

## ECOLOGICAL MODELLING OF GENETICALLY DIFFERENTIATED FORMS OF THE IBERIAN ENDEMIC GOLDEN-STRIPED SALAMANDER, *CHIOGLOSSA LUSITANICA*

J. W. ARNTZEN<sup>1,2</sup> AND J. ALEXANDRINO<sup>3</sup>

<sup>1</sup>National Museum of Natural History, The Netherlands

<sup>2</sup>Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Portugal

<sup>3</sup>Museum of Vertebrate Zoology, University of California, USA

We modelled the distribution of two genetically differentiated forms of the Golden-striped salamander, *Chioglossa lusitanica*, in Portugal with the help of a geographical information system. Models were derived with and without the assumption that the contact zone between the forms would be associated with the Mondego river valley and with and without the statistical Bonferroni correction. The model fit was high and ranged from  $kappa = 0.81-0.99$ . The southern form appears to live under harsher (but still tolerated) environmental conditions than the northern form (low precipitation, low air humidity, low summer temperatures, a high number of frost months, low NDVI vegetation index) and can be viewed as a mountain form. This is in line with the observation that the southern form has shorter extremities than the northern one.

*Key words:* amphibian distribution, GIS, Portugal, Urodela

### INTRODUCTION

For many widespread animal and plant species in Europe, genetic variation is more frequently observed in – or even restricted to – southern areas that remained largely unglaciated during the Pleistocene. These areas may have served as refugia during cold periods and as sources to the recolonization of Europe thereafter. The present-day general pattern of substantial genetic variation in the south and a low, depleted genetic variation in the north, is a witness to this process (Hewitt 1996, 1999). Endemic species are frequently taken as morphologically and genetically uniform. This conviction may be *a priori*, or be vaguely inferential, e.g. when it is suggested that a small range limits the scope for geographical subdivision and that a specialized niche limits the scope for differential adaptation. Current molecular techniques, revealing high levels of genetic variation and substantial genetic differentiation over small areas, help to demonstrate the inadequacy of these arguments. Nevertheless, patterns of genetic divergence may exist over small areas, independent of the size of the species ranges. Several case studies, including one on *Chioglossa lusitanica*, the Golden-striped salamander (Alexandrino *et al.*, 2000, 2002), illustrate this point.

*Chioglossa lusitanica* is an Iberian endemic amphibian with a small distribution and a distinct ecological niche. Its range is restricted to the north-western corner of the Iberian peninsula. The species inhabits the banks of swift running mountain streams in areas characterized by high rainfall, high topographical relief and low summer and winter temperatures (Arntzen, 1981; Teixeira *et*

*al.*, 2001; Teixeira & Ferrand, 2002). Morphological and ecological characteristics are shared with its sister-species, the salamandrid *Mertensiella caucasica* from the Caucasus and with some Nearctic plethodontid salamanders, revealing a remarkable convergent evolutionary adaptation (Arntzen, 1994). Geographic variation in *C. lusitanica* has been documented for several nuclear and cytoplasmic gene loci and coloration pattern. Variation at both nuclear and cytoplasmic gene loci suggests that the northern part of the distribution results from a post-glacial range expansion (Alexandrino *et al.*, 2000, 2002, and in Teixeira & Ferrand, 2002). Furthermore, the genetic data revealed significant substructuring, warranting the recognition of a southern form as different from a northern form. The forms may have been historically separated by the Mondego river valley (Alexandrino *et al.*, 2000, 2002; Teixeira & Arntzen, 2002). They appear to have recently resumed gene flow, at least in the western part of the Mondego area, where habitat conditions are more favourable to the species than the flat eastern Mondego range with an annual precipitation of less than 1000 mm.

The southern and northern form of *C. lusitanica* differ morphologically in the relative length of their limbs and digits. Similar morphological variation occurs gradually along a south to north axis across the species range, i.e. independent of taxonomy. The patterns of variation accumulate into a stepped cline of ‘short-legged’ salamanders in the south to relatively ‘long-legged’ salamanders in the north, with a ‘step’ coinciding with the Mondego valley (J. Alexandrino, unpublished data). The non-random signature of variation suggests that the morphological characters are under natural selection. We conclude that both forms are probably adapted to somewhat different environments. The agent of selection is unknown, but likely to be ecological-climatic.

*Correspondence:* J. W. Arntzen, National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands. *E-mail:* arntzen@nmm.nl

Here we model the distribution of two genetically differentiated forms of *Chioglossa lusitanica* with the help of GIS-modelling, adding a taxonomic component to spatial and spatio-temporal axes (Teixeira *et al.*, 2001; Teixeira & Arntzen, 2002).

#### MATERIAL AND METHODS

An extensive survey guided by the UTM-geographical system provides a sharp picture of the current range of *C. lusitanica* in Portugal with a  $10 \times 10$  km spatial resolution (Godinho *et al.*, 1999; Teixeira *et al.*, 2001). Data on species absence-presence cover 281 grid cells out of 967 grid cells countrywide (29%). Perceived absences cover 74 grid cells, mostly neighbouring the recorded range. Recorded presences cover 202 grid cells, plus five grid-cells covering less than 50% of continental Portuguese that were omitted from the analysis. At this scale of observation the distribution of *C. lusitanica* is continuous and the ranges of the northern and the southern forms are contiguous. The contact between them may coincide with the Mondego river (Ferrand de Almeida *et al.*, 2001). However, the observations that (1) population 5 (Várzeas), representing the southern form is situated north of the Mondego; and (2) population 6 (Buçaco) possesses both 'southern' and 'northern' cytoplasmic genetic markers, shows that the contact zone not fully and strictly coincides with the river valley (Fig. 1). A more conservative and cautious approach therefore is to recognize that the contact zone could be situated at any place in between the investigated populations 1-5 (representing the southern form) and populations 7-10 (representing the northern form). Models derived under these different considerations we refer to as model 1 and model 2, respectively.

To derive the models, the presence of one form is contrasted with the presence of the other form. Absence data are not taken into consideration and the analysis is restricted to the area covering the documented *C. lusitanica* range (Arntzen, 1999). Under model 1, data from three grid cells for which group allocation is ambiguous or mixed are omitted from the analysis, leaving 199 data points available for analysis (59 southern and 140 northern). In model 2, data from 39 grid cells for which group allocation is ambiguous are not taken into consideration, leaving 163 data points for model construction (54 southern and 109 northern).

Thirteen pre-selected environmental parameters are available as independent variables in logistic regression analysis, carried out with SPSS 10.0 for Windows (SPSS, 1999), with the distributional data as the dependent variable. Selection implied limited correlation between the environmental variables, with a Spearman correlation coefficient  $< 0.8$  and further as in Teixeira *et al.* (2001) and 'Atlas do Ambiente' (Direcção Geral do Ambiente, 1995): acidity of the soil (pH), altitude (m), frost days (annual number), frost months (annual number), hardness of subterranean water ( $\text{CaCO}_3$ ,  $\text{mg l}^{-1}$ ), humidity of the air (%), insolation (h), vegetation

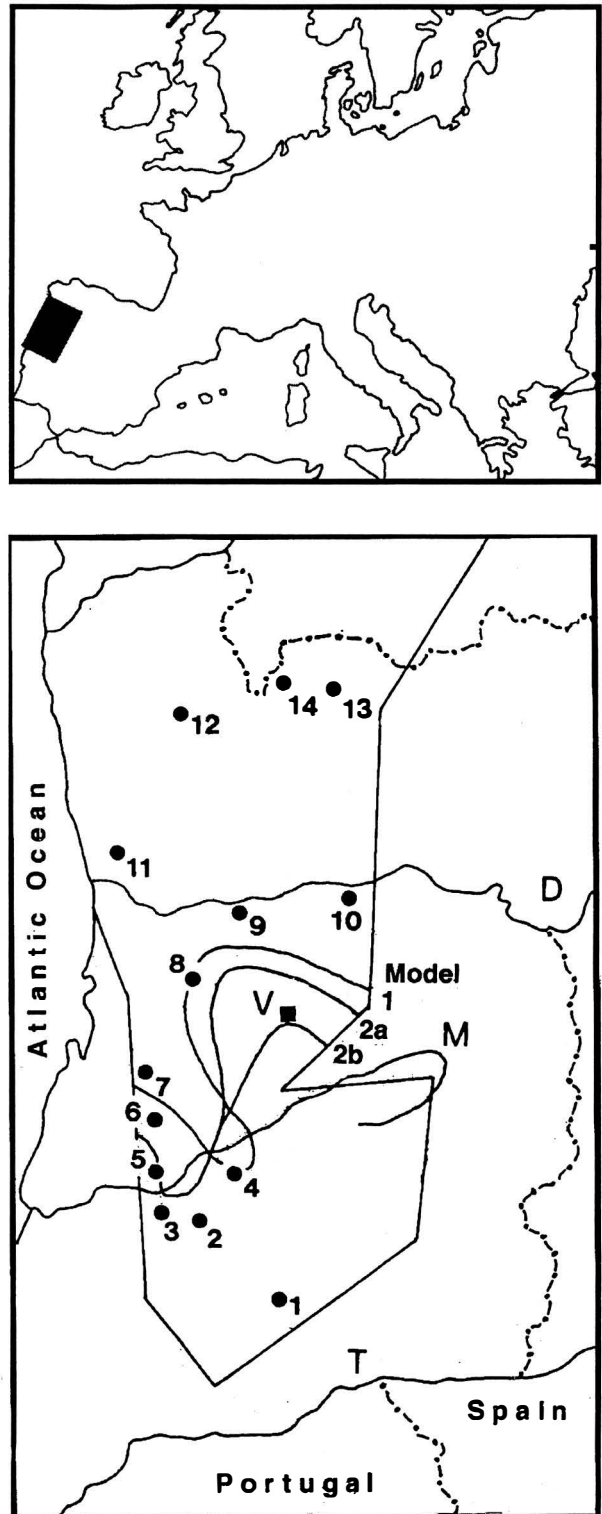


FIG. 1. Top, map of Europe with the area of interest shown in black. Bottom, the transition between northern and southern habitat types, following the GIS-models 1, 2a and 2b, are shown by curved lines. Note that models 1a and 1b are very similar in spatial terms and were merged. The range limit of *Chioglossa lusitanica* in northern Portugal is shown by straight lines (from Arntzen, 1999). Genetically studied populations are numbered 1 to 14 (from Alexandrino *et al.*, 2000). D = river Douro, M = river Mondego, T = river Tejo and V = Viseu.

index (NDVI), precipitation days (d yr<sup>-1</sup>), slope (%), annual mean temperature (°C) and July mean temperature (°C) as continuous variables and lithology as categorical variable (sedimentary and metamorphic versus igneous rock).

Variables were allocated to the explanatory models by a forward stepwise addition procedure, with the likelihood ratio as a selection criterion and alpha set at 0.05. Weights were applied to equilibrate the impact of differences in sample size between the northern and southern form. In the absence of significant imbalance between scores of correct and incorrect classification, the cut-off point was set at 0.5. Models were derived with and without Bonferroni correction (Holm, 1979). Models are presented on the basis of standardized continuous variables (with a mean of 0 and standard deviation of 1). The order of presentation reflects the input order of variables into the logistic equation. Model performance was evaluated with Cohen's kappa statistic. Kappa is a measure of the proportion of correct classification after accounting for chance effects. It provides a standardized, simple and effective statistic for evaluating or comparing presence-absence (and presence-presence) models (Manel *et al.*, 2001). The closer kappa is to 1.0 the higher is the accuracy of the data.

Values of over 0.8 are considered in medical applications to indicate a near perfect model performance (Landis & Koch, 1977). For visual representation, the variables composing the models are introduced into the Geographical Information System ILWIS 2.1 (ILWIS, 1997) as raster layers with a spatial resolution of 1 km.

## RESULTS

Model 1 includes three or six environmental variables. Under Bonferroni correction the probability of occurrence ( $P$ ) for the southern form versus the northern form is:

Model 1a –

$$P = (1 / (1 + \exp (-0.730 \times \text{humidity} - 0.272 \times \text{vegetation\_index} + 2.573 \times \text{lithology} + 85.249)))$$

Without Bonferroni correction:

Model 1b –

$$P = (1 / (1 + \exp (-1.158 \times \text{humidity} - 0.499 \times \text{vegetation\_index} + 2.485 \times \text{lithology} - 0.274 \times \text{precipitation} - 0.608 \times \text{July\_temperature} - 0.151 \times \text{frost\_months} + 224.394)))$$

TABLE 1. Environmental parameters selected in logistic regression modelling of the contiguous ranges of a southern and a northern form of the Golden-striped salamander, *Chioglossa lusitanica*, in Portugal, with the mean  $\pm$  standard deviation, or the number grid cells involved. The assumption that the contact zone between the forms follows the Mondego river (model 1) is relaxed under model 2 (details see text). † relative to countrywide average.

	under Bonferroni correction		without Bonferroni correction	
	southern form	northern form	southern form	northern form
<b>MODEL 1</b>				
Number of UTM-grid cells analysed	59	140	59	140
Allocated number of UTM-grid cells	69	133	73	129
Selected environmental parameters				
Frost months (number per year)			3.71 $\pm$ 1.443	2.93 $\pm$ 1.586
Humidity of the air (%)	71.1 $\pm$ 3.15	78.5 $\pm$ 3.03	71.5 $\pm$ 3.35	78.5 $\pm$ 3.23
Lithology (no. and frequency of grid cells)				
- sedimentary and metamorphic rock	57 (83%)	44 (33%)	56 (77%)	45 (35%)
- igneous rock	12 (17%)	89 (67%)	17 (23%)	84 (65%)
Vegetation index (NDVI) †	0.938 $\pm$ 0.0336	1.008 $\pm$ 0.0615	0.938 $\pm$ 0.0333	1.01 $\pm$ 0.0611
Annual total precipitation (mm/year)			1256 $\pm$ 269	1566 $\pm$ 334
Mean July temperature (°C)			17.9 $\pm$ 1.70	19.1 $\pm$ 1.02
<b>MODEL 2</b>				
Number of UTM-grid cells analysed	54	109	54	109
Allocated number of UTM-grid cells	87	115	81	121
Selected environmental parameters				
Altitude (m) †			1.44 $\pm$ 0.657	1.39 $\pm$ 0.832
Humidity of the air (%)	71.6 $\pm$ 3.01	79.3 $\pm$ 2.41	71.3 $\pm$ 2.99	79.1 $\pm$ 2.56

Kappa is 0.81 respectively 0.91. Model 2 includes one or two environmental variables. Under Bonferroni correction:

Model 2a –

$$P = (1 / (1 + \exp (-12.160 \times \text{humidity} - 3.298))).$$

Without Bonferroni correction:

Model 2b –

$$P = (1 / (1 + \exp (-18.080 \times \text{humidity} - 2.504 \times \text{altitude} - 5.011))).$$

Kappa is 0.95 respectively 0.99. The number of correctly and erroneously allocated data points across forms was not significantly out of equilibrium for either model ( $G$ -test of independence,  $0.322 < G < 2.625$ ,  $df=1$ ,  $P > 0.05$  in all four cases). The GIS spatial translation of the models is shown in Fig. 1. The area modelled as suitable for the southern form has, in climatic terms, low precipitation, low humidity, low summer temperatures and a high number of frost months, relative to the area modelled as suitable for the northern form. Also the southern area has a relatively low vegetation index, high altitude, and is mostly composed of sedimentary and metamorphic rock as opposed to igneous rock. The differential environmental characteristics of the southern and northern ranges are summarized in Table 1.

## DISCUSSION

The present analysis is different from other GIS-modelling exercises in that no 'false absences' (grid cells where the species occurs but has not been recorded) figure in the analysis. This is, firstly, because the distributional data set is near-complete at the  $10 \times 10$  km UTM grid scale, the spatial level chosen for analysis. Secondly, the analysis does not employ any absence data, but instead contrasts the presence of one form with the presence of another. An underlying assumption of the approach is that the ranges of the forms are contiguous and non-overlapping. The available data indicate this to be the case. GIS-modelling of the Portuguese portion of the range of 12 Iberian amphibian and reptile species shows that the descriptive fit of (absence-presence based) spatial models is related to accuracy in the correct prediction of ranges in adjacent areas (J. W. Arntzen, unpublished data). In other words, species range models that show tight fit for Portugal are likely to be good predictors for the range over the remainder of the Iberian Peninsula (i.e. Spain). Therewith, the high kappa values obtained for model 1 and model 2 descriptions of genetically differentiated forms within *C. lusitanica*, provide confidence that the selected environmental variables are meaningful descriptors of the *C. lusitanica* ranges and – directly or indirectly – relevant markers for the ecology of both form. It must be kept in mind, however, that the effect of the novel 'presence-presence' approach on kappa remains to be studied.

Considering the known ecological and habitat preferences of *C. lusitanica* as a species (Arntzen, 1981, 1999; Teixeira *et al.*, 2001), the parameters selected for the models can be interpreted as the southern group, stronger so than its northern counterpart, being a mountain form that is adapted to comparatively harsh environmental conditions. The possession of relatively short extremities is in line with a harsh climate (Nevo, 1972; Lee, 1993; J. Alexandrino, unpublished data). The ecological transition runs in the western part of the species range adjacent to the Mondego river. At the eastern side of the range, however, it runs up to 30–60 km north of the Mondego. The particular area around and to the south of Viseu belongs ecologically to the southern range (Fig. 1). This notion contrasts the range assumption underlying model 1, involving a contact between the forms along the Mondego river. Given the relatively harsh ecological conditions and the acknowledged effect of such an environment on the phenotype, within as well as between both forms, we predict that salamanders from this area will have shorter extremities than salamanders from populations at the same latitude at the western side of the range, or elsewhere in the northern range. Conversely, a historical biogeographical model reconstructed from genetic data and ecological modelling (Teixeira & Arntzen, 2002) predicts that the local populations belong to the northern form. This model postulates the separation of the ancestral *C. lusitanica* range along an east-west axis, coinciding with the Mondego river valley. Unfortunately, an adequate taxonomic description of both genetically differentiated units is not yet available and populations to the southeast of Viseu have not yet been studied. A complication is that the type locality of *C. lusitanica* (Buçaco, population 6) has, genetically as well as morphologically, mixed characteristics.

## ACKNOWLEDGEMENTS

We thank J. Brito and P. Sá Sousa for constructive comments.

## REFERENCES

- Alexandrino, J., Arntzen, J. W. & Ferrand, N. (2002). Nested clade analysis and the genetic evidence of population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Heredity* **88**, 66–74.
- Alexandrino, J., Froufe, E., Arntzen, J. W. & Ferrand, N. (2000). Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). *Molecular Ecology* **9**, 771–781.
- Arntzen, J. W. (1981). Ecological observations on *Chioglossa lusitanica* (Caudata, Salamandridae). *Amphibia-Reptilia* **1**, 187–203.
- Arntzen, J. W. (1994). Allometry and autotomy of the tail in the Golden-striped salamander, *Chioglossa lusitanica*. *Amphibia-Reptilia* **15**, 267–274.

- Arntzen, J. W. (1999). *Chioglossa lusitanica* Bocage 1864 - der Goldstreifensalamander. In: *Handbuch der Reptilien und Amphibien Europas*, 301-321, Grossenbacher K. & Thiesmeier B. (eds). Akademische Verlagsgesellschaft, Wiesbaden, Germany.
- Direcção Geral do Ambiente (1995) *Atlas do Ambiente – Versão 1.0*. Direcção Geral do Ambiente, Lisboa, Portugal.
- Ferrand de Almeida, N., Ferrand de Almeida, P., Gonçalves, H., Teixeira, J., Sequeira, F. & Ferrand de Almeida, F. (2001). *Anfíbios e Répteis de Portugal*. Guias Fapas, Porto.
- Godinho, R., Teixeira, J., Rebelo, R., Segurado, P., Loureiro, A., Álvares, F., Gomes, N., Cardoso, P., Camilo-Alves, C. & Brito, J. C. (1999). Atlas of the continental Portuguese herpetofauna: an assemblage of published and new data. *Revista Española de Herpetología* **13**, 61-82.
- Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.
- Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**, 87-112.
- Holm, S. (1979). A simple sequential rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65-70.
- ILWIS (1997). *ILWIS 2.1 for Windows*. International Institute for Aerospace Survey & Earth Sciences. Enschede, The Netherlands.
- Landis, J. R. & Koch, G. G. (1977). The measurements of observer agreement for categorical data. *Biometrics* **33**, 159-174.
- Lee, J. C. (1993). Geographic variation in size and shape of neotropical frogs: a precipitation gradient analysis. *Occasional Papers Museum Natural History University Kansas* **163**, 1-20.
- Manel, S. H., Williams, C. & Ormerod, S. J. (2001). Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**, 921-931.
- Nevo, E. (1972). Climatic adaptation in size of the green-toad, *Bufo viridis*. *Israel Journal of Medical Science* **54**, 1010.
- SPSS (1999). *SPSS for Windows*. Release 10.0. SPSS Inc., Chicago
- Teixeira, J. & Arntzen, J. W. (2002). Potential impact of climate warming on the distribution of the Golden-striped salamander, *Chioglossa lusitanica*, on the Iberian Peninsula. *Biodiversity and Conservation* **11**, 2167-2176.
- Teixeira, J. & Ferrand, N. (2002). The application of distribution models and Geographic Information Systems for the study of biogeography and conservation of herpetofauna: *Chioglossa lusitanica* as a case study. *Revista Española de Herpetología*, Special volume **2002**, 119-130.
- Teixeira, J., Ferrand, N. & Arntzen, J. W. (2001). Biogeography of the golden-striped salamander, *Chioglossa lusitanica*: a field survey and spatial modelling approach. *Ecography* **24**, 618-624.

Accepted: 23.6.03