European Journal of Soil Biology 75 (2016) 115–122

Contents lists available at ScienceDirect

European Journal of Soil Biology

journal homepage: http://www.elsevier.com/locate/ejsobi

Original article

Macroecological inferences on soil fauna through comparative niche modeling: The case of Hormogastridae (Annelida, Oligochaeta)

Daniel F. Marchán [a, *, 1], Pablo Refoyo [a, 1], Rosa Fernández [b], Marta Novo [c], Irene de Sosa [a], Darío J. Díaz Cosín [a]

[a] Departamento de Zoología y Antropología Física, Facultad de Biología, Universidad Complutense de Madrid, C/José Antonio Nováis 2, 28040 Madrid, Spain
[b] Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA
[c] Environmental Toxicology and Biology, Facultad de Ciencias UNED, C/Senda del Rey 9, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 25 January 2016
Received in revised form
12 May 2016
Accepted 17 May 2016
Available online 25 May 2016

Keywords:
Ecological niche models
ENMtools
Macroecological inferences
Soil fauna
Hormogastridae

ABSTRACT

Ecological Niche Modeling (ENM) through MaxEnt and quantitative comparison techniques using ENMtools could facilitate ecological inferences in problematic soil dwelling taxa. Despite its ecological relevance in the Western Mediterranean basin, the ecology of the endemic family Hormogastridae (Annelida, Oligochaeta) is poorly known. Applying this comparative approach to the main clades of Hormogastridae would allow a better understanding of their ecological preferences and differences. One hundred twenty-four occurrence data belonging to four clades within this earthworm family were used as input to infer separate MaxEnt models, including seven predictor variables. Niche breadth, niche overlap and identity tests were calculated in ENMtools; a spatial Principal Components Analysis (sPCA) was performed to contrast with the realized niches. The highly suitable predicted ranges varied in their ability to reflect the known distribution of the clades. The different analyses pointed towards different ecological preferences and significant ecological divergence in the four above-mentioned clades. These results are an example of wide-scale ecological inferences for soil fauna made possible by this promising methodology, and show how ecological characterization of relevant taxonomic units could be a useful support for systematic revisions.

1. Introduction

Macroecological studies comparing the ecological preferences of different soil taxa are almost absent from the literature (but see Ref. [1]): this is not necessarily a case of a lack of interest for this community, but most probably related to the difficulty of their study [2].

One approach which has facilitated the ecological inferences in these problematic groups is Ecological Niche Modeling (ENM), with MaxEnt [3] standing out among the different methodologies due to its high performance when including presence data only. It has been implemented in several groups (ground beetles [4], termites [5], millipedes [6]) including earthworms [7]: used MaxEnt to study the effect of large-scale ecological variables in the distribution of *Hormogaster elisae*, corroborating its high predictive power and the ability to reflect accurately its soil preferences.

Additionally, the implementation of several existing indices and statistical tests in the software ENMtools [8] has allowed the quantitative comparison of Ecological Niche Models (ENMs) between related species, including niche overlap, niche breadth and testing for statistically significant differentiation. Some recent studies have proven the usefulness of these methodologies to answer diverse biological questions, applying them to different animal groups. For example [9], found ecological niche differentiation in two cryptic beetle species, using the fact as support for their status as valid species. Ref. [10] studied niche overlap and niche breadth in three cryptic bat species complexes, as part of their research on how environmental factors and ecological interactions influenced their speciation. However, these promising

Abbreviations: ENM, Ecological Niche Modeling; sPCA, Spatial Principal Component Analysis; ROC-AUC, Receiver Operating Characteristic-Area Under the Curve, AUC in shorter form; TRANGE, Mean Diurnal Range; ISOTHER, Isothermality; TCOLD, Mean Temperature of the Coldest Quarter; PRDRY, Precipitation of Driest Month; LITHO, Lithology; VEGET, Land cover; ANTHRO, Human influence; HGI, Highest gain in isolation; HDGO, Highest decrease in gain when omitted.

* Corresponding author.
E-mail address: danifermch@gmail.com (D.F. Marchán).
[1] Daniel F. Marchán and Pablo Refoyo contributed equally.

http://dx.doi.org/10.1016/j.ejsobi.2016.05.003
1164-5563/

and insightful methodologies have been scarcely used to address similar questions in soil animals (and especially earthworms).

Several authors have justified the utility of both modeling distributional patterns and ecological niche characteristics above the species level (e.g., [11–13]) when individual species have scarce occurrence data.

This approach appears promising for a relatively obscure earthworm family, the Hormogastridae Michaelsen 1900. It is the second most diverse earthworm family in the western Paleartic region, after the Lumbricidae Claus 1876. They play an important ecological role as deep-burrowing endogeics in the western Mediterranean basin [14], processing great amounts of soils expelled as casts [15]. In some places, such as Sardinia, they were shown to be dominant in abundance in earthworm communities [16], being adapted to drought and impoverished soils [17,18]. To understand better the role of hormogastrids in soil ecosystems, it is necessary to comprehend their ecological preferences, adaptations and response to environmental variables. However, those have been scarcely studied, mainly focusing on one particular species: *Hormogaster elisae* Álvarez 1977. The larger body size and associated slower reproductive rate [19] of most species in the family, together with their scattered distribution and difficult capture (their deep burrowing requires intensive digging efforts), have discouraged their laboratory and field research. The only work on their ecological preferences [20] corroborated the presence of most hormogastrids in soils with low nutrients content, with a preference for soils more basic and fine-textured than the ones observed for *H. elisae*.

The phylogeny of Hormogastridae has been clarified in the last years using molecular evidence: after [21–23] it has been divided in 9 clades. *Xana*, *Vignysa*, *Hemigastrodrilus* and *Ailoscolex* had already been described as independent genus, but the other five remain artificially joined in the catch-all genus *Hormogaster*. Four of these latter clades (informally termed Central Iberian, Northeastern Iberian, Tyrrhenian, and Disjunct) have wide geographic ranges and a high number of known populations, which makes them suitable for an in-depth macroecological study. Understanding their ecological preferences and the differences in their ecological niches would be helpful as an additional support for their future definition as new genera in the taxonomic revision of the family.

The main aim of this work is to obtain macroecological inferences for the four main clades of Hormogaster through comparative niche modeling as an example of the potential of this methodology. We used all the available geo-referenced presence locations to obtain the ENMs for each of the clades, and ENMtools and niche space visualization to perform quantitative comparisons. Our objectives were: i) to predict the distribution of the main clades of Hormogaster in their home range; ii) to find the environmental variables with a stronger influence in their distribution; iii) to study the overlap, breadth and statistical differentiation of their ecological niches; and iv) to compare their potential and realized niches.

This study could be potentially useful as a base for wide-scale ecological inferences in other groups of earthworms and soil fauna, a key element on most ecosystems around the world.

2. Materials and methods

2.1. Training data

One hundred twenty-four presence localities were used to train the models (detailed in Suppl. Material 1): 44 for the Central Iberia clade, 29 for the Tyrrhenian clade, 30 for the Northeastern Iberia clade and 21 for the Disjunct clade. All the presence data were obtained in sampling campaigns by the authors, ensuring high reliability; they also constitute a good representation of the known ranges of the species, defined after more than a century of field works by other researchers in the Mediterranean.

The Central Iberia clade corresponds to the *H. elisae* morphospecies, which comprises at least five cryptic lineages [24]. [7] included several new populations, which considerably widened its known range; their genetic variability is currently being researched.

The Northeastern Iberia clade comprises a high number of closely related species (see Suppl. Material 1) with high morphological variability in their diagnostic characters [20]. They inhabit Northeastern Spain and a small region of Southeastern France, with most of their diversity located in Catalonia (Spain).

The Disjunct clade includes the Sardinian populations of *Hormogaster pretiosa* Michaelsen 1889 (a taxonomically problematic species into which other unrelated species were wrongfully assigned [21]) confined to the southwestern part of the island, *Hormogaster najaformis* Qiu & Bouché 1998 and *H regina* Rota 2016 from Catalonia (Spain) and an assembly of undescribed related forms in the latter region.

The Tyrrhenian clade includes *Hormogaster redii* Rosa, 1887, *Hormogaster samnitica* Cognetti, 1914 and their subspecies [25] found deep genetic divergence pointing to them being composed of cryptic lineages. They are distributed around the Tyrrhenian Sea, occupying most of Sardinia, Corsica, Tuscan Archipelago, the Tyrrhenian side of Italy from Tuscany to Naples, Sicily and a small area between northern Algeria and Tunisia [16].

2.2. Environmental variables

The large-scale variables potentially relevant for the biology of Hormogastridae were chosen as predictor variables to model its distribution, as described below.

Four bioclimatic variables were selected from Worldclim (http://www.worldclim.org/ accessed 12/05/2016): Mean Diurnal Range-BIO2 (**TRANGE**) and Isothermality-BIO 3 (**ISOTHER**) are suitable to represent the influence of extreme temperature variation (both daily and across the year) on earthworm distribution in the Mediterranean region. Mean Temperature of the Coldest Quarter-BIO11 (**TCOLD**) was chosen to reflect the impact of soil freezing and low temperatures on the activity of earthworms. Precipitation of Driest Month-BIO 14 (**PRDRY**) is likely to reflect the availability of water in the soil across the year (an essential requirement for earthworms) and the severity of drought periods.

As topographical variable we selected Lithology -PAR-MAT-DOM2, Second level code for the dominant parent material of the STU from the European Soil Database Raster Library 1 km × 1 km (http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDB_data_1k_raster_intro/ESDB_1k_raster_data_intro.html accessed 12/05/2016)- (**LITHO**). Lithology is likely to influence indirectly Hormogastridae ecology through a wide range of correlated variables, including the structure and biochemical characteristics of soils.

CORINE 2006 Land Cover (version 12/2009: http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-clc2006-100-m-version-12-2009 accessed 12/05/2016) – (**VEGET**) was chosen to incorporate information about vegetation and land use, which are widely known to influence earthworm distribution (e.g. [26]).

As large, deep burrowing species as most Hormogastridae are affected negatively by human disturbance [27], the 'Human footprint' data set -representing the human influence on land surface [28] – (**ANTHRO**) was selected to include the effect of anthropic activities on habitat suitability.

These variables were the same that successfully predicted the distribution of *H. elisae* (and relevant soil characteristics) in Ref. [7]. Precipitation of the coldest quarter was replaced by Mean