

## Geographic variation in the Western grass snake (*Natrix natrix helvetica*) in relation to hypothesized phylogeny and conventional subspecies

R. S. THORPE

Department of Zoology, University of Aberdeen, Aberdeen, Scotland, U.K.

(Accepted 11 October 1983)

(With 6 figures in the text)

Ordination analysis of the geographic variation in the Western grass snake reveals a V pattern with the northern populations at the apex and the south-eastern (Italian) populations differentiated clinally down one arm and the south western (Ibero-African) populations differentiated categorically along the other arm. There is a clear relationship between latitude and extent of differentiation; the populations in the south west being particularly highly differentiated. These phenetic patterns can be closely related to the numerically hypothesized phylogeny and can be explained by Pleistocene events. The relationship between the rejected conventional subspecies and these patterns is discussed and it is shown that some conventional subspecies are derived from arbitrarily sectioned clines and delimited by inappropriate physiographic features.

### Contents

	Page
Introduction .. .. .	345
Methods .. .. .	347
Samples .. .. .	347
Characters .. .. .	347
Numerical methods .. .. .	347
Results .. .. .	348
Geographic networks .. .. .	348
Ordination analyses .. .. .	348
Allopatric populations .. .. .	350
Population differentiation and sex .. .. .	351
Discussion .. .. .	352
References .. .. .	355

### Introduction

Quantitative analyses of the grass snake (*Natrix natrix*) species complex reveal two large phenetic categories at the semispecies level, i.e. the Eastern and Western grass snakes (Thorpe, 1975a, 1979, 1980a). These two main groups, and two island outlyers, are given formal trinomials (Thorpe, 1979); the Western grass snake being *N. natrix helvetica*. Numerical phylogenetic analysis (Thorpe, 1982, 1983a, In press b) indicates that this western subspecies, although originally based on phenetic criteria, is also monophyletic.

In central Europe this Western grass snake meets the Eastern grass snake (*N. n. natrix*) along a "hybrid" or transition zone, the exact nature of which is discussed in Thorpe (1979, 1980*b*). The east-west differentiation in *Natrix natrix* is similar to that in many other palaeartic species and it is probable that this pattern is the result of the advancing Pleistocene ice caps separating south-east and south-west refugia which expanded to meet along a "secondary" zone in post-glacial times, (Thorpe, 1975*a*). This is considered in depth with an alternative hypothesis by Thorpe (1979) and discussed in relation to allopatric speciation by White (1978). Phylogenetic analysis of range expansion (Thorpe, 1983*a*, In press *b*) also hypothesizes a secondary origin for this east/west zone.

The Western grass snake is distributed across North Africa, Iberia, Italian peninsula, France, Britain and parts of the Low Countries, Germany and Switzerland. It is also found on several islands including Sicily, Elba and Jersey (but not Corsica or Sardinia, which are inhabited by *N. n. corsa* and *N. n. cetti*, respectively).

This range transcends several 'conventional' subspecies (Hecht, 1930; Mertens, 1947; Kramer, 1970), i.e. subspecies recognized on conventional criteria *sensu* Thorpe (1980*b*, 1981) rather than phenetic or cladistic criteria. The current conventional subspecies within the range of the Western grass snake are as follows: *N. n. astreptophora* which occupies North Africa and Iberia, *N. n. helvetica* which occupies Britain, France, Germany west of the Rhine, Switzerland and northern Italy, *N. n. lanzai* which occupies mid Italy, *N. n. sicula* which occupies the southern tip of Italy and Sicily; and finally the western edge of *N. n. natrix* which occupies the area between the Rhine and the east/west transition zone.

Variation within the Western grass snake has previously been subject to numerical phylogenetic analysis (Thorpe, 1982, 1983*a*, In press *a*, In press *b*). These analyses reveal, 1) a close association between the geographical location of the population and the position along the cladogram; 2) a pattern of anagenesis that indicates a significant direction of phylogenetic change so that "phylogenesis" rather than "current ecology" is the most likely cause of the intraspecific variation (Thorpe, 1982, 1983*a*, In press *b*); 3) two sublineages, one with an Ibero/African root and one with a southern Italian root; 4) introgression between these two sublineages in northern Europe resulting in reticulate evolution (Thorpe, 1982).

The microphylogenesis in the Western grass snake can be interpreted in terms of range expansion due to Pleistocene ice retreats and advances (Thorpe, 1982, 1983*a*, In press *b*). It is evident from geological studies (West, 1968) that during the last Pleistocene advance (Weichsel/Waldai maximum) the ice cap extended down to *circa* 51.50° North in some western regions. The current distribution of the grass snake is *circa* 10° latitude lower than the Spitzbergen and Novaya Zemlya glaciers, respectively. Therefore, if the Pleistocene population of grass snakes kept its 10° latitude south of the ice caps it would be depressed to *circa* 41.50°. It is hypothesized (Thorpe, 1982, 1983*a*, In press *b*) that the western Pleistocene refugium was in Iberia-North Africa and that as the climate began to get warmer, about 20,000 years ago, the population expanded northwards out of Iberia and through southern France to colonize Italy. Then as the climate became suddenly colder again in the Late Weichselian (West, 1968) the population was depressed into temporary Iberian and Italian refugia. As the ice retreated again, *circa* 10,000 years ago, the populations expanded northwards from these two refugia to introgress in northern Europe.

The phenetic patterns of geographic variation in the Western grass snake have not previously been described. It is the aim of this study to describe these patterns and relate them to the conventional subspecies and hypothesized microphylogeny.

## Methods

### *Samples*

As the number of specimens per locality is generally insufficient for statistical purposes several localities are grouped into compound localities (referred to as samples) on the basis of collecting gaps and similarity of the individual specimens (Thorpe, 1979) assessed by multivariate analysis. The heterogeneity within the compound localities is commensurate with the required degree of resolution (Thorpe, 1976) and any geographical subdivision of the samples would not fundamentally reduce this heterogeneity.

The 15 sample localities which are illustrated in Thorpe (1980a, Fig. 1) are not always represented by both sexes. Analyses are always run separately for the sexes.

### *Characters*

Of the 160 characters initially recorded 41 were selected on the basis that they showed statistically significant geographic variation (analysis of variance) between 15 western samples for both sexes and that they had a low pooled within-sample correlation with other included characters (Thorpe, 1979). These 41 independent characters express a considerable amount of information and represent a range of character systems, i.e. colour pattern (20), scalation (10), internal morphology (5), dentition (3), and body proportions (3). The body proportions are head width, head length and body width. The other 37 characters are listed in Thorpe (In press, a).

The quantification of these characters is discussed by Thorpe (1975b) and the adjustment of the linear dimensions for growth independence is discussed by Thorpe (1975b, 1976).

### *Numerical methods*

The various multivariate techniques for assessing population affinities using quantitative phenotypic characters are reviewed in Thorpe (1976, 1983b). In this study geographical networks and ordination techniques are used based on taxonomic distance ( $d$ ) between sample means and Mahalanobis  $D^2$  between samples (canonical analysis). For the principal coordinate studies (based on taxonomic distance) the characters are standardized to zero mean and unit standard deviation to remove any scaling effect (Thorpe, 1976, 1979, 1980a). The  $D^2$  and canonical analyses are run on 37 rather than 41 characters because the adjusted linear dimensions (4 characters) are not available for all specimens (Thorpe, 1975b, 1979).

The direction of phenotypic change over geographic distance is illustrated by plotting canonical variate scores against geographical transects whilst the extent of phenotypic change per unit distance is illustrated by plotting the adjusted  $d$  or  $D^2$  values between adjacent contiguous populations against latitude. The latitudinal location of the metric is taken as halfway along the line joining the geographical centres of the samples.

The geographical distance between sample centroids differs slightly so the  $d$  or  $D^2$  values are adjusted by

$$\hat{\delta} = \frac{\delta}{\log_{10} \text{ km}}$$

where  $\delta$  is the similarity valued  $d$  or  $D^2$ ,  $\hat{\delta}$  is the adjusted value and km is the linear distance (in kilometres) between the geographical centres of the samples.

Bar charts (Fig. 6) are used to illustrate the divergence of allopatric populations. These are based on the analyses in Thorpe (1979) where a greater number of characters and samples were used.

## Results

### *Geographic networks*

The  $d$  and  $D^2$  values plotted between samples for females (Fig. 1(a)) and males (Fig. 1(b)) show that the populations from the south west (Ibero-African) have no close affinity for the populations from the south east (Italian peninsula).

There is also a general tendency for the level of differentiation between adjacent (contiguous) populations to decrease further north. This is illustrated more clearly by the plots of adjusted phenetic distance against latitude (Fig. 2). These plots show quite clearly that there is a high level of differentiation between contiguous populations in the south (especially in Ibero-Africa) which decreases quite markedly as one goes further north.

### *Ordination analyses*

The similarity between samples was summarized by canonical and principal coordinate analyses run separately for males and females. There is an extremely close agreement between all four of these ordination analyses, consequently only the canonical analysis of males is illustrated (Fig. 3).

The most obvious feature of these analyses is the V-shaped pattern of the scatter of samples with the northern populations at the crotch and the south east (Italian) and south west (Ibero-African) populations extending down each arm.

The first canonical variate expresses a south-west projection whilst the second canonical variate expresses a south-east projection.

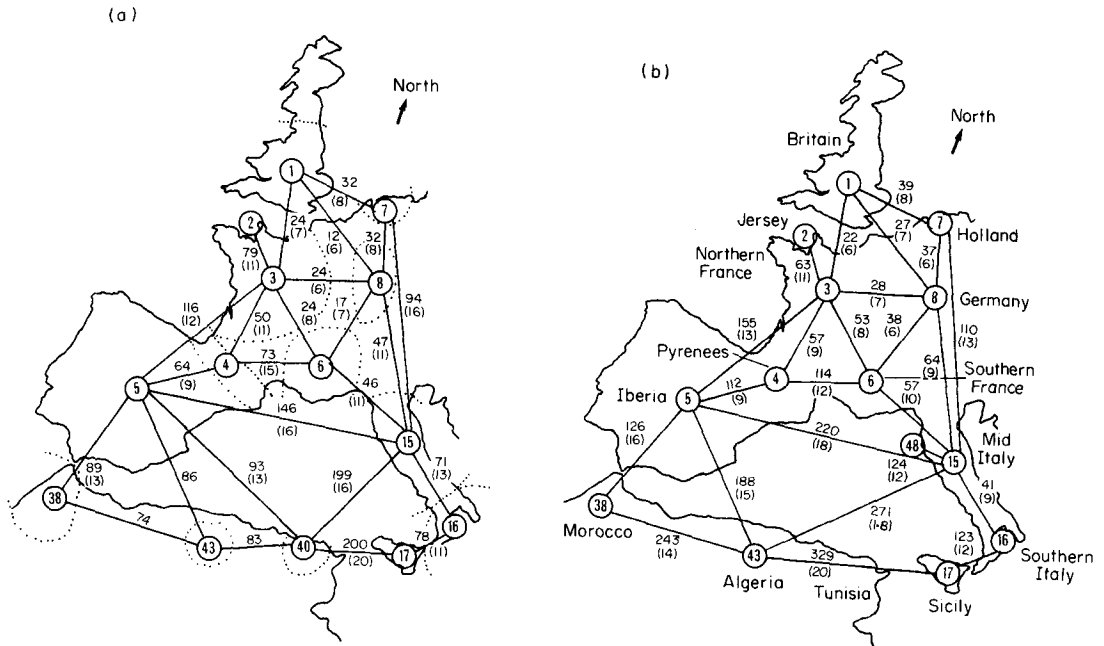


FIG. 1. Network diagrams with  $D^2$  and  $d$  values between (a) females and (b) males. Taxonomic distance values are multiplied by 10 to remove the decimal point and are bracketed. Sample numbers are encircled and the sample areas are indicated by a broken line in (a).

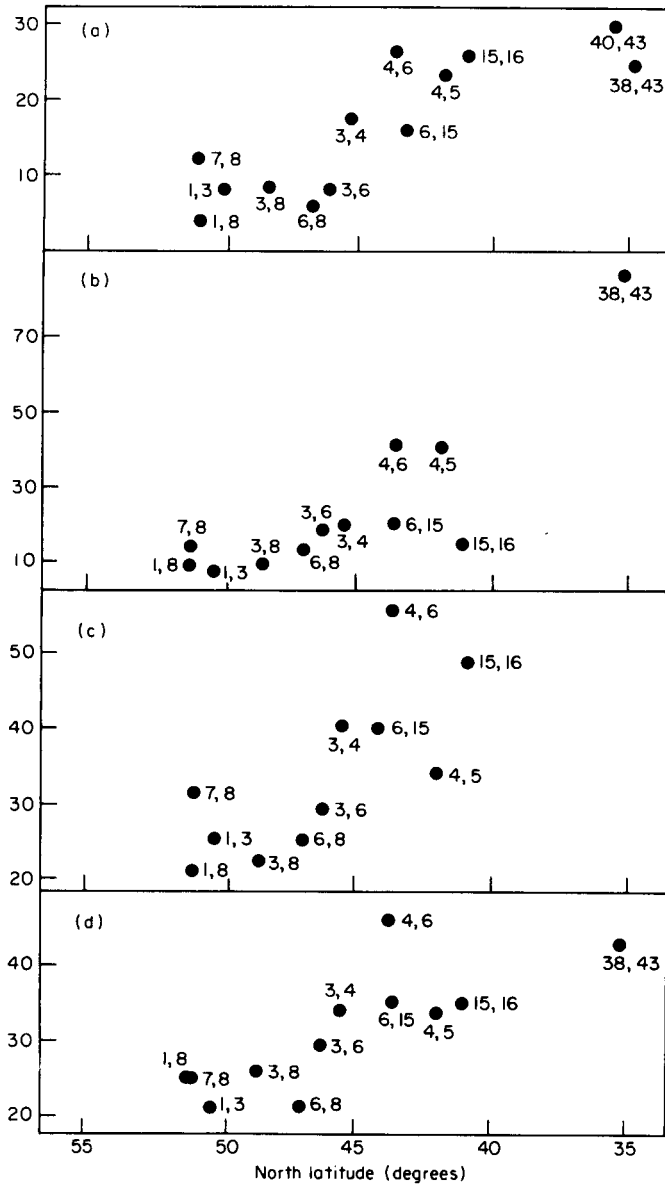


FIG. 2. The adjusted  $D^2$  and  $d$  values between adjacent-contiguous populations (vertical axis) plotted against latitude. Figures 2(a) and (b) are based on adjusted  $D^2$  between females and males, respectively and 2(c) and (d) are based on adjusted  $d$  (taxonomic distance) between females and males, respectively. Sample numbers as in Fig. 1.

The nature of the geographical variation along the south-western and south-eastern projections differs.

If one takes a south-westerly transect and plots the first canonical variate scores of individual male specimens against this (Fig. 4) the categorical nature of the geographic variation

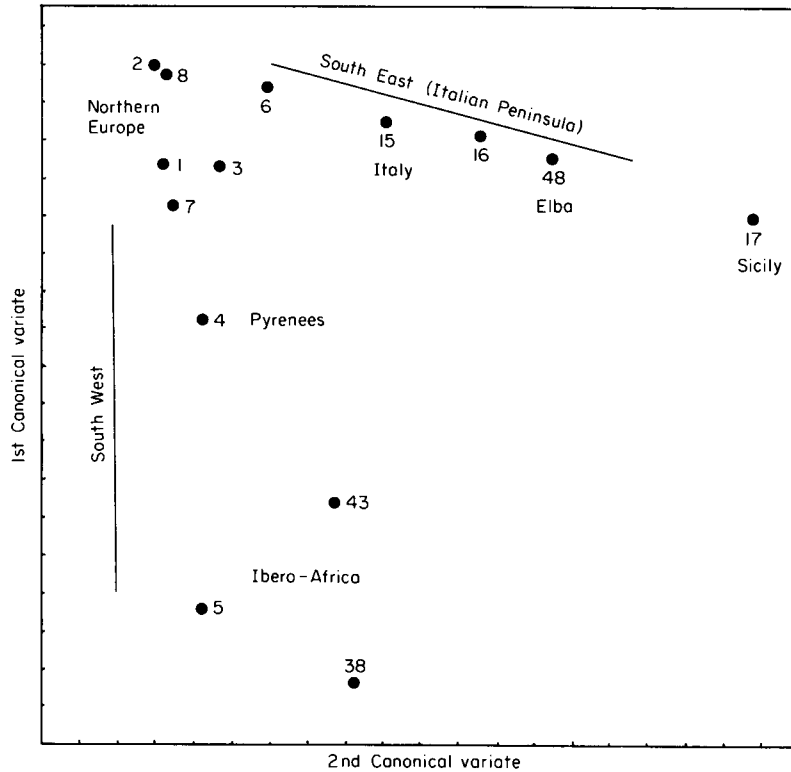


FIG. 3. Canonical analysis of male samples. The first and second variates (marked in standard deviations) express 38 and 27% of the variation, respectively. Ninety-five per cent of the individuals in a sample will be within approximately two standard deviations of the centroid. Sample numbers as in Fig. 1.

is apparent. There is clearly a northern category (e.g. Britain, Holland and France) which is distinct from a southern category (Iberia and north Africa) with the geographically and phenotypically intermediate population from the Pyrenees forming a step between these categories. It must be borne in mind that the first canonical variate does not express all the variation between these samples.

If one takes a south-easterly transect and plots the second canonical variate scores (Fig. 5) it is apparent that the mainland and northern island populations (southern Italy, mid-Italy, southern France, northern France, Holland, Jersey and Britain) vary clinically along this projection. This clinical variation is in distinct contrast to the categorical variation along the south-west projection.

#### *Allopatric populations*

The allopatric populations of the Western grass snake include the island populations of Britain, Jersey, Elba and Sicily.

As a generalization island populations tend to be divergent and this is true of most of the studied island populations of the Western grass snake. The noticeable exception is the British population which is extremely similar to the other northern European population (Fig. 5). The phenetic distance between the British population and adjacent mainland population

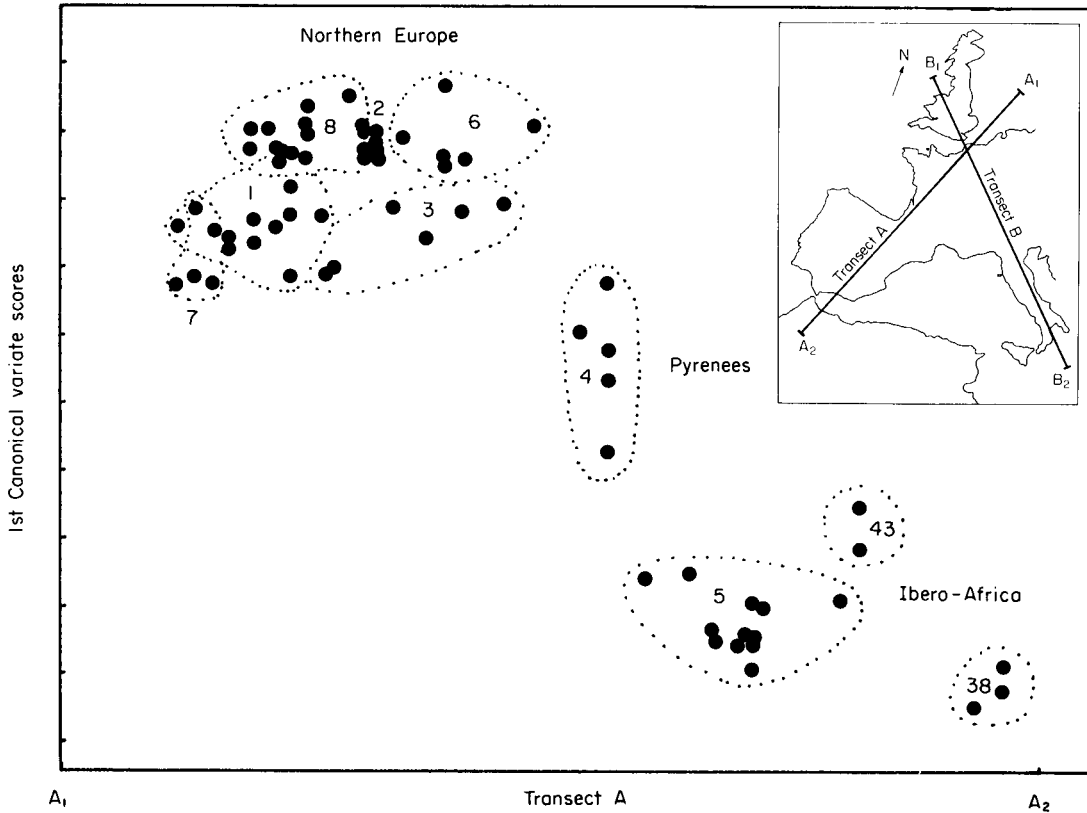


FIG. 4. Canonical variate (1st axis) scores of individual male specimens plotted against transect A (inset). Canonical variate (vertical axis) marked in units of two standard deviations. Sample numbers as in Fig. 1.

is basically the same as the difference between contiguous mainland populations in that northern region (Fig. 1).

The island populations of Jersey, Sicily and Elba are more divergent, at least in males, (Fig. 6) but do not approach the excessive divergence that can be exhibited by some island populations, such as that of Sardinia (*N. n. cetti*). The divergence of Jersey, Sicily and Elba are all generally greater than the level of divergence generally found between contiguous populations on the nearby mainland (Fig. 1).

Some characters, e.g. reduction in lunar markings (Thorpe, *In press a*), show convergence on island populations. This convergence has not influenced the general phenetic patterns insofar as Jersey and the Mediterranean island populations do not appear phenetically similar to one another but similar to their adjacent mainland populations (Fig. 3).

#### *Population differentiation and sex*

There is little difference between the sexes in their pattern of population differentiation whatever method (geographic networks, ordination, bar charts) is used. However, it is interesting to note that males do show greater levels of population differentiation than females

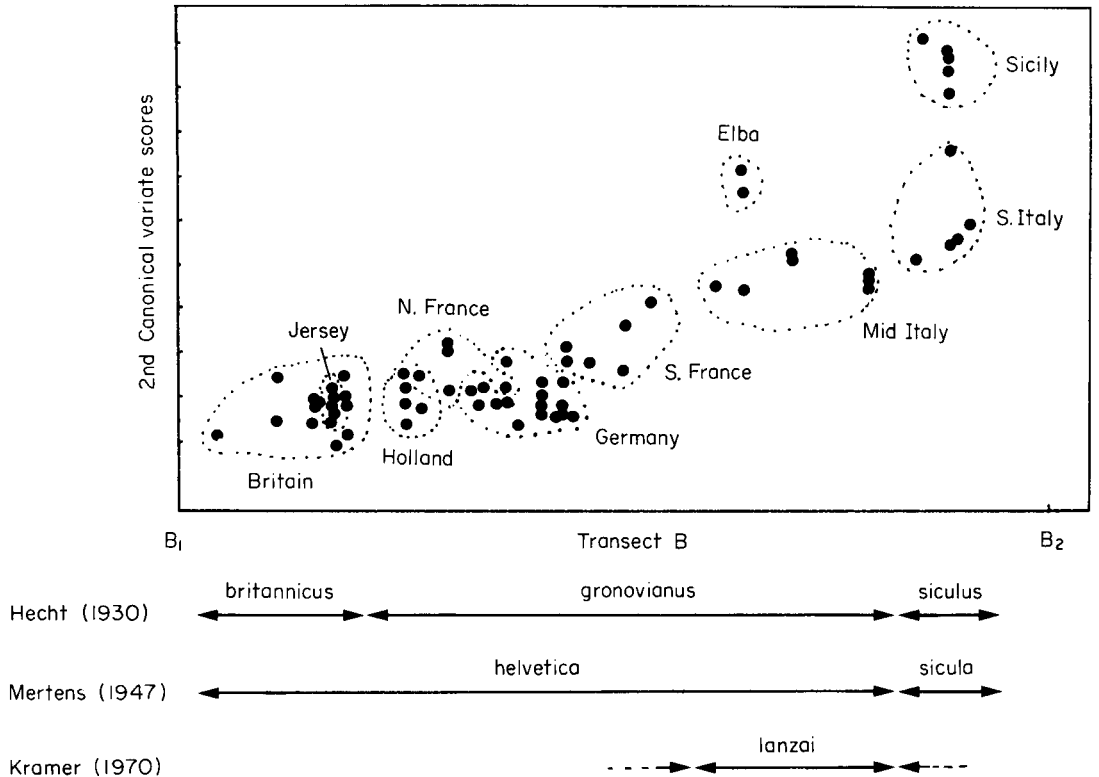


FIG. 5. Canonical variate (2nd axis) scores of individual male specimens plotted against transect B (inset Fig. 4). Canonical variate marked in units of two standard deviations. The sectioning of this cline into various conventional subspecies is indicated. Hecht (1930) recognized *britannicus*, (Britain), *gronovianus* (France and Italy), and *siculus* (Sicily and Calabria). Mertens (1947) recognized *helvetica* (Britain, France and Italy) and *sicula* (Sicily and Calabria) to which Kramer (1970) added *lanzai* (mid Italy).

even when the pattern of racial differentiation is the same (Figs 1, 6). This is apparent in previous studies of this species complex and of multivariate morphometric studies of other snakes (Benton, 1980).

### Discussion

The main phenetic pattern of geographic variation within the Western grass snake is clearly the divergence of the "Italian" and "Ibero-African" forms down two "arms" with the northern populations at the apex (Fig. 3). This is entirely commensurate with previous numerical phylogenetic analyses (Thorpe, 1982, 1983a, In press b) which reveal separate Ibero-African and Italian sublineages which introgress in northern Europe resulting in reticulate evolution. In this sense there is very close agreement between the phenetic and phylogenetic analyses.

The populations occupying the position of the hypothesized temporary refugia, i.e. Iberia/Africa and southern Italy, are among the most phenetically divergent populations as shown by ordination analyses and geographical networks. The northern populations of the



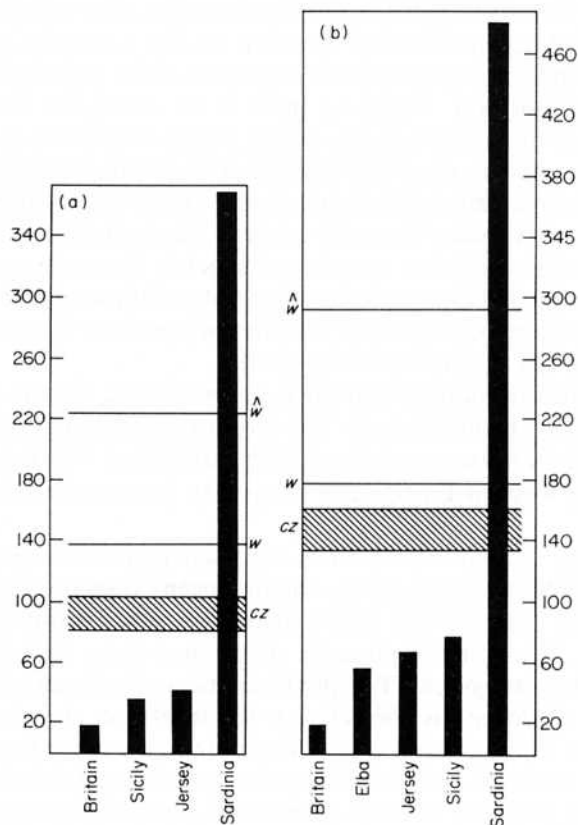


FIG. 6. Bar chart showing divergence of allopatric populations from their most similar adjacent mainland population. Vertical axis is  $D^2$  for (a) females and (b) males based on analyses in Thorpe (1979, figs 8, 9). The difference between terminal populations in a contiguous series (Iberia and Italy) is indicated by  $w$ , the difference between the most dissimilar pair of populations (allopatric populations included) is indicated by  $\hat{w}$ . For comparison the divergence of the Sardinian population is shown, as is the divergence associated with incipient speciation,  $CZ$  (difference across the contact zone between eastern and western incipient species).

Ibero-African lineage, e.g. northern France and Britain, are only slightly phenetically divergent (Figs 1, 3, 4) from the northern populations of the Italian lineage, e.g. Germany and southern France.

Whilst this phenetic pattern relates to the hypothesized phylogeny it cannot be readily explained by natural selection for current conditions, i.e. an "ecological" causation. Contemporary physical environmental factors are broadly similar climatically in Iberia/North Africa and Italy insofar as they have a Mediterranean climate. Summer temperature at sea level is similar (*circa* 24/26°C) as is the precipitation (Escardo, 1970; Cantu, 1977).

The other pronounced feature of the phenetic pattern of geographic variation is the greater differentiation between adjacent populations in the southern latitudes (Fig. 2). This can also be readily related to the hypothesized phylogeny. I have argued elsewhere (Thorpe, 1983a, In press b) that a pattern of progressive change in the extent of anagenesis in an intraspecific "phylogeny" indicates a clear phylogenetic direction of evolution and hence a

“phylogenetic” rather than “ecological” cause for the geographic variation. Selection for current conditions may well result in clinal variation, so that numerical phylogenetic analysis produces a “tree” in which the phylogenetic sequence of the populations is in accord with their actual geographic position. However, there is no ecological reason why this “tree” should have a pattern of progressive change in the extent of anagenesis.

In the Western grass snake, as in the Eastern grass snake, the numerical phylogenetic schemes show a progressive pattern of change in the extent of anagenesis and hence a clear phylogenetic direction of evolution from the roots in the south (where the populations have had time to accumulate considerable anagenic change) to the more recently evolved populations in the north (which had time to accumulate only a little anagenic change). This relates to the above phenetic pattern of geographic variation where there is more divergence between adjacent populations in the south than in the north.

On the other hand this latitudinal pattern is not explained by selection pressures in the south, being more spatially heterogeneous. The physical factors such as summer temperature and precipitation that have been recorded and mapped (Wallen, 1970; Arlery, 1970; Escardo, 1970, Cantu, 1977) for western Europe can be seen to be as diverse in the north as in the south.

Whilst there is a close relationship between the phenetic patterns of geographic variation and the previously hypothesized phylogeny the phylogeny does not allow one to predict all of the phenetic patterns. For example, the phenetic divergence down the leg of Italy (Italian lineage) is clinal (Fig. 5) whilst the phenetic divergence down into the Iberian peninsula (Ibero-African lineage) is categorical. The phylogenetic analysis cannot indicate this.

If the hypothesized phylogeny is correct then the transition zone in the Pyrenees is not at the point where two lineages meet and is therefore not a secondary transition zone but a primary transition zone, i.e. due to differentiation of the populations *in situ* (Thorpe, 1983*a*, In press *b*). The phenetic analysis cannot indicate this. Consequently both phenetic and phylogenetic analyses can be necessary to understand patterns of geographic variation.

The northern island populations of Jersey and Britain show a great disparity in their level of phenetic divergence from the mainland even though phylogenetic analysis hypothesizes a recent colonization for both these islands. Britain was connected to the mainland by a land bridge until *circa* 8000 years ago and it is thought that the retreat of the ice enabled mainland populations to expand into Britain prior to its isolation. The absence of the founder effect may explain the comparatively low level of divergence of the British population, i.e. no more divergent than adjacent mainland populations at that latitude. The Jersey population may be more divergent because of the founder effect as Jersey was not recently joined to France and could be colonized only by founder individuals floating or rafting from France to this small island. Genetic drift and strong selection on this small population could have contributed in this divergence.

The Western grass snake *sensu* Thorpe (1979) is a phenetic and phylogenetic entity that covers the complete range of four, and part of the range of a fifth, conventional subspecies. There is therefore basic disagreement between the conventional taxonomy and this phenetic and phylogenetic taxon. Details of this disagreement illustrate some of the limitations of “conventional” subspecies.

Even when distinct phenetic and phylogenetic entities do exist, i.e. the Eastern and Western grass snakes, the conventional subspecific boundaries tend to be linked to physiographic features i.e. the Rhine, rather than coinciding with the actual transition zone.

Clines are also likely to be arbitrarily sectioned by conventional subspecies and the grass snake provides an example of this. The phenetic variation from northern Europe down the leg of Italy is clinal (Fig. 5). Conventional taxonomists have arbitrarily sectioned this cline into subspecies but each author has a different scheme (Fig. 5). Sicily plus the southern tip of Italy is, however, consistently recognized as a conventional subspecies but the phenetic analyses show the populations at the tip of Italy to be part of the cline whilst the Sicilian population is divergent (Fig. 5). These conventional subspecies are no more in accord with the hypothesized phylogeny than with the phenetic patterns of geographic variation.

## REFERENCES

- Arlery, R. (1970). The climate of France, Belgium, The Netherlands and Luxembourg. In *Climates of Northern and Western Europe, World Survey of Climatology* 5: 135–193. Wallen, C. C. (Ed.). Amsterdam: Elsevier.
- Benton, M. J. (1980). Geographic variation in the garter snakes (*Thamnophis sirtalis*) of the north-central United States, a multivariate study. *Zool. J. Linn. Soc.* 68: 307–323.
- Cantu, V. (1977). The climate of Italy. In *Climates of Central and Southern Europe. World Survey of Climatology* 5: 127–183. Wallen, C. C. (Ed.). Amsterdam: Elsevier.
- Escardo, A. L. (1970). The climate of the Iberian Peninsula. *World Survey of Climatology* 5: 195–239. Wallen, C. C. (Ed.). Amsterdam: Elsevier.
- Hecht, G. (1930). Systematik, Ausbreitungsgeschichte und Oekologie der europäischen Arten der Gattung *Tropidonotus* (Kuhl) H. Boie. *Mitt. zool. Mus. Berl.* 16: 244–393.
- Kramer, E. (1970). Revalidierte und neue Rassen der europäischen Schlangenfauna. *Lav. Soc. Ital. Biogeogr.* 1: 667–676.
- Mertens, R. (1947). Studien zu Eidonomie und Taxonomie der Ringelnatter (*Natrix natrix*). *Abh. senckenb. naturforsch. Ges.* No. 476: 1–38.
- Thorpe, R. S. (1975a). Biometric analysis of incipient speciation in the ringed snake *Natrix natrix* (L.). *Experientia* 31: 180–182.
- Thorpe, R. S. (1975b). Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the ringed snake *Natrix natrix* (L.). *Biol. J. Linn. Soc.* 7: 24–43.
- Thorpe, R. S. (1976). Biometric analysis of geographic variation and population affinities. *Biol. Rev.* 51: 407–452.
- Thorpe, R. S. (1979). Multivariate analysis of the population systematics of the ringed snake *Natrix natrix* (L.). *Proc. R. Soc. Edinb.* 78B: 1–62.
- Thorpe, R. S. (1980a). 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.
- Thorpe, R. S. (1980b). Microevolution and taxonomy of European reptiles with particular reference to the grass snake *Natrix natrix* and the wall lizards *Podarcia sicula* and *P. melisellensis*. *Biol. J. Linn. Soc.* 14: 215–233.
- Thorpe, R. S. (1981). Racial affinities and geographic variation: a comparison of the use of multivariate morphometrics and conventional subspecies in relation to European reptiles: 5–7. In *Proceedings of the European Herpetological Symposium 1980*. Coburn, J. (Ed.) Burford, Oxon: Cotswolds Wildlife Park.
- Thorpe, R. S. (1982). Reticulate evolution and cladism: tests for the direction of evolution. *Experientia* 38: 1242–1244.
- Thorpe, R. S. (1983a). Phylogenetic analysis of range expansion in the grass snake: reticulate evolution: primary and secondary contact zones. In *Numerical taxonomy*: 464–468. Felsenstein, J. (Ed.). Berlin, Heidelberg, New York: Springer. (Proc. of NATO ASI Series G, 1).
- Thorpe, R. S. (1983b). A review of the numerical methods for recognising and analysing racial differentiation. In *Numerical taxonomy*: 404–423. Felsenstein, J. (Ed.). Berlin, Heidelberg, New York: Springer. (Proc. of NATO ASI Series G, 1).
- Thorpe, R. S. (In press a). Coding morphometric characters for constructing Distance Wagner networks. *Evolution, Lawrence, Kans.*
- Thorpe, R. S. (In press b). Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. *Evolution, Lawrence, Kans.*
- Wallen, C. C. (1970). Introduction. *Climates of Northern and Western Europe World Survey of Climatology* 5: 1–21. Wallen, C. C. (Ed.). Amsterdam: Elsevier.
- West, R. G. (1968). *Pleistocene geology and biology*. London: Longmans.
- White, M. J. D. (1978). *Modes of speciation*. San Francisco: W. H. Freeman & Co.