UNOBSERVED DIVERSITY IN DARWIN'S APPRAISAL OF THE AZORES

José A.P. Marcelino [1], Rosanna Giordano [2], Felipe Soto-Adames [2], Patrícia V. Garcia [1], Everett Weber [3], Luís Silva [1], António O. Soares [1]

[1] *University of the Azores, Department of Biology, Ponta Delgada, Azores, Portugal*
e-mail: jmarcelino@uac.pt
[2] *University of Illinois, Illinois Natural History Survey, Champaign, IL, USA*
[3] *Murray State University, Department of Biology, Murray, KY, USA*

ABSTRACT

Sampling data for Collembola (Arthropoda: Hexapoda) bioindicator species in thirteen ecosystem types, representing a gradient of anthropogenic influence on four islands of the Azorean archipelago and Madeira Island were analyzed using rarefaction curves to determine species richness per habitat type. Rarefaction curves show that despite the high degree of sampling we have yet to reach a plateau for species richness. Nestedness temperature of presence-absence matrices (T=9.6) indicates a non-random distribution of species along the gradient of habitats. Non-metric Multi-Dimensional Scaling (NMDS) for Arthropoda community composition across the gradient reveals a clear differential abundance of species between natural and artificial habitat types with a MDS Fit Plot for the 3 Dimensional solution showing an overall Goodness of fit of *ca.* 99%. The marked differential species assemblages might correlate with high nestedness in Azorean Arthropoda communities. Preliminary complete COI mitochondrial DNA sequences were used to evaluate the phylogeographic structure within and among population assemblages of Collembola species (N=50). Maximum genetic distances were observed between artificial production forest (i.e., *Cryptomeria japonica*), natural habitats (i.e., *Laurus azoricus*) and Madeira Island. Darwin may have thought that, in comparison with the Galapagos and South American flora and fauna, the Azores has not "much worth seeing"; however, our study indicates that this archipelago has high species diversity despite long-standing anthropogenic impacts.

RESUMO

Amostragens intensivas com o intuito de obter uma representação significativa de espécies bioindicadoras de Collembola (Arthropoda:

Hexapoda), foram efectuadas em treze ecossistemas dos Açores, representado estes um gradiente crescente de influência antropogénica em quarto ilhas do Arquipélago dos Açores e da Madeira. Os dados foram analisados usando curvas de rarefacção de espécies com o intuito de determinar a riqueza por cada tipo de habitat. As curvas de rarefacção indicam que apesar do elevado grau de amostragem ainda não se atingiu o topo de riqueza de espécies para os habitats dos Açores. Temperatura de Nestedness para matrizes de presença-ausência de espécies (T=9.6) indicam uma distribuição não aleatória das espécies ao longo do gradiente. Non-metric Multi-Dimensional Scaling (NMDS) para a composição das comunidades de Artrópodes através do gradiente, revelam uma clara diferenciação na abundância de espécies entre habitats tipo naturais e artificiais, com uma MDS FIT Plot fitness de aprox. 99%. A marcada diferenciação na composição de espécies pode estar correlacionada com a alta Nestedness encontrada nas comunidades de Artrópodes dos Açores. Dados preliminares de sequencias de ADN do gene mitocondrial COI (completo) foram usadas para avaliar a filogeografia entre populações de Collembola (N=50) e no interior das mesmas. Distâncias genéticas máximas foram observadas entre floresta artificial de produção (i.e., *Cryptomeria japonica*), habitats naturais (i.e., *Laurus azoricus*) e a ilha da Madeira. Darwin pode ter pensado que, em comparação com a flora e fauna das Galápagos e a América do Sul, os Açores não "têm muito meritório de se ver"; no entanto, o nosso estudo indica que os Açores têm uma alta diversidade de espécies apesar dos duradouros impactes antropogénicos no arquipélago.

INTRODUCTION

The Azorean islands (36° 35'-39° 43' N, 24° 45'-31° 17' W), located over 615 km across the Mid-Atlantic Ridge, are recognized worldwide as important conservation hotspots (UNESCO World Heritage spots, Biosphere and Natura 2000 Networks). The increase of agriculturally fragmented and monocultural landscapes, as a result of a growing domestic demand for arable crop land, has drastically altered the biodiversity (richness and abundance) of flora and fauna in many areas of the archipelago. Modifications in biotic food webs occur after habitat disturbance with long-term negative ecological and economic impacts being

consistently reported (Thies *et al.*, 2008; Tylianakis, 2008).

The critical importance of preserving biodiversity is widely recognized by groups within the Convention for Biological Diversity, the G8 and the European Environmental Agency (European Environment Agency, 2007). Economical gains from services provided by ecosystems, have been recently estimated as worth 19 trillion € (Costanza *et al.*, 2007) whilst the costs associated with biodiversity losses are estimated as 1.35 and 3.1 trillion € (Sukhdev, 2008). The groups mentioned above are active and politically committed to protect biodiversity hotspots, aiming to reduce the loss of biodiversity by 2010, the International Year for Biodiversity, and achieve environmental sustainability by 2015.

Invertebrates are used as bioindicators in the most advanced protocols for the analysis of biodiversity and ecosystem stability (Basset *et al.*, 2004; Pearsall, 2007). Invertebrate bioindicators are unique due to their high sensitivity and rapid response to changes in their habitat (Hodkinson & Jackson, 2005). Hence, they can accurately portray cryptic changes in the dynamics of ecosystems.

The present research endeavors to contribute to the current international directives aiming to safeguard the biological resources of the globe, with particular relevance to the Azorean Archipelago. Remote archipelagos enclose high endemism levels and a significant portion of terrestrial biodiversity (Holland & Hadfield, 2002; Martín *et al.*, 2008). We intend to detail the Collembola invertebrate communities on 13 representative habitat types distributed in four representative islands of the archipelago of the Azores (scaling up from pristine to increasing anthropogenic influenced habitats) and in Madeira Island (Madeira Archipelago). Genetic characterization and community phylogeographic structure analyses will be complemented by ecological studies in order to find population interconnectivity rates and diversity within and among island(s) and, if possible, population source(s).

The class Collembola (Arthropoda: Hexapoda) was selected for the molecular analysis in this study, as a model bioindicator organism for the Azorean Archipelago, due to two main factors: (i), their richness and abundance in a wide variety of ecosystems and (ii), their rapid response to changes in ecological and pedological patterns within a given ecosystem.

In addition, we correlate arthropod community assemblages' richness (β and α diversity), for the orders sorted so far, i.e., Hymenoptera, Diptera, Collembola and Araneae, and the degree of disturbance throughout an increasing gradient of anthropogenic influence in the ecosystems of the Azores.

Our ultimate goal is to combine genetic with ecological data, in order to define critical conservation hotspots and to make recommendations that will reduce the disturbance of anthropogenic activities on the Azorean Archipelago.

METHODS

Habitat selection

Habitats were selected to represent an increasing gradient of anthropogenic influence in the chosen ecosystems. A total of 13 major habitat types were chosen with the aim to depict the overall landscape diversity of the Azorean Archipelago (Table 1). Monitored islands were selected on the basis of the relative proportion of land allocated to agriculture, pristine and reservoir habitats, i.e.:

● São Miguel - SM (high level of agricultural development *vs.* low level of pristine habitat); special attention was given to the unique microclimate of the three main calderas on São Miguel Island.
● Terceira - TR (high level of agricultural development, high level of pristine habitat)
● Pico - PI (low agricultural development *vs.* high level of pristine habitat - World Heritage spot)
● Santa Maria - SMR (low agricultural development, low level of pristine habitat with relevant uniqueness value, i.e., oldest island in the archipelago).

In addition to the sampled areas herein, the Island of Madeira (MD) in the Madeira Archipelago was also sampled as as an additional area outside of the Azores archipelago.

Sampling & Molecular Analysis

To obtain representative samples of collembolan richness and faunal composition for a given habitat type, soil and litter fauna were obtained from June to December 2008. Sampling method and device were adjusted to sample type, i.e., epigeic Collembola (pitfall traps) and litter Collembola (Berlese-Tullgren traps).

TABLE 1. Habitat types sampled depicting an increasing level of anthropogenic impact. MD, Madeira Island; MID, Middle; RIM STL, Rim Santiago Lagoon; SH, Shore; SH ST, Shore Stream; SH W, Shore Woods; SM, São Miguel Island; SMR, Santa Maria Island; TOP-AD-HS, Top adjacent Hotspring; TOP-HS, Top Hotspring; TR, Terceira Island.

Habitat type	Land cover		Island sample
Pristine forest (Indigenous)	*Laurus azorica, Myrica faya, Frangula azorica, Erica azorica, Ilex perado, Juniperus brevifolia*		SM, SMR, TR, MD
Mixed woodlands	*Acacia melanoxylon, Persea indica, Lantana camara, Tradescantia* spp.		SM, SMR
Caldera #1 (Lagoa do Fogo)	RIM	*Hedychium gardneranum*	SM
	MID	*Pteridium aquilinum, Rubus ulmifolius, Argeratina adenophora*	
	SH	*Colocasia esculenta, Sphaeropteris cooperi*	
Caldera #2 (Lagoa das Furnas)	TOPHS	*Hedychium gardneranum*	SM
	TOP AD HS	*Pteridium aquilinum, Rubus ulmifolius, Argeratina adenophora*	
	MID	*Cyperus longus*	
	SH ST	*Colocasia esculenta, Sphaeropteris cooperi*	
	SH W	*Colocasia esculenta, Sphaeropteris cooperi*	
Invasive plant scrubs	*Pittosporum undulatum*		SM
Production forest #1 (Introduced)	PN	*Pinus pinaster*	SM
Production forest #2 (Introduced)	EU	*Eucalyptus globulus*	SM
Production forest #3 (Introduced)	CR	*Cryptomeria japonica*	SM, TR, SMR, PIC
Semi natural pasture (old sown areas)	*Sporobolus indicus, Briza minor, Lotus subbiflorus, Daucus carota, Sherardia arvensis, Anagallis arvensis*		TR
Artificial pasture (current grazing)	*Lolium perenne, Bromus willdenowii, Trifolium repens, Poa* spp., *Holcus lanatus, Rumex obtusifolius, Plantago lanceolata, Dactylis glomerata, Sporobolus indicus*		SM, TR, PIC
Cereals	*Zea mays*		SM
Agro-cultures	*Beta vulgaris*		SM
Urban areas (garden)	*Hydrangea macrophylla, Robinia pseudoacacia*		SM
Vineyard	*Vitis vinifera*		MD

Increasing Anthropogenic Impact

Pitfall traps (PF) consisted of plastic cylinder cups 78 mm deep and 42 mm in diameter filled with *ca*. 80 ml of a mixture of 96% alcohol and 0.05% liquid detergent in order to maintain insect specimens viable for genetic work and to hinder surface tension, respectively. Each trap was placed in the soil so that the lid was flush with the surface. The traps were covered with a plastic plate at *ca*. 3 cm height to avoid desiccation, flooding and insectivore predation. Traps remained in the soil for 7 days prior to re-collection.

Berlese-Tullgreen traps (BT) consisted of two plastic darkened containers assembled together to provide an upper vented area (14 cm diameter x 11.5 cm high) with 4 openings (1 cm diameter covered with a 0.3 x 0.3 mm diameter mesh) and coupled with a 15 W lamp on top, and a lower collecting area (13 cm diameter x 10 cm high) partially filled with *ca*. 80 ml of the same mixture used in the pitfall traps. Litter samples were placed on a 1.8 x 1.8 mm mesh attached to a plastic funnel positioned in the assembling zone between the two halves of the device. In order to avoid heat and dryness, collembola crawl downward in the littler sample and drop through the funnel into the collecting mixture. Litter samples remained for 72 h in BT before processing.

Thirteen pitfall trap replicates (*ca.* 25 m apart), per habitat type, were set. Litter samples were collected, when present, adjacent to pitfall traps and ranged from 6 to 13 for a given habitat type. Additional sampling was made for the unique microclimate of the catchments areas (calderas). In these, three sampling sites were established, namely, at the base, middle and top of the lagoons. All sampling sites were geo-referenced.

Specimens were sorted under a stereomicroscope and assigned to morphospecies groups (i.e., Entomobryidae, Sminthuridae and Poduridae). A reference specimen collection was generated for each habitat type and island surveyed, at the Entomology Laboratory, Department of Biology of the University of the Azores, Portugal. All specimens were preserved in 96% ethanol and allocated to a storage chamber at 4 °C, prior to higher taxonomic identifications at the Institute of Natural Resources Sustainability, Biodiversity Section, University of Illinois, EUA. Genetic extraction was conducted following Marcelino *et al*. (2008) and molecular analyses using PAUP software (Swofford 2001). Sixteen primers where designed to sequence the complete COI gene for all target species represented in the majority of the islands sampled.

Following the initial sampling described herein more transects and islands were sampled to increase data robustness on insect species richness and distribution. Two replicates per habitat type and island were made (i.e., 2 replicates x 8 habitat types x 5 islands) for a total of 80 sampling sites and four collection methods representing all

strata present in a given habitat, i.e., Pitfall trapping (epigean fauna); Berlese-Tullgren trapping (micro epigean fauna); Vacuum (fauna with aerial vagility) and sweeping (canopy fauna). Data presented herein for Collembola community assemblages analysis concern the specimens found in Berlese traps for São Miguel island. Data for non-metric Multi-Dimensional Scaling (MDS) for community composition across the ecosystems' gradient concerns arthropod communities of Collembola, Hymenoptera, Diptera and Araneae.

Statistical analysis

We calculated the Nestedness temperature of presence-absence matrices using BINMATNEST software (Rodríguez-Gironés & Santamaría, 2006) as a measure of the degree of correlation (Nestedness) between a community assemblage and a given habitat. Temperature nestedness range from 0° to 100°. The lower the temperature the higher the nestedness.

Rarefaction curves, which give an estimate of the richness of species yet to be sampled, as a function of the number of individuals sampled, where also plotted. Higher slopes indicate that a large fraction of species

is yet to be sampled (Gotelli & Colwell, 2001). Rarefaction curves where computed using EstimateS® 8.0 (Colwell, 2006).

Multi-dimensional scaling (MDS) displaying the similarity between habitats sampled for species composition (calculated using SAS®, Statistical Analysis System) are also depicted. Points are arranged in space so that the distance between pairs of habitats reflects their similarity in terms of species community assemblages (Borg & Groenen, 2005).

RESULTS & DISCUSSION

A preliminary molecular analysis using 1300bp of Cytochrome oxidase I (COI) shows the results from a subset of Collembola species sampled. *Heteromorus major* was sampled from SMG (N=3) and TR (N=2). Uncorrected pairwise p distances between specimens from the two islands were 10%. Collembola usually show a greater than 10% divergence between different species, and species differing approximately between 10 to 12% may not show clear signs of morphological differences. However, using these measures and the general mitochondrial calibration of 2% per

Million years (Papadopoulou *et al.*, 2010), these populations have been separated for 5 Million years. Thus, we conclude that this species is likely to be endemic to the Azores. Subsequent pitfall and Berlese funnel trapping showed that this species is restricted to pasture on São Miguel Island (data from Terceira Island has not yet been analyzed). The molecular evidence would seem to indicate that this species is a native that has lost habitat or was able to take advantage of the change in habitat resulting from the expansion of pasture landscapes.

Ceratophysella engandinensis sampled from SMG (N=6), SMR (N=2) and MD (N=2) showed the presence of two haplotypes differing by 19%, suggesting that they probably are cryptic species. Both haplotypes were found on SMG while the haplotypes in MD and SMR differed. However, more sampling needs to be done to determine if this pattern is maintained. The differences described above may indicate that this species is endemic. Two species (*Pogonognathelus longicornis* and *Lepidocyrtus curvicollis*) have molecular and distribution patterns that would tend to indicate that they are in-

vasive species. *Pogonognathelus longicornis* (N=7) was sampled on the island of SMG from 6 far distant localities. All the 7 analyzed specimens had the same haplotype. Three individuals of *Lepidocyrtus curvicollis* were sampled on SMG from 5 localities and these were also identical. In subsequent field sampling, both species were found on five islands and 8 habitat types. The similarity of these organisms determined through the molecular data combined with the ubiquitous nature of this species seems to indicate the rapid distribution characteristic of an invasive species; however, there were few sites where this species clearly comprised more than 50% of the total Collembola species present. Thus, although these species may quickly disperse throughout the region, they do not seem to be disrupting the existing species.

We have sequenced 25 individuals of *Onychiurus* sp. from SMR (N=8); SMG (N=12) and MAD (N=5) and found that there are three cryptic species with a range of p values from 14 to 19%. One of the species is found both on SMR and SMG, a second is found on both SMG and MD while the third species is

found only on SMR. Subsequent sampling found the species on three islands and in seven habitats (data not shown). There are twenty-two locations where this species has fewer than six individuals in a site and most of those are singletons or doubletons, and three locations where there were a large number of individuals, two of which were inclusively dominated by this species (a seminatural pasture and a crop site).

These preliminary results using COI indicate that within this group we were able to identify both endemic and newly arrived or invasive species, making Collembola a good bioindicator tool to assess the degree of disturbance in a given ecosystem.

Rarefaction curves of the Collembola species collected and identified so far (Figure 1), indicate that the natural sites are likely to be the most diverse with respect to Collembola species, and pasture and *Cryptomeria* sites will likely be the least diverse. It is also clear that a lot of sampling is still required particularly for the natural pristine sites as evidenced by the steep slope on the rarefaction curve. Pasture and *Cryptomeria* production forests had both relatively
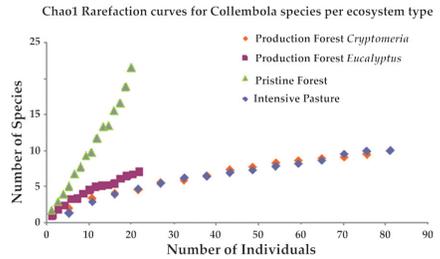

FIGURE 1. Chao1 Rarefaction curves (calculated using EstimateS® 8.0, Colwell, 2006) of Collembola species abundance collected through Berlese funnels in four habitats on São Miguel Island. The curves have been corrected for the number of individuals from each sample.

low diversities. None of the species richness rarefaction lines became asymptotic, indicating that more sampling is required to capture all of the diversity at the research sites.

Non-metric Multi-Dimensional Scaling (NMDS) for community composition across the gradient (Figure 2) reveals a clear separation of species richness between natural and artificial habitat types with a NMDS Fit Plot for the 3 Dimensional solution showing an overall Goodness of fit of *ca.* 99%.

Results indicate that Arthropod species richness and the uniqueness of species present in Pristine Forests is very high when contrasted with artificial,
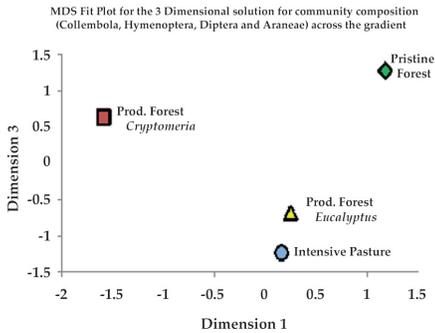
MDS Fit Plot for the 3 Dimensional solution for community composition
(Collembola, Hymenoptera, Diptera and Araneae) across the gradient

FIGURE 2. Multi-dimensional scaling displaying the structure of distance-like data as a geometrical picture (calculated using EstimateS® and SAS®, Statistical Analysis System). Similar habitats are represented by two points that are close together, and two dissimilar habitats are represented by two points that are far apart.

human generated habitats, i.e., Production forests and Intensive Pastures. Nutrient-enriched ecosystems (i.e., intensive pastures) showing higher arthropod abundance but low richness, represent depauperate communities dominated by the most tolerant and competitive species (i.e., invasive species), hence, altering both species richness and evenness of community assemblages.

A nestedness of T=9.6° was obtained for the Arthropod community assemblages and habitats analyzed so far, hence, indicating a non-random distribu-

tion of species along the gradient of habitats. Such species assemblage corroborates the theories of island biogeography (Losos & Ricklefs., 2010), where a high correlation between species and habitats is usually found.

In spite of the intensive sampling we have still not yet reached a plateau for the number of species putatively present in the robust sample of habitats representing major ecosystem types of the archipelago of the Azores. The data presented give an indication of the current species richness in the archipelago and its relevance within Azorean Pristine ecosystems. This project is a multi-year effort and we will continue to collect data in order to capture a larger proportion of the species present at each site, use ecological analysis of species interactions to indicate ecosystem disturbance and expand the use of molecular data to help identify dispersal rates.

ACKNOWLEDGEMENTS

LITERATURE CITED

BASSET, Y., J.F. MAVOUNGOU, J.B. MIKISSA, O. MISSA, S.E.MILLER, R.L. KITCHING & A. ALONSO, 2004. Discriminatory power of different arthropod data sets for the biological monitoring of anthropogenic disturbance in tropical forests. *Biodiversity and Conservation*, 13: 709-732

BORG, I., & P. GROENEN, 2005. *Modern Multidimensional Scaling: theory and applications* (2nd ed.). Springer-Verlag, New York.

COLWELL, R.K., 2006. *EstimateS*: *Statistical estimation of species richness and shared species from samples*. Version 8.

COSTANZA, R., N.E. GOLUBIEWSKI & C.J. CLEVELAND, 2007. Value of the world's ecosystem services: the influence of a single paper. *In*: CUTLER, J., & C.J. CLEVELAND (eds.), *Encyclopedia of Earth.* (Washington, D.C.: Environmental Information Coalition, National Council for Science and the Environment).

DIAS, E., D. PEREIRA, V. MEDEIROS, J. MENDES & R.B. ELIAS, 2007. Distribuição das principais manchas florestais. Açores. *In:* SILVA, J.S. (ed.), *Árvores e florestas de Portugal. Açores e Madeira. A floresta das ilhas*, pp. 299-322. Público - Comunicação Social, SA e Fundação Luso-Americana para o Desenvolvimento.

EUROPEAN ENVIRONMENT AGENCY, 2007. *Halting the loss of biodiversity by 2010: proposal for a first set of indicators to monitor progress in Europe*, 182 pp. Technical report Nr. 11/2007. Copenhagen, Denmark.

GOTELLI, N., & R. COLWELL, 2001. Quantifying biodiversity: procedures and pitfalls in measurement and comparison of species richness. *Ecology Letters*, 4: 379-391.

HODKINSON, I.D., & J.K. JACKSON, 2005. Terrestrial and aquatic invertebrates as bioindicators for environmental monitoring, with particular reference to mountain ecosystems. *Environmental Management*, 35: 649-666.

HOLLAND, B.S., & M.G. HADFIELD, 2002. Islands within an island: phylogeography and conservation genetics of the endangered Hawaiian tree snail *Achatinella mustelina*. *Molecular Ecology*, 11: 365-375.

LOSOS, J.B., & R.E. RICKLEFS (eds.), 2010. *The theory of island biogeo-*

graphy revisited, 476 pp. Princeton University Press, Princeton and Oxford.

MARCELINO, J., R. GIORDANO, S. GOULI, V.V. GOULI, B.L. PARKER, M. SKINNER, R. CESNIK & D. TEBEEST, 2008. *Colletotrichum acutatum* var. *fioriniae (*teleomorph: *Glomerella acutata* var. *fioriniae* var. nov.)* infection of a scale insect. *Mycologia*, 100: 353-374.

MARTÍN, J.L, M. ARECHAVALETA, P.A.V. BORGES & B. FARIA (eds.), 2008. *TOP 100 - As cem espécies ameaçadas prioritárias em termos de gestão na região europeia biogeográfica da Macaronésia*, 500 pp. Consejería de Medio Ambiente y Ordenación Territorial, Gobierno de Canarias.

PAPADOPOULOU, A., I. ANASTASIOU & A.P. VOGLER, 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology and Evolution*, 7: 1659-1672.

PEARSALL, I.A., 2007. Carabid beetles as ecological indicators. *In*: *Monitoring the effectiveness of biological conservation conference*, pp. 389-399. Richmond, British Columbia, USA, 2-4 November 2004.

RODRÍGUEZ-GIRONÉS, M.A., & L. SANTAMARÍA, 2006. A new algorithm to calculate the nestedness temperature of presence-absence matrices. *Journal of Biogeography*, 33: 924-935.

SUKHDEV, P. (Lead Author) 2008. *The economics of ecosystems and biodiversity. The TEEB report,* 64 pp. Welzel+Hardt, Wesseling, Germany. European Communities 2008.

SWOFFORD, D.L., 2001. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.

THIES, C., I., STEFFAN-DEWENTER & T. TSCHARNTKE, 2008. Interannual landscape changes influence plant-herbivore-parasitoid interactions. *Agriculture, Ecosystems and Environment*, 125: 266-268.

TYLIANAKIS, J.M., T. TSCHARNTKE & O.T. LEWIS, 2007. Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, 455: 202-205.