The utility of open-access biodiversity information in representing anurans in the Brazilian Atlantic Forest and Cerrado

Tiago S. Vasconcelos and Bruno T. M. Nascimento

Departamento de Ciências Biológicas (DCB), Faculdade de Ciências (FC), Universidade Estadual Paulista (UNESP), 17033-360 Bauru, SP, Brazil. E-mail: zoologia@ig.com.br.

Abstract

The utility of open-access biodiversity information in representing anurans in the Brazilian Atlantic Forest and Cerrado. Many geographic coordinates of species are only available in biological collections, and sometimes it is difficult to access these data. However, recent initiatives promise to compile and organize such biodiversity data at a global scale. We evaluated the effectiveness of open-access biodiversity information for forecasting the occurrence of anurans in two Brazilian hotspots - the Atlantic Forest (AF) and the Cerrado (CER). We compiled all available point-occurrence records for anuran species in both hotspots and identified the regions in each of the two hotspots having the highest and lowest number of anuran occurrences based on information from the Global Biodiversity Information Facility (GBIF) and the SpeciesLink project. A total of 13,130 point-occurrence records were recovered for 409 anurans from the AF (~77% of the known species) and 12,729 records for 176 species (~85% of the known species) from the CER. Density of point occurrence data is not randomly distributed within the hotspots. The greatest density of Atlantic Forest records is in southeastern Brazil, and the densest areas in Cerrado occur in the transitional zones to the Atlantic Forest. Comparison of these results with previous studies based on traditional museum information revealed that many important collections of anurans from these hotspots are not included in GBIF and SpeciesLink.

Keywords: amphibians, GBIF, point density analysis, SpeciesLink.

Resumo

A utilidade de informações de livre acesso à biodiversidade para a representatividade de anuros da Mata Atlântica e Cerrado brasileiro. Muitas coordenadas geográficas de espécies estão disponíveis apenas em coleções biológicas, sendo muitas vezes de difícil acesso. No entanto, existem iniciativas recentes para compilar e organizar esses dados de biodiversidade em escala global. Nós

avaliamos a efetividade dessas ferramentas gratuitas de acesso à informação biológica para representar a ocorrência de anuros nos hotspots brasileiros Cerrado (CER) e Mata Atlântica (MA). Nós compilamos os registros de pontos de ocorrência de todas as espécies disponíveis em ambos os hotspots e identificamos regiões com alta e baixa representatividade de ocorrências de anuros para ambos os hotspots utilizando duas ferramentas gratuitas de busca por informação biológicas (*The Global Biodiversity Information Facility* - GBIF; projeto SpeciesLink). Foram registrados 13130 pontos de ocorrência para 409 anuros da Mata Atlântica (~77% das espécies conhecidas) e 12729 pontos para 176 espécies do Cerrado (~85% das espécies conhecidas). A densidade dos pontos de ocorrência não está distribuída aleatoriamente ao longo de ambos os hotspots. A maioria dos registros da MA está concentrada no sudeste do Brasil, sendo que as áreas mais densas no CER ocorrem em áreas de transição com a MA. Comparando com estudos prévios que utilizaram dados de coleções não implementadas nas ferramentas utilizadas neste estudo, concluímos que o GBIF e SpeciesLink não hospedam coleções biológicas importantes para anfíbios do CER e MA.

Palavras-chave: anfíbios, análise de densidade pontual, GBIF, SpeciesLink.

Introduction

Point-occurrence data are the basic units used to generate the geographic distributions of species by methods such as "point-to-grid," "expert," or "ecological niche model" maps (Elith and Burgman 2002, Graham and Hijmans 2006, Vasconcelos *et al*. 2012). Although occurrence records are the available scientific literature, new, web-based initiatives facilitate the access to the geographic information about species (e.g., NatureServe: http://www.natureserve.org/; International Union for Conservation of Nature: http://www.iucnredlist.org); some serve specific taxonomic groups, such as amphibians (AmphibiaWeb 2014, Frost 2014). However, much of the raw data (i.e., the geographic coordinates) has not been captured electronically or is only available in electronic files of individual herpetological collections. Consequently, these data are not readily accessible and their lack of availability may limit the advance of basic and applied ecological studies (Beck *et al*. 2014 and references therein). The Global Biodiversity Information Facility (GBIF; Yesson *et al*. 2007) and the SpeciesLink (http://splink. cria.org.br/) projects seek to compile and organize point-occurrence data from biological surveys and biocollections globally and provide open access to this information via an online data search portal. It is not unusual for there to be a spatial bias and/or incomplete dataset for any taxonomic group included in these databases owing to sampling error, and data storage and mobilization (Yesson *et al*. 2007, Beck *et al*. 2014). Nevertheless, data derived from these sources are being used increasingly for purposes such as taxonomic revisions (Pennisi 2000), environmental niche modeling (Vasconcelos *et al*. 2012), compilation of "redlists" of threatened species (Shaffer *et al*. 1998), and biodiversity assessment (Ponder *et al.* 2001, Yesson *et al*. 2007 and references therein).

The Atlantic Forest (AF) and the Cerrado (CER) are two of five South American biodiversity hotspots (i.e., highly biodiverse formations that are critically endangered by human populations; *sensu* Mittermeier *et al*. 2004). The Atlantic Forest is a complex forest formation along the Brazilian Atlantic coast, whereas the Cerrado is a tropical savanna mainly distributed in central Brazil (MMA and IBAMA 2010, IBGE 2012). There are approximately 550 species of anurans in the AF and 80% of species are endemic (Haddad *et al*. 2013), and 209 anuran species in the CER, of which 108 are

endemic (Valdujo *et al.* 2012). Given the importance of the AF and CER to anuran biology, we evaluate the utility and accuracy of open-access biodiversity information for anurans in these two hotspots. We compiled all available point-occurrence records for anuran species in both hotspots and identified the regions in each of the two hotspots having the highest and lowest number of anuran occurrences based on information from the Global Biodiversity Information Facility (GBIF) and the SpeciesLink project.

Materials and Methods

Examples of studies using GBIF and SpeciesLink databases can be found in Vasconcelos *et al.* (2012), Beck *et al.* (2014), and references therein. From June to October 2013, we downloaded the point observations species-by-species, based on anuran lists available for the AF (Haddad *et al.* 2013) and CER (Valdujo *et al.* 2012). Although we followed the amphibian nomenclature of Frost (2014), the species recorded in the surveys also included synonyms. For example, occurrence points of *Rhinella schneideri* also involve synonymous species, such as *Bufo paracnemis*, *B. schneideri*, and *Chaunus schneideri*. Subsequent validation of the records and database involved deletion of: (1) records lacking geographic coordinates; (2) records based on indeterminate geographical information (e.g., municipalities, farms, and/or wide areas); and (3) duplicate geographic coordinates.

Point-density analyses in ArcGIS 10.1 were performed on the validated datasets of GBIF and SpeciesLink records for AF and CER anurans. This analysis generates a rasterized map depicting the magnitude of point occurrences per unit area that fall within a neighborhood (~100 km) around each record. The final map of each hotspot will indicate areas with highest and lowest densities of occurrence records, thereby allowing us to identify areas that are either well- or under-represented by the open-access biodi-

versity data; a similar approach was used by Rossa-Feres *et al.* (2011).

Results

We obtained 13,130 point-occurrence records for 409 anurans from the AF and 12,729 records for 176 species from the CER. The abundance of records per species varied from 1 to 685 occurrence points, but most AF species have fewer than 10 records (Appendix I), whereas in CER most species have fewer than 25 records (Appendix II).

In the AF extent, the point-occurrence data (7,879 records) are concentrated in southeastern Brazil, mainly in the Atlantic coastal areas in the states of São Paulo and Espírito Santo, and inland areas of São Paulo State (Figure 1). The coastal areas in the states of Paraná and Rio Grande do Sul also have significant occurrence records (Figure 1). Although having lower densities, all other areas along the AF have at least some level of representative records, except for few empty cells in transitional areas to CER and Caatinga biome (Figure 1). On the other hand, the density of occurrence records found for anurans within the CER extent (2,835 records) evidences that this hotspot is under-estimated compared to the AF (Figure 2). There are more empty cells, i.e. no record within a ~100-km radius in the central-east and northern CER, than in the AF. The areas of greatest record density in CER occur in transitional regions between the CER and AF (in São Paulo and Minas Gerais states; Figure 2) and, to a lesser extent, in the state of Mato Grosso (Figure 2).

Discussion

Based on the works of Haddad *et al.* (2013) and Valdujo *et al.* (2012), we recorded approximately 77% and 85% of the known AF and CER anuran species, respectively. Thus, the GBIF and SpeciesLink provided qualitatively satisfactory results, despite the fact that some species (e.g., the recently described species
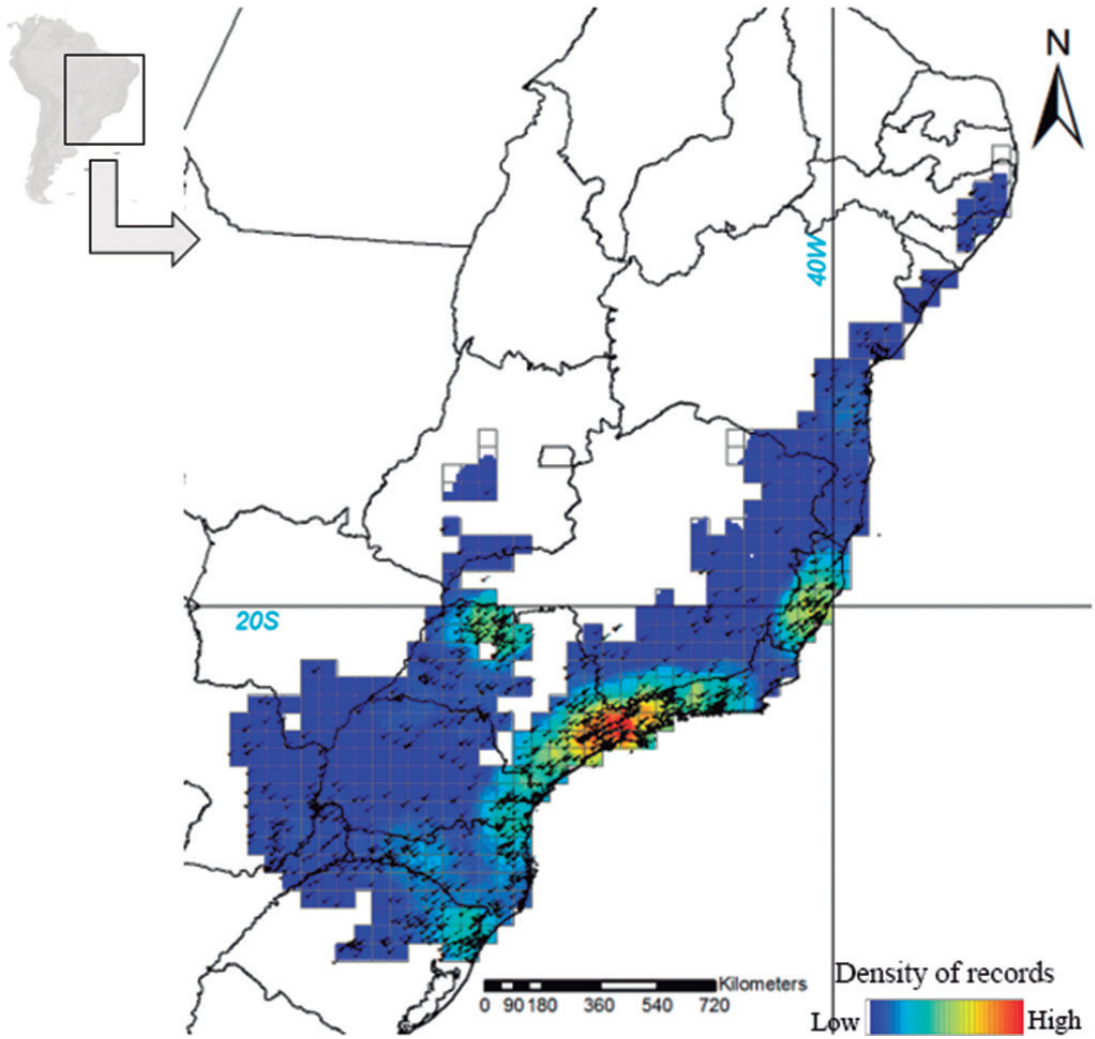
Figure 1. Map indicating the point occurrence records for the AF anurans (black dots) and the density of point records based on the point density analysis - circular neighborhood. The grid cells represent the AF original extent.

Brachycephalus pulex and *B. toby*) were not included in the databases. Although the total number of records for the AF anurans is higher than those for CER, the mean number of records per species is lower in the AF (32.18 vs. 71.61 for the CER). This may be because the AF has more anurans with small distributions than does

the CER. For instance, Haddad and Prado (2005) recorded 137 endemic anuran species known only from one locality in the AF. Since 2005, much more locality data for AF species has been reported, and some taxa thought to be endemic are now known to have broader distributions. At the same time, however, many more new species
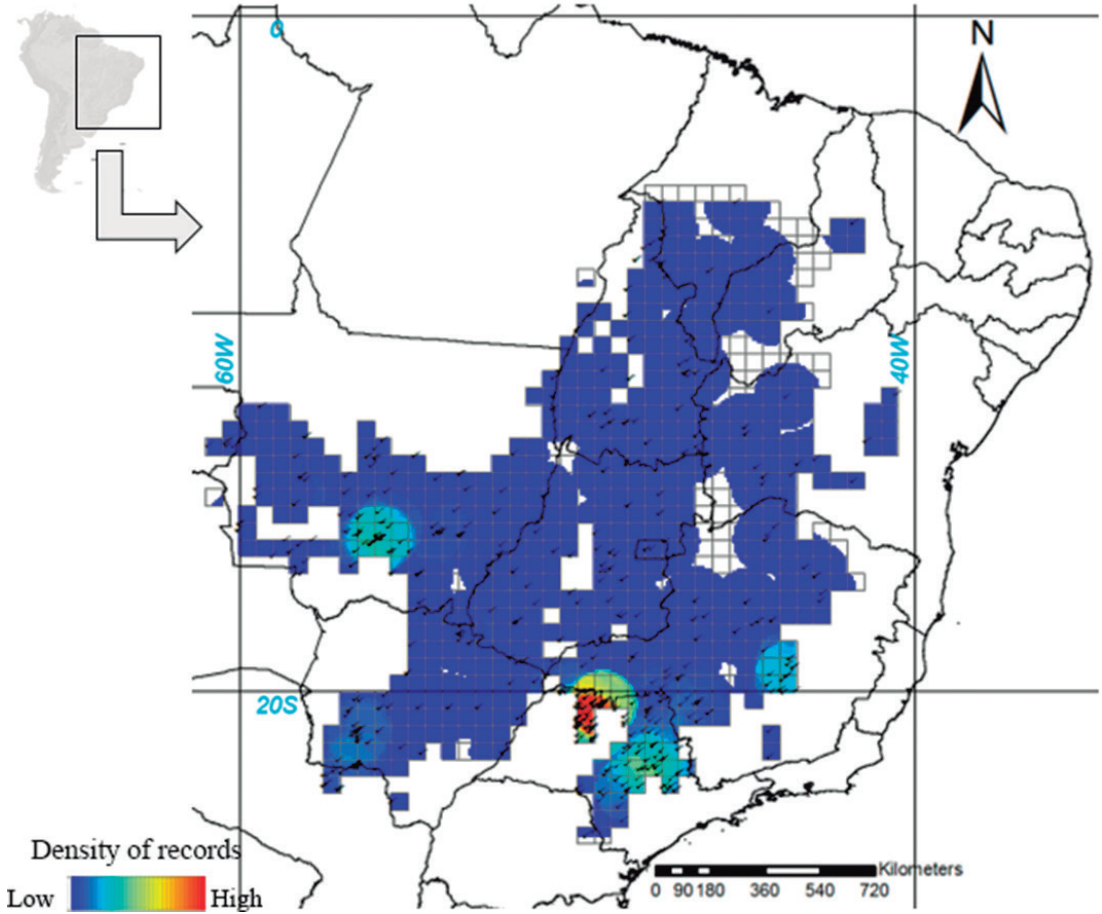
Figure 2. Map indicating the point occurrence records for the CER anurans (black dots) and the density of point records based on the point density analysis - circular neighborhood. The grid cells represent the CER original extent.

have been described from a single locality in the AF (e.g., Haddad *et al.* 2013). In contrast, 108 endemic anurans are recorded in the CER and included are species endemic to narrow areas along the CER (Valdujo *et al.* 2012). Hence, the higher numbers of species with limited distributions in the AF are presumed to generate low numbers of point occurrence records per species.

In contrast to these qualitative data derived from GBIF and SpeciesLink, we found that the point-occurrence density data indicate that the data are not randomly distributed across both hotspots, thereby reducing our confidence in our knowledge of the anurans that occur in them. This is not unanticipated because these databases are prone to errors and geographical bias (Yesson *et al.* 2007, Beck *et al.* 2014). The highest

density of occurrence records within the AF and CER is in southeastern Brazil (mainly in the state of São Paulo), and doubtless is a result of a recent initiative to advance the biodiversity knowledge of the state of São Paulo, with an increase in anuran surveys (the Biota-FAPESP Program, see Rossa-Feres *et al*. 2011). The vouchers from most of the anuran inventories supported by this program were deposited in herpetological collections that are hosted in the SpeciesLink database (e.g., Vasconcelos and Rossa-Feres 2005, 2008, Santos *et al*. 2007, 2009, Zina *et al*. 2007, Silva *et al*. 2012). Although other areas of Atlantic Forest have enough point records to be considered minimally represented, this is clearly not the case for anurans in the CER.

There are obvious gaps in the density of occurrence points of anurans in the CER. Although we might conclude that these areas represent distribution gaps, the compilation of point records of Valdujo *et al*. (2012) suggests that GIBF and SpeciesLink are incomplete. For example, in the case of some well-surveyed CER localities, the voucher specimens were deposited in collections not hosted by GIBF and SpeciesLink (e.g., Valdujo *et al.* 2009, 2011, Santos *et al.* 2014). Consequently, these localities seem to be under-represented by our density analysis; examples include the Estação Ecológica Serra Geral do Tocantins (states of Tocantins and Bahia, Brazil: Valdujo *et al*. 2011) and southeastern Goiás State (central Brazil: Santos *et al.* 2014). If the aim of a study is to analyze general macroecological patterns of anurans in the CER, then GIBF and SpeciesLink should not be the only source of point occurrence records (e.g., Valdujo *et al*. 2013).

Likewise, if the goal is to generate range maps with ecological niche models (Elith and Burgman 2002), then GBIF and SpeciesLink should be augmented with other occurrence records to provide a more complete and reliable dataset. These sources might include (1) local/regional/major herpetological collections that neither is digitalized nor hosted on online

databases (e.g., Coleção Zoológica da Universidade Federal de Goiás – ZUFG, Brazil; Museu Nacional do Rio de Janeiro – MNRJ, Brazil; Museu de Zoologia da Universidade Federal da Bahia – MZUFBA, Brazil; Museu de Zoologia da Universidade de São Paulo – MZUSP, Brazil; reference Valdujo *et al.* [2012] for more complete list); and (2) recent literature that appears in journals focusing on natural history, species lists, and distributional records.

In Brazil, as elsewhere, biological information is being garnered from monitoring plans or preliminary biological assessments involved in environmental impact studies. Although some of this information is published (e.g., Brasileiro *et al*. 2008), most probably remains unpublished and unavailable to the scientific community. We suggest that it is important to implement mandatory surveys as a part of such environmental studies, with properly identified voucher specimens being deposited in recognized institutional collections. These collections require either public or private funding to acquire the software (some available at no cost) and to train staff to georeference locality data, and organize and capture data for integration into online databases, such as GBIF. Taxonomists must be hired to verify the identities of specimens.

The Brazilian agencies responsible for issuing collecting permits require reports of specimens collected and/or recorded during fieldwork (e.g., see the ICMBio portal: http://www.icmbio.gov.br/sisbio/relatorio-de-atividades.html). Current Brazilian law/policy does not allow these agencies to share biological data with the scientific community; legally, the owner of the collecting permit is the only one who can use the biological data (see also the ICMBio portal). A database containing all of the biological information from environmental assessments should be available for the scientific community. It is imperative for these agencies to establish the rules for the access to these databases by the scientific community. And users must be aware that the data that they access

should be verified; specimens may have been misidentified or taxonomies may have changed. In summary, an extensive discussion is needed to change/adequate the current law and policy restrictions that regulates collecting permits in Brazil for the implementation of our suggestions, but our main aim at this time is to stimulate the debate in order to make the biological information more accessible for the scientific community.

Open-access biodiversity databases are an invaluable resource in the compilation and organization of otherwise dispersed and fragmented data. However, we corroborate the results of previous studies (e.g., Yesson *et al*. 2007, Beck *et al*. 2014) in finding that such databases are spatially biased for the AF and CER anurans. We encourage the incorporation and continuous updating of new collections to these open-access resources. To accomplish the latter, Brazilian decision-makers should promote the creation of Brazilian online databases and fund collection improvement to accelerate biological studies that are urgently needed to inform conservation and land-use management. With deforestation proceeding apace (MMA & IBAMA 2010), the data describing Brazil's spectacular diversity must be captured now.

Acknowledgments

References

AmphibiaWeb. 2014. *Information on amphibian biology and conservation*. Electronic Database accessible at http://amphibiaweb.org. Berkeley, California, USA. Captured on 31 January 2014.

Brasileiro, C. A., E. M. Lucas, H. M. Oyamaguchi, M. T. C. Thomé, and M. Dixo. 2008. Anurans, Northern Tocantins River Basin, states of Tocantins and Maranhão, Brazil. *Check List 4:* 185–197.

Beck, J., M. Böller, A. Erhardt, and W. Schwanghart. 2014. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics 19:* 10–15.

Elith, J. and M. Burgman. 2002. Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. Pp. 303–314 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, M. G. Raphael, W. A. Wall and F. B. Samson (eds.), *Predicting species occurrences: issues of accuracy and scale*. Covelo, California. Island Press.

Frost, D. R. (ed.) 2014. *Amphibian Species of the World: an Online Reference*. Version 6.0 (31 January 2014). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA. Captured on 31 January 2014.

Graham, C. H. and R. J. Hijmans. 2006. A comparison of methods for mapping species range and species richness. *Global Ecology and Biogeography 15:* 578–587.

Haddad, C. F. B. and C. P. A. Prado. 2005. Reproductive modes in frogs and their unexpected diversity in the Atlantic Forest of Brazil. *BioScience 55:* 207–217.

Haddad C. F. B., L. F. Toledo, C. P. A. Prado, D. Loebmann, J. L. Gasparini, and I. Sazima. 2013. *Guia de Anfíbios da Mata Atlântica: diversidade e biologia*. São Paulo. Anolisbooks. 544 pp.

IBGE (Instituto Brasileiro de Geografia e Estatística). 2012. *Manual técnico da vegetação brasileira*. Rio de Janeiro: IBGE, Brazil. 275 pp.

Mittermeier, R. A., P. R. Gil, M. Hoffmann, J. Pilgrim, T. Brooks, C. G. Mittermeier, J. Lamoreux, and G. A. B. Fonseca. 2004. *Hotspots revisited*. Mexico City: CEMEX. 392 pp.

MMA (Ministério do Meio Ambienta) and IBAMA (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis). 2010. *Monitoramento do desmatamento nos biomas brasileiros por satélite: monitoramento do bioma Mata Atlântica*. Brasília: MMA, Brasil. 42 pp.

Pennisi, E. 2000. Taxonomic revival. *Science 289*: 2306-2308.

Ponder, W. F., G. A. Carter, P. Flemons, and R. R. Chapman. 2001. Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology 15*: 648–657.

Rossa-Feres, D. C., R. J. Sawaya, J. Faivovich, J. G. R. Giovanelli, C. A. Brasileiro, L. Schiesari, J. Alexandrino, and C. F. B. Haddad. 2011. Anfíbios do estado de São Paulo, Brasil: conhecimento atual e perspectivas. *Biota Neotropica 11*: http://www.biotaneotropica.org.br/v11n1a/pt/abstract?inventory+bn0041101a2011

Santos, D. L., S. P. de Andrade, E. P. Victor-Jr, and W. Vaz-Silva. 2014. Amphibians and reptiles from southeastern Goiás, Central Brazil. *Check List 10:* 131–148.

Santos, T. G., D. C. Rossa-Feres, and L. Casatti. 2007. Diversidade e distribuição espaço-temporal de anuros em região com pronunciada estação seca no sudeste do Brasil. *Iheringia, Série Zoologia 97:* 37–49.

Santos, T. G., T. S. Vasconcelos, D. C. Rossa-Feres, and C. F. B. Haddad. 2009. Anurans of a seasonally dry tropical forest: Morro do Diabo State Park, São Paulo State, Brazil. *Journal of Natural History 43:* 973–993.

Shaffer, H. B., R. N. Fisher, and C. Davidson. 1998. The role of natural history collections in documenting species declines. *Trends in Ecology and Evolution 13:* 27–30.

Silva, F. R., M. Almeida-Neto, V. H. M. Prado, C. F. B. Haddad, and D. C. Rossa-Feres. 2012. Humidity levels drive reproductive modes and phylogenetic diversity of amphibians in the Brazilian Atlantic forest. *Journal of Biogeography 39:* 1720–1732.

Valdujo, P. H., R. S. Recoder, M. M. Vasconcellos, and A. S. Portella. 2009. Amphibia, Anura, São Desidério, western Bahia uplands, northeastern Brazil. *Check List 5:* 903–911.

Valdujo, P. H., A. Camacho, R. S. Recoder, M. Teixeira Junior, J. M. B. Ghellere, T. Mott, P. M. S. Nunes, C. Nogueira, and M. T. Rodrigues. 2011. Anfíbios da Estação Ecológica Serra Geral do Tocantins, região do Jalapão, estados do Tocantins e Bahia. *Biota Neotropica 11:* 251–262.

Valdujo, P. H., D. L. Silvano, G. Colli, and M. Martins. 2012. Anuran species composition and distribution patterns in Brazilian Cerrado, a Neotropical hotspot. *South American Journal of Herpetology 7:* 63–78.

Valdujo, P. H., A. C. O. Q. Carnaval, and C. H. Graham. 2013. Environmental correlates of anuran beta diversity in the Brazilian Cerrado. *Ecography 36:* 708–717.

Vasconcelos, T. S. and D. C. Rossa-Feres. 2005. Diversidade, distribuição espacial e temporal de anfíbios anuros (Amphibia, Anura) na região noroeste do Estado de São Paulo, Brasil. *Biota Neotropica 5:* http://www.biotaneotropica.org.br/v5n2/pt/abstract?article+BN01705022005.

Vasconcelos, T. S. and D. C. Rossa-Feres. 2008. Habitat heterogeneity and use of physical and acoustic space in anuran communities in Southeastern Brazil. *Phyllomedusa 7:* 127–142.

Vasconcelos, T. S., M. Á. Rodríguez, and B. A. Hawkins. 2012. Species distribution modelling as a macroecological tool: a case study using New World amphibians. *Ecography 35:* 39–548.

Yesson, C., P. W. Brewer, T. Sutton, N. Caithness, J. S. Pahwa, M. Burgess, W. A. Gray, R. J. White, A. C. Jones, F. A. Bisby, and A. Culham. 2007. How global is the Global Biodiversity Information Facility? *PLoS ONE.* 2: e1124. doi:10.1371/ journal.pone.0001124.

Zina, J., J. Ennser, S. C. P. Pinheiro, C. F. B. Haddad, and L. F. Toledo. 2007. Taxocenose de anuros de uma mata semidecídua do interior do Estado de São Paulo e comparações com outras taxocenoses do Estado, sudeste do Brasil. *Biota Neotropica 7:* http://www.biotaneotropica.org.br/v7n2/pt/abstract?article+bn00607022007.

Editor: Jaime Bertoluci