and the rear width (and to a lesser extent the carapace width); males have a greater carapace length relative to the rear width than have females. Both canonical variates result from contrasts between body dimensions, implying that the resulting canonical variates discriminate in terms of the shape of the animal.

Fig. 2 shows that there is complete separation of the blue and orange forms along the first canonical axis. Thus all animals, including the randomly excluded animals, were discriminated perfectly for colour form.

The second canonical variate largely discriminated on the basis of the sex of the crabs; thus variates contributing to this are probably sex-linked. Nineteen of the 20 excluded crabs were correctly identified for both colour form and sex (one blue male was identified as a blue female). There was only minimal overlap on the second canonical axis for both the blue and the orange forms.

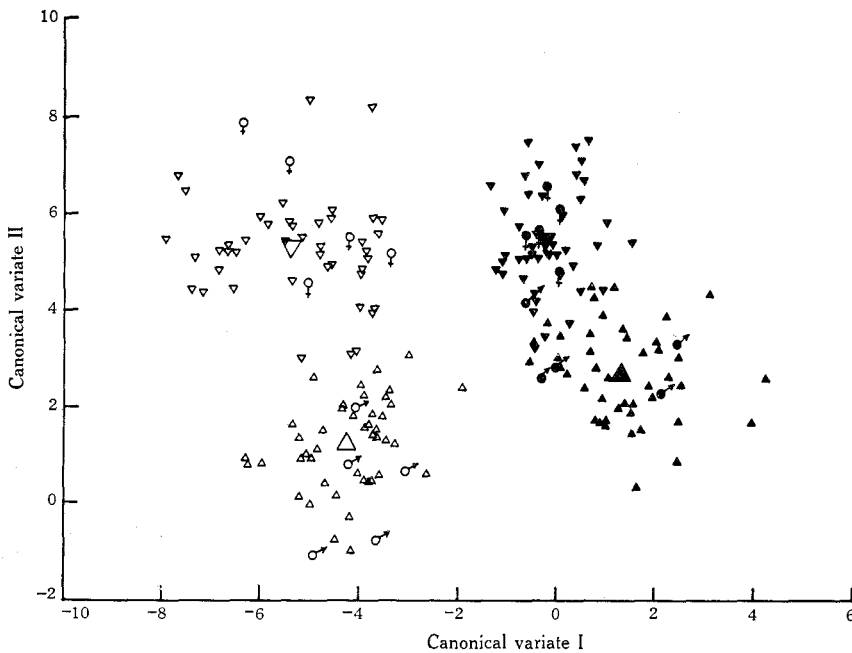


Fig. 2. Discrimination of blue males (Δ), blue females (∇), orange males (\blacktriangle) and orange females (\blacktriangledown), with respect to the first two canonical axes. Large symbols, means of groups of 45. Male and female symbols, five of each group arbitrarily excluded (see text).

All variables contributed to either the colour form or sex discrimination or both.

The colour form and sex interpretations for the first two canonical analyses were confirmed by calculation of the canonical variates for the corresponding matrices from the multivariate analysis of variance. The marked similarity between the two sets of results is shown in Table 3. The only difference is the relative unimportance of carapace width in discriminating for sex. The results also confirm that the major discrimination is in terms of colour form.

Effect of Heterogeneity of the Covariance Matrices

Canonical analysis assumes that the within-group variation is similar for each group. This is not always found in biological studies. To examine the effect of heterogeneity of the covariance matrices on the resulting canonical variates, Mahalanobis' D^2 was calculated for each pair of groups, based on the pooled covariance matrix, on the individual covariance matrices for each pair of groups, and on Anderson's

Table 2. Canonical analysis of the four colour form and sex groups of *Leptograpsus*
 Values in parentheses are canonical variates weighted by within-group standard deviation

Front lip	Rear width	Carapace length	Carapace width	Body depth	
Between-groups matrix; 3 d.f.					
490·549	265·629	714·772	654·758	452·752	
	266·797	321·909	322·510	216·764	
		1124·236	1057·617	666·983	
			1037·743	597·779	
				425·673	
Within-groups matrix; 176 d.f.					
1687·499	1195·421	3626·865	4080·533	1651·064	
	913·993	2608·634	2941·978	1185·273	
		7919·014	8901·151	3598·105	
			10039·066	4046·596	
				1658·169	
Within-groups correlation matrix					
1·0	0·963	0·992	0·991	0·987	
	1·0	0·970	0·971	0·963	
		1·0	0·998	0·993	
			1·0	0·992	
				1·0	
Canonical variates					Canonical root
CVI					
-1·538	-0·684	-0·120	1·478	-1·374	7·713
(-4·763)	(-1·559)	(-0·804)	(11·161)	(-4·218)	
CVII					
0·175	1·536	-1·177	0·726	-0·522	3·200
(0·541)	(3·500)	(-7·892)	(5·486)	(-1·601)	

Table 3. Comparison of canonical analyses based on all four groups, and on the covariance matrices corresponding to the multivariate factorial effects of colour and sex

Values in parentheses are canonical variates weighted by within-groups standard deviation

	Canonical variates					Canonical root
	Front lip	Rear width	Carapace length	Carapace width	Body depth	
CVI ^A	-1·538 (-4·763)	-0·684 (-1·559)	-0·120 (-0·804)	1·478 (11·161)	-1·374 (-4·218)	7·713
Colour	-1·510 (-4·677)	-0·394 (-0·898)	-0·319 (-2·142)	1·582 (11·973)	-1·445 (-4·444)	7·377
CVII	0·175 (0·541)	1·536 (3·500)	-1·177 (-7·892)	0·726 (5·486)	-0·522 (-1·601)	3·200
Sex	0·607 (1·880)	1·674 (3·814)	-0·947 (-6·361)	0·121 (0·916)	-0·057 (-0·176)	3·353

^A Canonical variate based on all four groups.

(1958) solution to the Behrens–Fisher problem for unequal covariance matrices. Gower (1966) has shown that a principal coordinates analysis of the matrix of D^2 values (based on the pooled within-groups covariance matrix) provides the same configuration in the space defined by the first q principal coordinate axes as that provided by the transformed group means in the space defined by the first q canonical axes. If Gower's principal coordinates analyses of the D^2 matrices defined above show similar configurations, it can be concluded that the within-group heterogeneity has not had a marked effect on the relative positions of the groups.

Principal coordinates analyses of the three similarity matrices derived from Mahalanobis' D^2 are given in Figs 3a–3c. The principal coordinates analysis of the similarity matrix derived from D^2 based on the pooled covariance matrix (Fig. 3a) agrees with the results of the canonical analysis (Fig. 2) since the two approaches are equivalent.

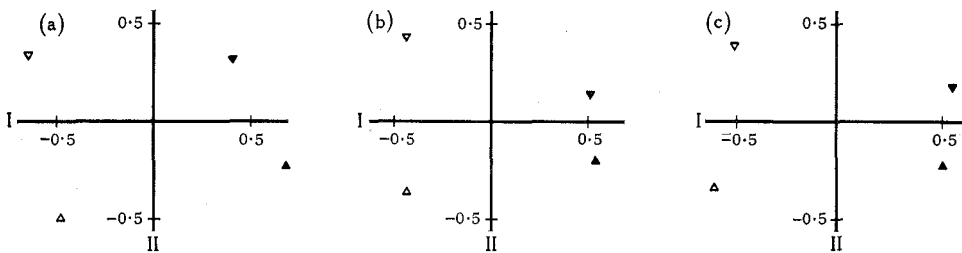


Fig. 3. Co-ordinates of blue males (\triangle), blue females (∇), orange males (\blacktriangle) and orange females (\blacktriangledown) with respect to the first two principal coordinate axes. (a) Derived from the matrix of D^2 values based on the pooled covariance matrix. (b) Derived from the matrix of D^2 values based on the individual covariance matrices. (c) Derived from the matrix of D^2 values based on Anderson's (1958) solution for unequal covariance matrices.

Comparison of Figs 3a–3c shows that lack of homogeneity of the covariance matrices has no effect on the discrimination between the blue and orange forms, nor on the separation of orange males from females. However, separation of blue males from females is less marked. Thus the results of the canonical analysis (Fig. 2 and Table 2), with regard to major discrimination based on colour form, and less marked discrimination due to sex (effectively within each colour form) are valid.

Principal Components Analysis

Principal components analysis may be used to study variation within each of the four groups for growth patterns, by transforming the original observations to independent linear combinations of the original variables. Geometrically, the approach can be considered as the rotation of the axes of the original coordinate system to new orthogonal axes, called principal axes.

The results of the principal components analyses of the size measurements, based on the correlation matrices, showed that the first principal component accounted for more than 98% of the total variation for each group. The corresponding eigen-vectors were positive for each variable, indicating a general 'size' component. The eigen-vectors did not differ from the hypothetical vector (0.447, 0.447, 0.447, 0.447), the cosine of the angle between the observed and the hypothetical eigen-vector being

greater than 0.999 for each group. This first component is the usual 'mean' vector found in many growth studies. The second principal component accounted for less than 1% of the total variation; this implies that the crabs are very uniform in shape, varying only in size, with a very consistent relationship between dimensions within each colour form and sex group.

Discussion

While body size dimensions are the variables used in the canonical analysis, it is apparent from the nature of the canonical variates that absolute size is not the criterion used to discriminate between colour forms (i.e. species); it is, rather, the combination of size measurements to give differences in shape (i.e. relative proportions) which separate the species (and sexes within species). Since the crabs are highly uniform in shape within species (for both males and females), the differences in shape revealed by the canonical variates are characteristic of most individuals. This high uniformity of shape, shown by the results of the principal components analysis, explains in part the high degree of separation achieved.

The results of the canonical analysis show that the first canonical variate represents a contrast between the carapace width on the one hand and the width of the front lip and the depth of the body on the other, while the second canonical variate represents a contrast between the rear width and the carapace length. Plots of width of front lip against carapace width, body depth against carapace width, and carapace length against rear width, are given in Figs 4a-4c. Neither the colour forms nor the sexes would be separated by consideration of each variable separately (i.e. a univariate approach) as shown by the overlap of the animals for each variable.

The scatter diagrams (bivariate approach) in Figs 4a and 4b show very little overlap between the blue and orange forms, though there is no separation of the sexes. However, a few animals would be incorrectly identified. Fig. 4c shows separation between the sexes, though not between species; overlap is again minimal. Thus while the scatter diagrams (bivariate plots) show separation between the blue and orange forms, with only minimal overlap, the results of the canonical analysis show complete separation of the species, with the correct identification for colour form for all animals.

Previous morphometric analyses of the rock crab *L. variegatus* have been limited to a comparison of individual measurements and to ratios of measurements, and to regressions of one variable on another (Griffin 1973). With the overlap evident in Figs 4a and 4b the results could be interpreted to suggest that there is no apparent difference, on the basis of the measured characters, in the two colour forms. Stephenson (1964) also used a regression approach in a morphometric analysis of swimming crabs of the genus *Portunus*.

A regression approach in these situations must be used with care, since both variables are generally subject to random variation. While a regression of one variable on the other is appropriate if the intention is to estimate one variable from the other, it is not appropriate if it is the underlying relationship that is of interest. Jolicoeur and Heusner (1971) and Ricker (1973) have discussed this problem.

Neither of the studies described above was concerned with the identification of individual crabs; morphometric studies of that type are rare. Identification of individuals by multivariate techniques, particularly canonical analysis, has been

used in anthropology (Howells 1969; Blackith and Reyment 1971), though its use is not widespread in other fields. Lubischev (1959) studied situations in the chrysomelid beetles of the genus *Halticus*, though generally the use of multivariate techniques has been restricted to study of geographic variation and possible speciation (see, e.g. Blackith and Reyment 1971), and has not been concerned with the identification of individuals of a species.

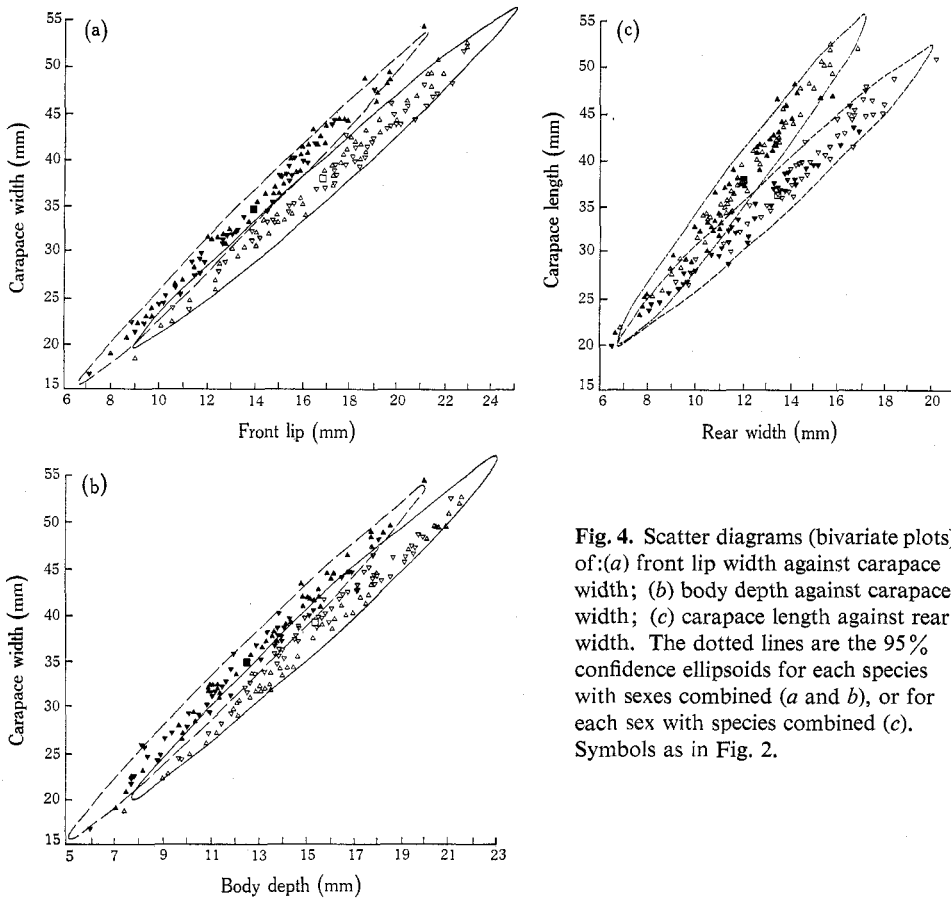


Fig. 4. Scatter diagrams (bivariate plots) of: (a) front lip width against carapace width; (b) body depth against carapace width; (c) carapace length against rear width. The dotted lines are the 95% confidence ellipsoids for each species with sexes combined (a and b), or for each sex with species combined (c). Symbols as in Fig. 2.

Since some species of animals, e.g. sibling species of *A. gambiae* (Davidson *et al.* 1967), are almost indistinguishable in external appearance, identification of individuals by simple measures such as ratios is unlikely to be successful in these species. Multivariate techniques are more appropriate since the variation of the individual measurements is considered simultaneously with the correlations between them. In our study, all of the characters examined gave frequency distributions which overlapped almost completely, while the scatter diagrams for two characters showed some overlap. However, the results of the canonical analysis showed complete separation of the two colour forms, while the resulting canonical variate provided an effective tool for identifying the two species. A logical extension of this study is to use the canonical variate to identify type specimens and to study geographic variation of the two colour forms.

The clear separation of the blue and orange forms achieved by canonical analysis supports the previously determined specific status of the two forms. The existence of two morphologically distinct, sympatric populations with definite ecological and behavioural differences (Shield 1959; Mahon 1973) further suggests that the two forms are indeed two species.

Acknowledgments

The authors thank Dr H. E. Paterson, Zoology Department, University of Western Australia, and Mrs D. M. E. Leadbetter, CSIRO, for many helpful suggestions in the writing up of this work.

References

- Anderson, T. W. (1958). 'An Introduction to Multivariate Statistical Analysis.' (John Wiley and Sons: New York.)
- Balss, H. (1935). Brachyura of the Hamburg Museum expedition to south-western Australia, 1905. *J. Proc. R. Soc. West. Aust.* **21**, 113-51.
- Blackith, R. E., and Reyment, R. A. (1971). 'Multivariate Morphometrics.' (Academic Press: London and New York.)
- Davidson, G., et al. (1967). The *Anopheles gambiae* complex. In 'Genetics of Insect Vectors of Disease'. (Eds J. W. Wright and R. Pal.) (Elsevier: Amsterdam.)
- Fisher, R. A. (1958). 'Statistical Methods for Research Workers.' (13th Ed.) (Oliver and Boyd: London.)
- Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325-38.
- Griffin, D. J. G. (1969). Notes on the taxonomy and zoogeography of the Tasmanian grapsid and ocpodid crabs (Crustacea, Brachyura). *Rec. Aust. Mus. Sydney* **27**, 323-47.
- Griffin, D. J. G. (1973). A revision of the two southern temperate shore crabs *Leptograpsus variegatus* (Fabricius) and *Plagusia chabrus* (Linnaeus) (Crustacea, Decapoda, Grapsidae). *J. R. Soc. N.Z.* **3**, 415-40.
- Howells, W. W. (1969). The use of multivariate techniques in the study of skeletal populations. *Am. J. Phys. Anthropol.* **31**, 311-4.
- Jolicoeur, P., and Heusner, A. A. (1971). The allometry equation and the analysis of the standard oxygen consumption and body weight of the white rat. *Biometrics* **27**, 841-55.
- Lubischev, A. A. (1959). The application of biometrics in taxonomy. *Vestn. Leningr. Univ. Ser. Biol.* **14**, 128-36.
- Mahon, R. J. (1974). A study of rock crabs of the genus *Leptograpsus*. Ph.D. Thesis, University of Western Australia.
- Phillips, B. F., Campbell, N. A., and Wilson, B. R. (1973). A multivariate study of geographic variation in the whelk *Dicathais*. *J. Exp. Mar. Biol. Ecol.* **11**, 29-63.
- Rathbun, M. J. (1918). The grapsoid crabs of America. Bull. U.S. Nat. Mus. No. 97.
- Ricker, W. E. (1973). Linear regressions in fishery research. *J. Fish. Res. Board Can.* **30**, 409-34.
- Shield, J. W. (1959). Population studies in the littoral at Rottneest Island. In 'Rottneest Island: The Rottneest Biological Station and Recent Scientific Research'. (Eds E. P. Hodgkin and K. Sheard.) *J. R. Soc. West. Aust.* **42**, 65-95.
- Stephenson, W. (1964). A morphometric analysis of certain western American swimming crabs of the genus *Portunus* Weber, 1795. In 'Proceedings of Symposium on Crustacea'. Part I.