Insect Conservation and Diversity

Insect Conservation and Diversity (2016) **9,** 475–479 doi: 10.1111/icad.12187

SHORT COMMUNICATION

An integrative taxonomy approach unveils unknown and threatened moth species in Amazonian rainforest fragments

GREG P. A. LAMARRE,[1,2,*] THIBAUD DECAËNS,[3,*] RODOLPHE ROUGERIE,[4] JÉRÔME BARBUT,[4] JEREMY R. DEWAARD,[5] PAUL D. N. HEBERT,[5] DANIEL HERBIN,[6] MICHEL LAGUERRE,[7] PAUL THIAUCOURT[8] and MARLUCIA BONIFACIO MARTINS[9] [1]INRA - UMR Ecologie des Forets de Guyane, Kourou, French Guiana, [2]Biology Centre, Czech Academy of Science, Institute of Entomology, Ceske Budejovice, Czech Republic, [3]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175, CNRS–Univ Montpellier–Univ Paul-Valéry Montpellier–EPHE, Montpellier, France, [4]Institut de Systématique, Evolution, Biodiversité, UMR 7205, CNRS–MNHN–UPMC–EPHE–Muséum national d'Histoire naturelle–Sorbonne Univ, Paris, France, [5]Centre for Biodiversity Genomics, Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, Canada, [6] Garidech, France, [7] Léognan, France, [8] Paris, France and [9]Laboratório de Ecologia de Invertebrados, Coordenação de Zoologia, Museu Paraense Emilio Goeldi, Belém, Brazil

Abstract. 1. This study focuses on the importance in hyperdiverse regions, such as the Amazonian forest, of accelerating and optimising the census of invertebrate communities.

2. We carried out low-intensity sampling of tropical moth (Lepidoptera) assemblages in disturbed forest fragments in Brazil.

3. We combined DNA barcoding and taxonomists' expertise to produce fast and accurate surveys of local diversity, including the recognition and census of undescribed and endemic species.

4. Integrating expert knowledge of species distributions, we show that despite limited sampling effort, our approach revealed an unexpectedly high number of new and endemic species in severely threatened tropical forest fragments.

5. These results highlight the risk of silent centinelan extinctions and emphasise the urgent need for accelerated invertebrate surveys in high-endemism and human-impacted tropical forests.

Key words. Amazonian forest, Belém center of endemism, centinelan extinction, conservation, DNA barcoding, Lepidoptera, species discovery.

Introduction

Forest fragmentation and habitat loss are among the major causes of global biodiversity decline (Wilson, 2002; Ceballos *et al*., 2015). Our capacity to record and measure species extinctions and the associated loss of ecosystem services in impacted environments is strongly impeded by incomplete knowledge of this diversity, as best illustrated by hyperdiverse invertebrate communities in tropical rainforests (Basset *et al*., 2012; Lamarre *et al*., 2016). In insects, the likelihood for a species to become extinct has recently been shown to be far more likely than previously thought (Regnier *et al*., 2015). Hence, most species extinctions remain silent and their impact unnoticed (Fonseca, 2009). In a context of escalating threats in the Amazon basin due to land-use changes (Fearnside, 2005; Asner *et al*., 2010) and climate anomalies (Malhi *et al*., 2008), there is a critical need to develop and apply efficient approaches to address this knowledge gap and therefore obtain a more comprehensive understanding of the impact of these changes on communities (Diniz-Filho *et al*., 2010).

Entomological surveys in tropical rainforests are notoriously challenging. The study of the diversity of insect

Correspondence: Greg Lamarre, Biology Centre, Czech Academy of Science, Institute of Entomology, Branisovska 31, Ceske Budejovice 37005, Czech Republic. E-mail: greglamarre973@gmail.com

*Contributed equally to the paper

communities requires vast sampling effort and broad taxonomic expertise, both usually coming up short because of field constraints and lack of available resources. Consequently, reaching the asymptote in surveys of insect diversity is generally elusive (Fontaine *et al.*, 2012; Ashton *et al.*, 2014; Lees *et al.*, 2014). This issue may limit our ability to devise conservation strategies and policies accounting for this dominant segment of global biodiversity.

Our study investigates the value of a limited sampling effort on multiple moth families through standardised snap-shots using a combination of DNA barcoding and taxonomic expertise. We apply this approach on disrupted forest fragments in the Belém Center of Endemism (Pará and Maranhão States, Brazil; hereafter BCE) with a special focus on the discovery of undescribed and endemic species that can be used to rapidly inform the need for local conservation policies to prevent centinelan extinctions, i.e. extinction before species are even discovered and described (Wilson, 2002).

Materials and methods

Forest fragments in the BCE

Sampling was carried out in April 2009 at one site within each of four forest fragments in the Brazilian states of Pará and Maranhão (see Fig. S1): (i) Mojú: the Emprapa forest reserve in Mojú (Pará, 2°10′48.1″S, 48°48′06.1″W); (ii) Mokambo: the urban Park of Mokambo (Belém, Pará: 1°26′00.5″S, 48°24′40.0″W); (iii) Gunma: the Ecological Park of Gunma (Belém, Pará, 1°12′47.6″S, 48°17′22.6″W) and (iv) Gurupi: the Integral Biological Reserve of Gurupí (Açailândia, Maranhão 4°00′05.2″S, 46°50′13.9″W). All these sites are covered with primary and/or old secondary rainforest, and altogether belong to the BCE (Amaral *et al.*, 2012).

Moth sampling and identification procedure

We performed a limited standardised sampling of moth assemblages for two nights at each site (Fig. 1a). Specimens were attracted using a standard light trap consisting of a 125W mercury vapour bulb illuminating a white sheet of 2 m by 3 m. All macro-moths were manually collected, killed in cyanide jars or by injection of ammonia, and then placed in labelled glassine envelopes. Sampling extended for the entire night (Lamarre *et al.*, 2015) during the dark phase of the moon to maximise sampling efficiency. Specimens were sorted into morphospecies based on external morphological characters, and up to five specimens per morphospecies and per collecting night were processed through DNA barcoding. DNA was extracted from a single dry leg or a leg fragment using a routine high-throughput automated silica-based protocol in 96-well plates (Ivanova *et al.*, 2006), at the Canadian Centre for DNA Barcoding (Biodiversity Institute of Ontario, University of Guelph, Ontario, Canada). The barcode region of the COI gene (Hebert *et al.*, 2003) was amplified and sequenced following routine protocols

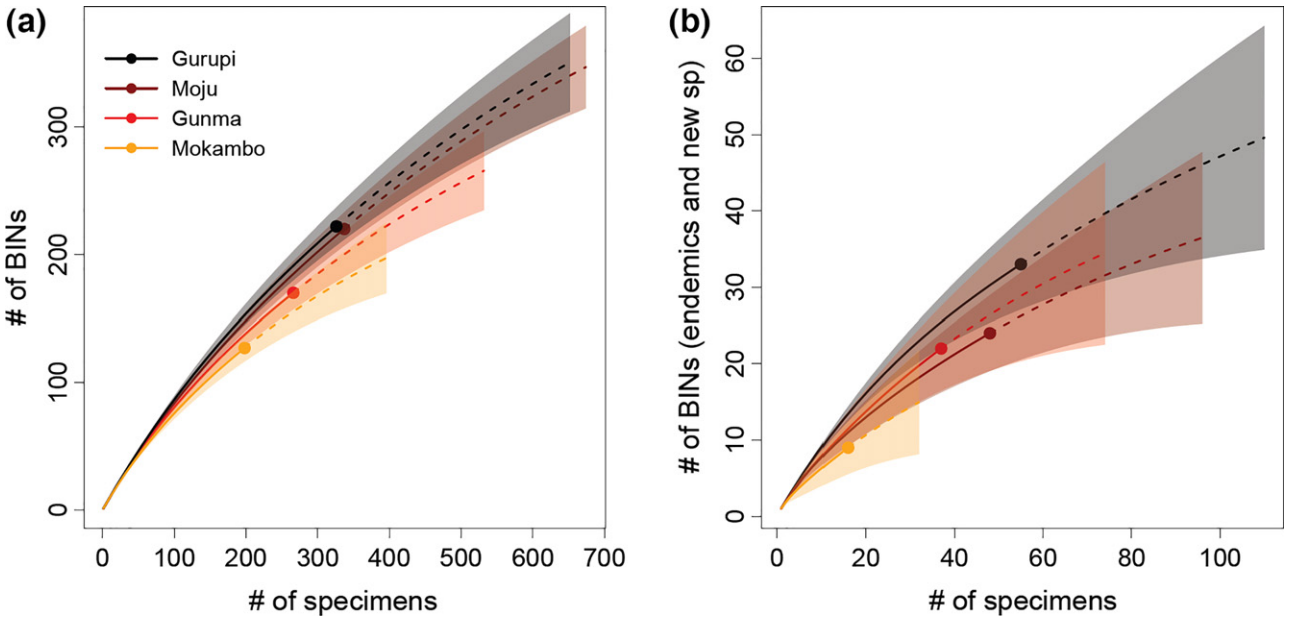

Fig. 1. Rarefaction and extrapolation curves showing (a) the observed species richness for each forest fragment (continuous line) and the predicted species richness if sampling effort was doubled (dotted line or shaded area); (b) the number of new and endemic species observed for each forest fragment (continuous line) and the predicted number of new and endemic species if sampling effort was doubled (dashed line or shadow).

(Hebert *et al.*, 2004; Hajibabaei *et al.*, 2005, see also http://www.ccdb.ca/resources.php for online protocols). All sequences, electropherograms, specimen data and pictures are publically available in the public dataset named DS-INCT01 (doi: 10.5883/DS-INCT01) in the Barcode of Life Data Systems (BOLD; www.boldsystems.org).

Voucher specimens were deposited in the Museu Paraense Emílio Goeldi (Belém, Pará, Brazil).

Data analysis

As a first analytical step, we used automatically generated Barcode Identification Numbers (or BINs, as implemented in BOLD) as a proxy for species (Hausmann *et al.*, 2013; Ratnasingham & Hebert, 2013). BINs proved to be good estimators of species richness in Lepidoptera (Zahiri *et al.*, 2014; Zenker *et al.*, 2016) and alternative methods such as ABGD (Puillandre *et al.*, 2011) and PTP (Zhang *et al.*, 2013) produced comparable outcomes for our dataset (results not shown). The BIN results were used to plot rarefaction and extrapolation curves and to obtain species richness estimates using the iNEXT package for R (Chao & Jost, 2012). We used characters of the habitus visible from images, as well as DNA sequence comparisons on BOLD (Ratnasingham & Hebert, 2007), to make a family-level taxonomic assignment for each BIN as well as genus- or species-level identifications when sequences match to taxa in the DNA barcode reference library permitted it. In seven families (see Table 1), each specimen was examined and studied by expert taxonomists who (i) confirmed or provided species-level identifications for known species, (ii) identified specimens they considered to represent undescribed species, and (iii) classified all recognised species as either endemic or non-endemic to BCE. By combining both barcoding results and taxonomy expertise, we were able to estimate the number of putative endemics and new species in the studied forest fragments.

Table 1. Summary of the diversity in all families, showing the number of individuals sampled (*n*), the number of BINs observed (# BINs), the number and percentage of endemic species for BCE (# end. sp. and % end.) and of new taxa for science (# sp. nov. and % sp. nov.). Lines highlighted in grey represent the families for which taxonomic expertise was available.

Families	*n*	# BINs	# end. sp.	% end.	# sp. nov.	% sp. nov.
Apatelodidae	33	8	3	37.5	3	37.5
Bombycidae	8	2	1	50.0	1	50.0
Cossidae	13	7				
Crambidae	28	17				
Dalceridae	2	2				
Depressariidae	2	2				
Erebidae	304	179	32	17.8	36	20
Geometridae	159	80				
Hesperiidae	1	1				
Lasiocampidae	33	22				
Limacodidae	17	9				
Megalopygidae	23	15				
Mimallonidae	17	12	2	16.7	4	33.3
Noctuidae	33	24				
Nolidae	3	3				
Notodontidae	201	113	40	35.4	29	25.6
Psychidae	1	1				
Pyralidae	11	6				
Riodinidae	1	1				
Saturniidae	168	66	15	22.7	16	24.2
Sphingidae	54	22	0	0	0	0
Thyrididae	9	4				
Tineidae	1	1				
Uraniidae	2	2				
Not identified	3	3				
Total	1127	601	93		91	

Results and Discussion

Our study examined 1127 specimens belonging to 25 macro-moth families and 601 distinct BINs that were treated as species-level proxies. Gurupi, the largest forest fragment in our study, was the most diverse site (222 BINs), followed by Mojú (220 BINs), Gunma (170 BINs) and Mokambo (127 BINs). With the crucial taxonomists' expertise integrated into our survey (Table 1), we identified 94 BCE endemic species and 90 presumptively undescribed species (Table 1, see details in Table S2). Even in the least diverse urbanised site of Mokambo, our results revealed 8.6% endemic and 7.1% of the BINs could correspond to undescribed species. In Erebidae, Notodontidae and Saturniidae, the three most diverse families examined by expert taxonomists, endemic species accounted for 18%, 35% and 23%, while undescribed species comprised almost 20%, 26% and 24% of the BINs respectively. Though less diverse, half of the diversity of BINs in the family Apatelodidae from the BCE was either endemic or undescribed species. Unsurprisingly, the well-studied, strongly flying hawkmoths (Sphingidae) showed neither endemics nor undescribed species.

Our results show that despite sampling for only two nights per site and one site per forest fragment (Fig. 1a), we collected numerous new and endemic species in isolated tropical forest fragments, some of which are severely threatened by human activity (in spite of current protection measures in Gurupi). The ratio of BCE endemics and undescribed species unveiled is striking (Table 1, Fig. 1b, Table S2), even in large, well-documented moths such as the wild silkmoths (Saturniidae). Considering that almost 70% of the forested area in BCE has now been cleared (Da Silva *et al.*, 2005), our results suggest that a large portion of insect species have already faced centinelan extinction, meaning that they may have become extinct before being described or discovered (Wilson, 2002). This result strongly stresses the importance of preserving forest remnants in the BCE (Da Silva *et al.*, 2005; Pimm *et al.*,

2010) and the urgency for new conservation strategies to mitigate human land-use and deforestation for both large and small forest remnants of the BCE (see Koh & Sodhi, 2004 for the importance of urban fragments in butterfly conservation).

Assessing the conservation status of insect communities in tropical regions is strongly impeded by the lack of knowledge concerning their diversity and the distribution ranges of species. The snapshot of moth diversity and endemism produced in this study revealed the deep lack of understanding of biodiversity in these communities. Although our approach required time-consuming taxonomic expertise, a major limitation to its scaling up, we stress the very powerful contribution made by integrating DNA barcoding in the workflow. Reference libraries of DNA barcodes for Lepidoptera are the best developed to date for any group of organisms (>1M sequenced records on BOLD; www.lepbarcoding.org). These reference sequences made it possible for us to rapidly identify a large number of our specimens to a known species through sequence matching, greatly reducing the burden on expert taxonomists and allowing them to focus on identification and study of samples whose DNA sequences represented new BINs in the BOLD database.

Our study demonstrates that molecular tools using available library provide a basis for a greatly accelerated census that allows the documentation of species that would otherwise remain unnoticed. As such, it provides the volume of data needed to diagnose ongoing challenges in tropical ecosystem. This approach will become increasingly powerful when the barcode reference library gains further parameterisation because it overcomes the taxonomic impediment which leads to decade-long delays in the formal description of species new to science (Mora *et al*. 2011) which may be endemics of a given region. Because DNA libraries are still incomplete, taxonomic classification is still fundamentally needed. Our study highlights that taxonomic expertise holds great potential to detect novelty and/or extinction risks among tropical forest fragments when gleaned through DNA barcoding. Further development of comprehensive DNA barcode libraries for described species of invertebrates, as is the case globally for the Saturniidae and Sphingidae families, along with distributional databases integrating expert knowledge (Jetz *et al*., 2012) represents a key endeavour to promote the integration of invertebrate communities into conservation studies aiding the measurement and prevention of centinelan extinctions.

Acknowledgements

Research was supported by INCT/CNPq collaborative grant (No. 574008/2008-0) «*Perda de Biodiversidade no centros de endemismo do arco de desmatamento*» that supported field work and specimens deposition at Museu Paraense Emílio Goeldi (Belém, Pará, Brazil). Greg P.A. Lamarre benefited from an "*Investissement d'Avenir*" (ANR-10-LABX-25-01) and an ERC Grant (No. 669609) during the writing of the paper. Financial support for barcode analysis was provided by the Ontario Ministry of Research and Innovation and by the Government of Canada through Genome Canada and the Ontario Genomics Institute in support of the International Barcode of Life project. We thank Louise Ashton for constructive comments on early version of the manuscript.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: doi: 10.1111/icad.12187:

Figure S1. Location of the four forest fragments that were studied within the Belém Center of Endemism in Brazil, including local-scale GIS images of land use metrics; 1) the Emprapa forest Reserve in Mojú; 2) the Urban Park of Mokambo; 3) the Ecological Park of Gunma and 4) the Integral Biological Reserve of Gurupí.

Table S2. Data summary for new and endemic species collected at each site during this study. n=number of specimens; # BINs= number of BINs observed; Chao1= Chao1 estimator of species richness (Chao & Jost, 2012); ACE= Abundance-based coverage estimator of species richness (Chao, 2005); C.hat = Estimator of sample coverage (Chao & Jost, 2012); New species refers to putatively new species for science found in our sample set.

References

Amaral, D.D., Vieira, I.C.G., Salomão, R.P., Almeida, S.S. & Jardim, M.A.G. (2012) The status of conservation of urban forests in eastern Amazonia. *Brazilian Journal of Biology*, **72**, 257–265.

Ashton, L.A., Barlow, H.S., Nakamura, A. & Kitching, R.L. (2014) Diversity in tropical ecosystems: the species richness and turnover of moths in Malaysian rainforests. *Insect Conservation and Diversity*, **8**, 132–142.

Asner, G.P., Loarie, S.R. & Heyder, U. (2010) Combined effects of climate and land-use change on the future of humid tropical forests. *Conservation Letters*, **3**, 395–403.

Basset, Y., Cizek, L., Cuenoud, P., Didham, R.K., Guilhaumon, F., Missa, M., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J., Tishechkin, A.K., Winchester, N.N., Aberlenc, H.P., Bail, J., Barrios, H., Bridle, J.R., Castano-Meneses, G., Corbara, B., Curletti, G., Darocha, W.D., Debaker, D., Delabie, J.H.C., Dejean, A., Fagan, L.L., Floren, A., Kitching, R.L., Medianero, L.E., Miller, S.E., Gama De Oliveira, E., Orivel, J., Pollet, M., Rapp, M., Ribeiro, S.P., Roisin, Y., Roubik, D.W., Schmidt, J.B., Sorensen, L. & Leponce, M. (2012) Arthropod diversity in a tropical forest. *Science*, **338**, 1481–1484.

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., Garcia, A., Pringle, R.M. & Palmer, T.M. (2015) Accelerated modern human–induced species losses: entering the sixth mass extinction. *Science Advances*, **1**, e1400253.

Chao, A. & Jost, L. (2012) Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, **93**, 2533–2547.

Da Silva, J.M.C., Rylands, A.B. & Da Fonseca, G.A.B. (2005) The fate of the Amazonian areas of endemism. *Conservation Biology*, **19**, 689–694.

Diniz-Filho, J.A.F., De Marco, P Jr & Hawkins, B.A. (2010) Defying the curse of ignorance: perspectives in insect macroecology and conservation biogeography. *Insect Conservation and Diversity*, **3**, 172–179.

Fearnside, P.M. (2005) Deforestation in Brazilian Amazonia: history, rates, and consequences. *Conservation Biology*, **19**, 680–688.

Fonseca, C.R. (2009) The silent mass extinction of insect herbivores in biodiversity hotspots. *Conservation Biology*, **23**, 1507–1515.

Fontaine, B., Perrard, A. & Bouchet, P. (2012) 21 years of shelf life between discovery and description of new species. *Current Biology*, **22**, 943–944.

Hajibabaei, M., deWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M. & Hebert, P.D.N. (2005) Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society B*, **360**, 1959–1967.

Hausmann, A., Godfray, H.C., Huemer, P., Mutanen, M., Rougerie, R., van Nieukerken, E.J., Ratnasingham, S. & Hebert, P.D.N. (2013) Genetic patterns in European geometrid moths revealed by the Barcode Index Number (BIN) system. *PLoS ONE*, **8**, e84518.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B*, **270**, 313–321.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H. & Hallwachs, W. (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, **101**, 14812–14817.

Ivanova, N.V., deWaard, J.R. & Hebert, P.D.N. (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, **6**, 998–1002.

Jetz, W., McPherson, J.M. & Guralnick, R.P. (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology and Evolution*, **27**, 151–159.

Koh, L.P. & Sodhi, M.S. (2004) Importance of reserves, fragments, and parks for butterfly conservation in a tropical urban landscape. *Ecological Applications*, **14**, 1695–1708.

Lamarre, G.P.A., Hérault, B., Fine, P.V.A., Vedel, V., Lupoli, R., Mesones, I. & Baraloto, C. (2016) Taxonomic and functional composition of arthropod assemblages across contrasting Amazonian forests. *Journal of Animal Ecology*, **85**, 227–239.

Lamarre, G.P.A., Mendoza, I., Rougerie, R., Decaëns, T. & Bénéluz, F. (2015) Stay out (almost) all night: contrasting responses in flight activity among tropical moth assemblages. *Neotropical Entomology*, **44**, 109–115.

Lees, D.C., Kawahara, A.Y., Rougerie, R., Oshima, I., Kawakita, A., Bouteleux, O., De Prins, J. & Lopez-Vaamnonde, C. (2014) DNA barcoding reveals a largely unknown fauna of Gracillariidae leaf mining moths in the Neotropics. *Molecular Ecology Resources*, **14**, 286–296.

Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W. & Nobre, C.A. (2008) Climate change, deforestation, and the fate of the Amazon. *Science*, **319**, 169–172.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011) How Many Species Are There on Earth and in the Ocean? *PLoS Biology*, **9**, e1001127.

Pimm, S.L., Jenkins, C.N., Joppa, L.N., Roberts, D.L. & Russel, G.J. (2010) How many endangered species remain to be discovered in Brazil? *Natureza & Conservação*, **8**, 71–77.

Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2011) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology*, **21**, 1864–1877.

Ratnasingham, S. & Hebert, P.D.N. (2007) BOLD: the Barcode of Life Data System. *Molecular Ecology Notes*, **7**, 355–364.

Ratnasingham, S. & Hebert, P.D.N. (2013) A DNA-based registry for all animal species: the Barcode Index Number (BIN) System. *PLoS ONE*, **8**, e66213.

Regnier, C., Achaz, G., Lambert, A., Cowie, R.H., Bouchet, P. & Fontaine, B. (2015) Mass extinction in poorly known taxa. *Proceedings of the National Academy of Sciences*, **112**, 7761–7766.

Wilson, E.O. (2002) *The future of life*. Vintage Books, New York city, New York.

Zahiri, R., Lafontaine, J.D., Schmidt, B.C., Dewaard, J.R., Zakharov, E.V. & Hebert, P.D. (2014) A transcontinental challenge - a test of DNA barcode performance for 1,541 species of Canadian Noctuoidea (Lepidoptera). *PLoS ONE*, **9**, e92797.

Zenker, M.M., Rougerie, R., Teston, J.A., Laguerre, M., Pie, M.R. & Freitas, A.V.L. (2016) Fast census of moth diversity in the neotropics: a comparison of field-assigned morphospecies and DNA barcoding in tiger moths. *PLoS ONE*, **11**, e0148423.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**, 2869–2876.

Accepted 5 July 2016
First published online 13 August 2016

Editor: Raphael Didham
Associate editor: Tom Oliver