

Congruence between independent character systems across a hybrid zone: patterns in geographic space

By R. S. THORPE

Received on 1. October 1986

Abstract

The hybrid zone between eastern and western grass snakes, which has been independently hypothesized to be due to secondary contact, is investigated using univariate and multivariate analysis of a large number of independent, quantitative characters. Investigation of the patterns of geographic variation show that 5 out of 6 character systems show a congruent pattern of "stepped clinal" change across the zone. Insofar as this case is concerned, secondary contact 'hybrid' zones are associated with concordant changes in many character systems and "speciation" involves the origin of a large number of small differences in most character systems rather than the major reorganisation of a single system.

Key words: Hybrid zone – geographic variation – character congruence – multivariate analysis – stepped clines – speciation

Introduction

Transition zones, or so called hybrid zones, between parapatric forms may arise in situ due to primary differentiation in response to current selection pressures, i. e. "ecologically" caused. They may also arise due to secondary contact of population that have differentiated in past allopatry. The former situation, if it has arisen recently, may not involve congruent geographic variation, i. e. coincident "stepped clines" (see THORPE 1985a for definitions of a stepped cline), in a wide range of independent character systems. There is a paucity of clear "ecologically" caused congruent geographic variation in a wide range of systems, most studies of natural selection showing variation in single character or system (see Table 5.1 in ENDLER 1986). However, if two lineages have been allopatric for sufficient time, detectable differences between them may accumulate in a wide range of character systems independent of coadaptation. Subsequent secondary contact between the two lineages may result in parapatric forms joined by a hybrid zone across which characters from a wide range of systems show a congruent "stepped clinal" change.

The hybrid zone between the eastern and western mainland European forms of the grass snake *Natrix natrix* is used to test this proposition. Previous phenetic studies have suggested that this zone is due to secondary contact after allopatry induced by Pleistocene events (THORPE 1979). The hybrid zone in *N. natrix* has a similar, but not identical, position and orientation to hybrid zones in other species (THORPE 1979) and these zones are not related to any pertinent environmental changes or physical barriers. Multivariate analysis of the zone (THORPE 1979) has shown that in the centre of transition the populations contain intermediate specimens but not the entire range of purely eastern or western forms. Consequently, these intermediates are not direct F1 hybrids between purely eastern and western parents. However, as ENDLER (1977) has pointed out, it is not possi-

ble to deduce the cause of a zone from its nature. What is needed (THORPE 1984a) is a procedure to hypothesise whether or not a zone is due to secondary contact independent to its nature so that a later investigation of its nature can be related to its cause without circular logic. This has been provided in this case by range expansion studies involving statistical analysis of anagenic trends in infraspecific lineages [fully explained in THORPE (1984a) and not duplicated here] as opposed to considering only the cladogenesis which does not allow 'ecological' and 'phylogenetic' causes to be distinguished (ENDLER 1982, 1983; THORPE 1982, 1983, 1984a). The importance of testing the pattern of anagenesis has not been appreciated by some population geneticists, e. g. BARTON and HEWITT (1985) who have consequently misunderstood the role of these techniques.

In recent reviews of such zones it is apparent that systematic studies of congruence between different character systems are rare and that studies of character congruence within systems may be based on very few characters, e. g. 3 to 6 enzymes. The zone under investigation offers the advantages, for the current study, of being based on an atypically large number of significant characters (71) from an atypically large number (6) of independent character systems. This allows a test of whether each independent character system reflects the hybrid zone by a geographically coincident "stepped clinal" pattern.

Furthermore, since the two forms in question are approaching species status (THORPE 1979) this also allows one to test, in this instance, whether the differences between closely related sister-species is due to small changes across a wide range of character systems, or changes limited to a single character system. This can then contribute to a broader consensus concerning the nature of differences between closely related sister-species and the nature of changes during speciation.

Methods

Geographic region

The broad pattern of change *across* the zone is considered within Europe (Fig. 1) between populations from Britain (1), northern France (3), southern France (6), Holland (7), West Germany (8) and mid-Italy (15) in the west, an intermediate population in the hybrid zone in northern Italy (14) and eastern populations from southern Germany (9), northern Germany (11), Scandinavia (13), Austria (20), Bulgaria (23), Moldavia (24) and Poland

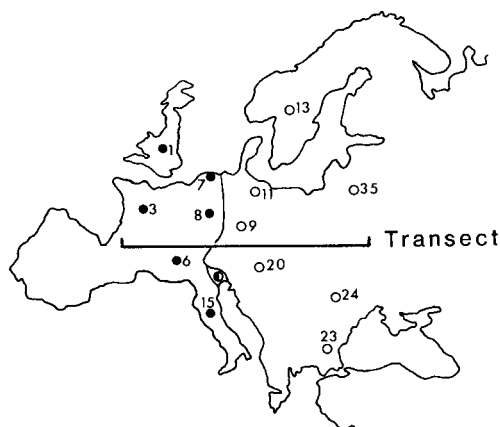


Fig. 1. Populations used to study changes across the hybrid zone. A map of Europe showing western (solid circles) and eastern (hollow circles) populations either side of hybrid zone running approximately north-south. Population 14 is within the transition

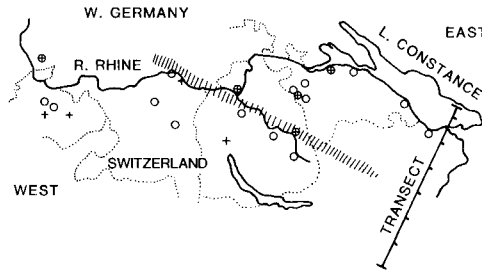


Fig. 2. Populations used to study changes within the hybrid zone. A map of northern Switzerland showing localities of males (crosses) and females (circles) and the centre of the hybrid zone (cross hatching). The transect is marked in 10 km units

(35). These 14 populations (6 west, 1 intermediate, 7 east) are projected against a longitudinal geographic transect illustrated in Fig. 1. Supplementary to this, an attempt is made to investigate the pattern of changes *within* the zone in northern Switzerland (Fig. 2).

Characters and character systems

In the broad area under consideration (plus Jersey, see THORPE 1985a) 71 out of 93 characters showed a statistically significant difference between populations. This large set of characters (measured on male specimens and listed in THORPE 1985a) was selected to study the variation across the hybrid zone. This character set is composed of characters from 6 independent character systems, 3 of which supply the majority of characters, i. e. the colour pattern (27), internal morphology (20) and scalation (14) and 3 of which supply additional characters, i. e. the dentition (4), body proportions regressed against body length (4) and dermal sense organs (2).

These characters have a low-within locality correlation (THORPE 1985b) indicating their independence (THORPE 1987). Moreover, the characters are largely continuous or meristic and so are likely to be polygenically controlled. Consequently, this character set is likely to reflect variation across a very large number of loci.

Assessing patterns and their congruence across the hybrid zone

The pattern of geographic variation in an individual character across the broad area considered in this study is taken as the variation between the 14 population means (Fig. 1). The primary pattern of geographic variation in a multivariate set of characters is taken as being represented by the major principal component/coordinate, i. e. the component associated with the largest, eigen-value in a principal component/coordinate analysis on the normalized means of the 14 populations.

A model stepped cline, representing the eastern and western races and the hybrid zone between them, can be taken as -1 for the 6 western populations, zero for the 'hybrid' populations, and $+1$ for the 7 eastern populations (i. e. west to east, $-1, -1, -1, -1, -1, -1, 0, 1, 1, 1, 1, 1, 1, 1$). This model, projected against the transect in Fig. 1 is illustrated in Fig. 3a. Although this is an oversimplification insofar as the hybrid population is assumed to be exactly intermediate and populations either side of the zone are assumed to be entirely eastern or western with no gene flow effects, it serves as a worthwhile model for geographic variation across this broad area.

A principal component/coordinate analysis was run on the 'total' set of 71 characters and separately on each of the 6 character systems. The primary geographic variation in the total system and each separate character system and their comparison to the stepped cline

model (Fig. 3a) can be visualized by the projection of the mean scores against the transect (Fig. 3b-f). A range of quantitative measures of congruence (taken as the absolute product-moment correlation coefficient across the 14 scores) and conformity to the stepped cline model are also available.

The congruence between the pattern of variation portrayed by the major component and the stepped cline model is assessed for each character system. Also the average of the correlations between individual characters and the model is computed for each character system as is the percentage of characters showing a significant ($P < 0.05$) correlation with the model. Correlations are used here primarily as a coefficient expressing the quantitative similarity between patterns of geographic variation. It must be borne in mind that even if two geographic patterns 'differ' they may still be significantly correlated as for example in THORPE and BAEZ (1987) where the position of the transition zone in the scalation differs somewhat to that implied by the subspecies even though they are significantly correlated. A study of a similar (but not identical) set of populations has shown that the pattern of geographic variation, based strongly on the total set of characters, is extremely stable (THORPE 1985a), i. e. on average it is not influenced by the inclusion or exclusion of characters when more than about 6 characters are used. Consequently, the component scores for the total set of characters can be used as an end-point against which to judge the geographic variation in a particular character system. As well as the congruence between the primary pattern of a specific character system and the 'total' system, the average correlation between the individual characters in a character system and the 'total-pattern' is also computed.

Canonical variate analysis could only be used with 4 of the character systems, i. e. colour-pattern, internal morphology, scalation and dentition. In each case the major canonical variate was strongly correlated ($r = 0.97$ on average) with the major principal component indicating very similar results. This is predicted from previous studies (THORPE 1980, 1985b) and is largely the result of the very low within group correlation between the characters. The congruence between the primary canonical variate and the stepped cline model is given but otherwise the results are not presented here for the sake of conciseness.

Assessing pattern and their congruence within the hybrid zone

The details of the pattern of change within the zone in northern Switzerland can be represented by the canonical variate scores of female individuals (from Diag. 8, THORPE 1979) projected against the transect of Fig. 2. This transition is illustrated in Fig. 4a and can be loosely approximated by the S curve defined as:

$$S_{cv} = \frac{u + le^{s(i-d)} - (u+l)e^{-sd}}{1 + e^{s(i-d)}}$$

where the S_{cv} is the mean canonical variate score, u is the upper asymptote (eastern value), l is the lower asymptote (western value), d is the distance along the transect, i is the inflexion point (i. e. centre of the hybrid zone) and s is the slope.

Since the individual localities within the zone do not provide sufficient specimens to compute locality mean scores formula 1 is applied to individual female specimens thereby compounding the within and between locality variation. Nevertheless it provides an evaluation of the parameters giving a model S curve for the transition along the transect (Fig. 4a) within the zone which can be used to provide a benchmark comparison for the transition in males. As with the females there are an insufficient number of males per locality to provide meaningful mean scores. Consequently, this investigation of the nature of the within-zone transition has only a supplementary role.

Four character systems (colour pattern, internal morphology, scalation and dentition) and the total sum of these 65 characters can be used to study the transition within the zone. Canonical variate analysis is run on the eastern and western populations and the scores of the 24 male specimens (within the zone) for the major canonical variate are then computed. The congruence between the geographic variation (compounded by individual within locality variation) is then computed between the major canonical variate scores derived from each character system and the scores predicted from the model S curve in Fig. 4a.

Results

Congruence across the hybrid zone

The primary pattern of geographic variation in the total character set is an almost perfect stepped cline as can be seen from visual (Fig. 3 a, b) and quantitative ($r = 0.99$) comparison with the model stepped cline. Of the separate character systems the colour pattern (Fig. 3 c, $r = 0.98$), internal morphology (Fig. 3 d, $r = 0.96$) and dentition (Fig. 3 f, $r = 0.96$) have primary patterns of geographic variation extremely similar to the stepped cline model whilst that of the scalation (Fig. 3 e, $r = 0.89$) and body proportions (Fig. 3 g, $r = 0.78$) are also similar to the stepped cline model. The primary pattern of the total set and the above 5 character systems are all significantly correlated to the stepped cline model at $P < 0.001$. Only the dermal sense organs do not have a primary pattern that reflects the existence of the hybrid zone (Fig. 3 h), i. e. is uncorrelated to the stepped cline model ($r = 0.45$).

The above results are confirmed by the average correlation between the

Fig. 3. Changes across the hybrid zone. The horizontal axis indicates the transect in Fig. 1 with the populations in Fig. 1 in the west-east sequence 3, 1, 6, 15/8, 7, 14, 9, 11/20, 13, 23, 24, 35. The vertical axis indicates the hypothetical score for Fig. 3a and the major principal component score (normalized) for Fig. 3b to h. Fig. 3a shows the model stepped cline with the west-east scores of -1, -1, -1, -1, -1, 0, 1, 1, 1, 1, 1, 1. The predominant patterns of geographic variation is shown for the total character set (3 b), the colour pattern (3 c), the internal morphology (3 d), the scalation (3 e), dentition (3 f), body proportions (3g) and dermal sense organs (3h)

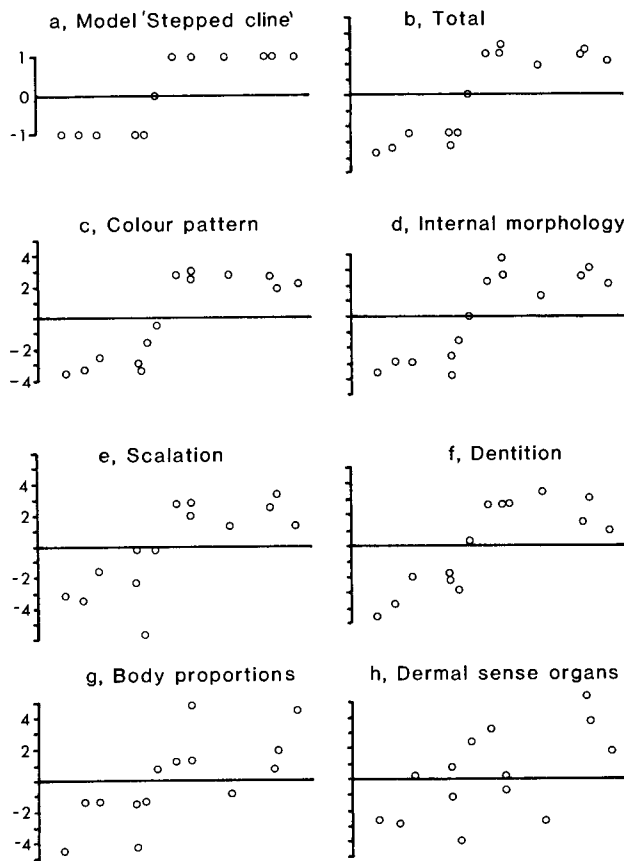


Table 1

Congruence between character systems and model or 'total' patterns of geographic variation

	A	B	C	D	E	F	G
Total	.99	.68	75	/	.70	.98	.74
Colour pattern	.98	.69	74	.98	.69	.98	.89
Internal morphology	.96	.75	85	.99	.76	.93	.44
Scalation	.89	.60	64	.90	.61	.88	.59
Dentition	.96	.89	100	.94	.91	.92	.61
Body proportions	.78	.66	75	.81	.69	/	/
Dermal sense organs	.45	.43	0	.50	.48	/	/

A: Congruence (r) between the major principal component and the stepped cline model of Fig. 3 a ($r = 0.53$ at $p = 0.05$); B: Average congruence (r) between individual characters and the stepped cline model of Fig. 3 a; C: Percentage of individual characters significantly correlated ($r = 0.53$ at $p = 0.05$) to the stepped cline model of Fig. 3 a; D: Congruence (r) between the major principal component of a character system and that of the total character set; E: Average congruence (r) between individual characters and the major principal component of the total character set; F: Congruence (r) between the major canonical variate and the stepped cline model of Fig. 3 a ($r = 0.53$ at $p = 0.05$); G: Congruence (r) between canonical variate scores of individual specimens and the S model of transition (Fig. 4 a) within the hybrid zone ($r = 0.404$ at $p = 0.05$)

individual characters and the stepped cline model and the percentage of individual characters that are significantly correlated to the model (Table 1b, c). The 5 systems, i. e. colour pattern, internal morphology, scalation, dentition and body proportions each have 64% or more of their characters significantly correlated to the model with an average correlation to the model of at least $r = 0.60$, within each character system. Once again only the dermal sense organs contradict this trend and have no characters correlated to the model ($r = 0.43$).

If one uses the primary pattern of geographic variation of the 'total' analysis as the baseline for the comparison rather than the 'stepped cline' model then the same pattern emerges. the primary pattern of geographic variation in the colour pattern, internal morphology, scalation, dentition and body proportions all show a similar change across the hybrid zone to that of the 'total' pattern (r between .81 and .99, Table 1d) and the individual characters within each character system have an average correlation between .61 and .91 (Table 1e) to the changes across the hybrid zone as represented by the 'total' primary pattern. As with the previous comparisons, only the dermal sense organs do not show similar changes across the hybrid zone (Table 1d, e).

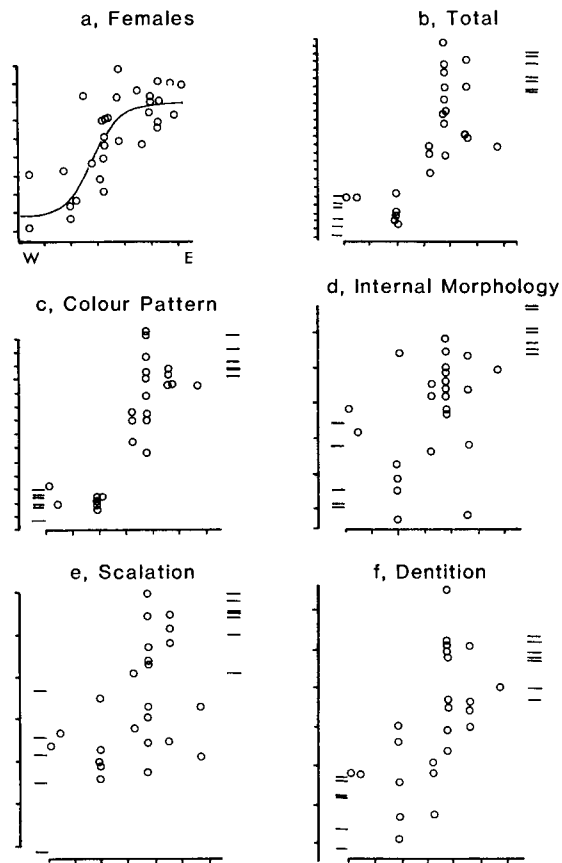
Results from canonical variate analyses were available for the colour pattern, internal morphology, scalation and dentition and the combined total of these character systems. The primary patterns of geographic variation shown by canonical variate analysis of each of these systems are strongly correlated to the stepped cline model (Table 1f) and this supports the results from the principal component analyses.

Congruence within the hybrid zone

The S curve model (formula 1) can be seen to approximate the transition from west to east within the zone for the canonical variate scores of female specimens that were used to evaluate the parameters. Since both geographical and within locality variation are pooled the correlations between the specimen scores and the model cannot be unity since it must, at best, be unity minus the deviation due to within locality variation. With the canonical variate scores for female specimens used to evaluate the parameters (Fig. 4a) the correlation between the scores and the S model was $r = 0.8$.

Bearing in mind the caveat detailed in the Methods section the canonical variate scores from the total analysis (Fig. 4B) can also be approximated by the S curve model of the transition, showing a correlation of .74 with the curve in Fig. 4a (Table 1g). Of the individual character systems, the colour pattern yields canonical variate scores indicating a transition that is strongly correlated ($r = 0.89$) with the S model whilst those of the dentition ($r = 0.61$) and scalation ($r = 0.59$) are slightly less highly correlated with this model. However, even the internal morphology yields canonical variate scores that are significantly correlated ($r = 0.44$, $P < 0.05$) to the S model of transition within the hybrid zone.

Fig. 4. Changes within the hybrid zone. The horizontal axis represents the transect (west to east) in Fig. 2, marked in units of 10 km, whilst the vertical axis represents the canonical variate score, marked in units of within-group standard deviation. Fig. 4a shows canonical variate scores for individual females from which the S curve is computed. Fig. 4b to 4f shows canonical variate scores for individual males for the 'total' character set (Fig. 4b), the colour pattern (Fig. 4c), the internal morphology (Fig. 4d) and the dentition (Fig. 4f). The horizontal bars either side of the transect indicate six western (left) and seven eastern (right) group means from Fig. 1



Discussion

It is apparent that the divergence across the hybrid zone between the western and eastern forms is due to differences accumulated across a large number of characters involving a very wide range of character systems. The broad changes across the hybrid zone predominate in all but one of the character systems irrespective of how this is measured. Moreover, the extent to which each of the 5 character systems show the existence of the hybrid zone is very pronounced.

The only character system not showing a coincident stepped clinal pattern is the dermal sense organs. It is probable that this is, at least in part, due to the number of characters representing this system, i. e. two. It is known that congruence is strongly related to the number of characters when the number of characters is low (THORPE 1985a). Consequently, it is possible that a sample of just two characters, taken from any character system, would not represent the hybrid zone.

Bearing in mind the difficulties of comparing the patterns of change within the zone

all 4 tested character systems showed a sharp step transition within the zone commensurate with the congruence indicated by the "across-zone" comparisons.

The primary aim of this study is to test the hypothesis that when a hybrid zone is formed due to secondary contact after vicariance over a substantial period of time (as opposed to primary differentiation due to adaptation to current ecological conditions) then the transition across the zone should be the predominant pattern of geographic variation in a large number of characters (THORPE 1985a) and a wide range of character systems. Since the east/west hybrid zone in *N. natrix* is independently hypothesised to be of secondary origin (THORPE 1984a) we can relate its nature to its cause without circularity and show that in this case, a secondary contact zone has resulted in the predominant pattern in a wide range of character systems being a congruent "stepped cline".

Since the eastern and western lineages either side of the hybrid zone are borderline species (THORPE 1979) this study indicates, that in this case, the difference between sister 'species' are reflected in numerous, relatively small changes, distributed throughout most of the character systems studied rather than major changes in one system. A large number of quantitative characters are significantly involved in this, i. e. 53 characters, or 75 % of the total showing geographic variation. These independent, quantitative characters are likely to reflect variation in a very large number of loci. As well as indicating the nature of differences between closely related sister species and the consequent nature of changes in allopatric speciation it is interesting to note that this *N. natrix* zone contradicts the generalization that zones involving many loci tend to be broader than those involving few. This zone is likely to involve a large number of loci yet, in some areas, it is one of the sharpest zones studied (THORPE 1987).

The congruence between the character systems is unlikely to be due to a genetic revolution during founder effect speciation. The size of the refugia can be estimated from current distribution in relation to the ice caps and post ice-cap distribution (THORPE 1984b). Whilst the hypothesised refugia are likely to be only a minute fraction of the current range, they were unlikely to be small enough to produce the bottleneck populations necessary for the complete developmental reorganisation during quantum speciation sensu STANLEY (1979).

Zusammenfassung

Übereinstimmungen zwischen unabhängigen Merkmalskomplexen in einer Hybridzone: Muster im geographischen Raum

Ringelnattern östlicher und westlicher Herkunft werden in einem Hybridisierungsbereich, von dem angenommen wird, daß er durch einen sekundären Kontakt zustande kam, in einer großen Zahl voneinander unabhängiger, quantitativer Merkmale mit Hilfe univariater und multivariater Analysen untersucht. Die Untersuchung der geographischen Variationen lehrt, daß fünf von sechs Merkmalskomplexen einen „stepped clinal“ Wandel zeigen. Im erörterten Beispiel zeigen sich in der sekundären Hybridzone konkordante Veränderungen in mehreren Merkmalskomplexen und die „Speciation“ besteht in stärkerem Ausmaß in der Entstehung einer großen Zahl kleiner Unterschiede in den meisten Merkmalskomplexen als in größerer Umorganisation eines Einzelsystems.

Literature

- BARTON, N. H.; HEWITT, G. M., 1985: Analysis of hybrid zones. *Ann. Rev. Ecol. Syst.* **16**, 113–148.
- ENDLER, J. A., 1977: *Geographic variation, speciation and clines*. Princeton, N. J.: Princeton Univ. Press.
- 1982: Problems in distinguishing historical from ecological factors in biogeography. *Amer. Zool.* **22**, 441–452.
- 1983: Testing causal hypotheses of geographical variation. *Numerical Taxonomy: Proc. NATO Advanced Studies Inst.* Ed. by FELSENSTEIN, J.; NATO Advanced Studies, Series G (Ecol. Sci.) No. 1. Berlin, Heidelberg, New York: Springer Verlag, pp. 424–443.
- 1986: *Natural selection in the wild*. Monographs in population biology 21. Princeton, N. Y.: Princeton Univ. Press.

- STANLEY, S. M., 1979: Macroevolution. Patterns and Process. San Francisco: W. H. Freeman & Co.
- THORPE, R. S., 1979: Multivariate analysis of the population systematics of the ringed snake *Natrix natrix*. Proc. R. Soc. Edinburgh **788**, 1–62.
- 1980: A comparative study of ordination techniques in numerical taxonomy in relation to racial variation in the ringed snake *Natrix natrix* (L.). Biol. J. Linn. Soc. **13**, 7–40.
 - 1982: Reticulate evolution and cladism: tests for the direction of evolution. Experientia **38**, 1242–1243.
 - 1983: Phylogenetic analysis of range expansion in the grass snake: reticulate evolution: primary and secondary contact zones. Numerical Taxonomy: Proceedings of a NATO Advanced Studies Institute. NATO Advanced Studies Institute Series G (Ecological Sciences), No. 1, Berlin, Heidelberg, New York: Springer Verlag. pp. 464–468.
 - 1984a: Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. Evolution **38**, 233–243.
 - 1984b: Geographic variation in the western grass snake (*N. natrix helvetica*) in relation to hypothesised phylogeny and conventional subspecies. J. Zool. Lond. **203**, 245–355.
 - 1985a: Character number and the multivariate analysis of simple patterns of geographic variation: categorical or ‘stepped clinal’ variation. Syst. Zool. **34**, 127–139.
 - 1985b: Clines: character number and the multivariate analysis of simple patterns of geographic variation. Biol. J. Linn. Soc. **26**, 201–214.
 - 1987: Geographic variation: a synthesis of cause, data, pattern and congruence in relation to subspecies, multivariate analysis and phylogenesis. Bull. Zool. **53** (in press).
- THORPE, R. S.; BAEZ, M., 1987: Geographic variation within an island: univariate and multivariate contouring of scalation, size and shape of the lizard *Gallotia galloti*. Evolution **41**, 256–268.

Author's address: Dr. ROGER S. THORPE, Dept. of Zoology, University of Aberdeen, Tillydrone Avenue, Aberdeen, AB9 2TN, Scotland

Erratum

Zeitschrift für Zoologische Systematik und Evolutionsforschung
Band 25 (1987), Heft 3

VERLAG PAUL PAREY

In the paper ‘Congruence between independent character systems across a hybrid zone: patterns in geographic space’ by R. S. THORPE the formula and the legend to the formula on page 164 are given incorrectly by an error. They should read as follows:

$$S_{cv} = \frac{u + le^{s(i-d)} - (u+l)e^{-sd}}{1.0 + e^{s(i-d)}}$$

where the S_{cv} is the mean canonical variate score, u is the upper asymptote (eastern value), l is the lower asymptote (western value), d is the distance along the transect, i is the inflexion point (i. e. centre of the hybrid zone) and s is the slope.

Since the individual localities within the zone do not provide sufficient specimens to compute locality mean scores formula 1 is applied to individual female specimens thereby compounding the within and between locality variation. Nevertheless it provides an evaluation of the parameters giving a model S curve for the transition along the transect (Fig. 4 a) within the zone which can be used to provide a benchmark comparison for the transition in males. As with the females there are an insufficient number of males per locality to provide meaningful mean scores. Consequently, this investigation of the nature of the within-zone transition has only a supplementary role.