

**O (DES)CONHECIMENTO DA BIODIVERSIDADE: UMA  
SISTEMATIZAÇÃO SOBRE LACUNAS, LIMITAÇÕES, VIESES,  
DÉFICITS E RUÍDOS**

Rodrigo Antônio Castro-Souza<sup>1,4\*</sup> , Geiziane Tessarolo<sup>6</sup> , Nicolas Silva Bosco<sup>1,4</sup> , Andréia de  
Carvalho Santos<sup>2</sup> , Afonso Kempner<sup>1,4</sup> , Thairik Mateus Silva Marques<sup>1,4</sup> , Vinícius Costa-  
Silva<sup>3,5</sup> , José Alexandre F. Diniz-Filho<sup>6</sup>  & Thadeu Sobral-Souza<sup>4</sup> 

<sup>1</sup> Universidade Federal de Mato Grosso, Programa de Pós-Graduação em Ecologia e Conservação da Biodiversidade, Rua Quarenta e Nove, 2367, Boa Esperança, CEP 78060-900, Cuiabá, MT, Brasil.

<sup>2</sup> Universidade Federal de Mato Grosso do Sul, Instituto de Biociências, Programa de Pós-Graduação em Ecologia e Conservação, Rua UFMS, s/n, Cidade Universitária, CEP 79070-900, Campo Grande, MS, Brasil.

<sup>3</sup> University of Pretoria, Department of Zoology & Entomology, Private Bag X20, Hatfield, 0028, South Africa.

<sup>4</sup> Universidade Federal de Mato Grosso, Instituto de Biociências, Departamento de Botânica e Ecologia, Laboratório de Macroecologia e Conservação da Biodiversidade, Av. Fernando Correa da Costa, 2367, Boa Esperança, CEP 78060-900, Cuiabá, MT, Brasil.

<sup>5</sup> Universidade Federal de Mato Grosso, Instituto de Biociências, Laboratório de Scarabaeoidologia, Av. Fernando Correa da Costa, 2367, Boa Esperança, CEP 78060-900, Cuiabá, MT, Brasil.

<sup>6</sup> Universidade Federal de Goiás, Instituto de Ciências Biológicas, Departamento de Ecologia, Avenida Esperança s/n, Câmpus Samambaia CEP 74690-900, Goiânia, GO, Brasil.

E-mails: rodrigodesouzaac@gmail.com (\*autor correspondente); geites@gmail.com; nicolas-bosco@hotmail.com; afonso.hk@gmail.com; andreaecologa@gmail.com; thairiksilvamarques@gmail.com; silvavinicius92@gmail.com; jafdinizflho@gmail.com; thadeusobral@gmail.com

**Resumo:** A partir dos avanços das ferramentas de bioinformática e ecoinformática tornou-se viável usar grandes quantidades de informação sobre biodiversidade para descrever padrões biológicos e tentar avaliar os processos ecológicos e evolutivos subjacentes a eles, bem como estabelecer ações mais efetivas e abrangentes para conservá-la. Contudo, as informações existentes acerca da biodiversidade ainda são incompletas, possuindo lacunas, limitações, vieses, déficits e ruídos. Neste sentido, reconhecemos a necessidade de discutir problemáticas relacionadas ao uso dos grandes bancos de dados de biodiversidade e a sua qualidade, discussões estas que precisam ser conceitualizadas de forma mais estruturada. Neste texto, iremos descrever de forma sistematizada as lacunas, limitações, e vieses do conhecimento da biodiversidade, problemas que estão diretamente associados à obtenção de informações das espécies na natureza. Em seguida, delimitaremos os tipos de déficits e ruídos, que são problemas associados aos bancos de dados da biodiversidade. Isso se deve ao fato de que as informações presentes nos registros primários e nos grandes bancos de dados frequentemente são usadas de forma inadequada. Além disso, apontamos também perspectivas sobre como abordar essas deficiências em trabalhos ecológicos/biogeográficos, por meio de uma adequada curadoria dos bancos de dados e da construção de mapas de ignorância. Por fim, destacamos desafios futuros, como a necessidade de novas coletas de dados, a formação de novos taxonomistas e o fornecimento de subsídios financeiros para a infraestrutura e a digitalização das coleções biológicas.

**Palavras-chaves:** déficits de conhecimento; lacunas da biodiversidade; limitações de conhecimento; ruídos do conhecimento; vieses de conhecimento.

**BIODIVERSITY (UN)KNOWLEDGE: A SYSTEMATIZATION OF GAPS, LIMITATIONS, DEFICITS, BIASES, AND NOISE:** With the advancements in bioinformatics and ecoinformatics tools, it has become feasible to utilize vast amounts of biodiversity information to describe biological patterns and attempt to assess the ecological and evolutionary processes underlying them, as well as to establish more effective and comprehensive conservation actions. However, existing information

regarding biodiversity remains incomplete, with gaps, limitations, biases, deficits, and noise. In this regard, we acknowledge the need to discuss issues related to the use of large biodiversity databases and their quality, discussions that need to be conceptually structured. In this text, we will systematically describe the gaps, limitations, and biases in biodiversity knowledge, issues that are directly associated with the acquisition of species information in nature. Subsequently, we will outline the types of deficits and noise, which are problems associated with biodiversity databases. This is because the information present in primary records and large databases is frequently used inadequately. Additionally, we also point out perspectives on how to address these deficiencies in ecological/biogeographical work through appropriate database curation and the construction of maps of ignorance. Finally, we highlight future challenges, such as the need for new data collections, the training of new taxonomists, and the provision of financial support for the infrastructure and digitization of biological collections.

**Keywords:** gaps of biodiversity; knowledge limitations; knowledge deficits; knowledge biases; knowledge noise.

---

## INTRODUÇÃO

Estudos biogeográficos e macroecológicos, bem como suas aplicações em conservação da biodiversidade, dependem diretamente do conhecimento sobre a biodiversidade que foi construído e armazenado por diversos cientistas ao longo do tempo. A fonte básica do conhecimento da biodiversidade é definida pelos registros primários, que são informações de quais espécies estão descritas e onde elas estão (ou estavam) no tempo e no espaço (Soberón & Peterson 2004). Além das informações taxonômicas, espaciais e temporais, muitos registros podem conter também informações genéticas e de traços funcionais (veja Cornwell *et al.* 2019) ou até mesmo de bioacústica. No entanto, focaremos aqui apenas na tríade taxonomia x espaço x tempo, visto que ela é o componente fundamental em praticamente todo estudo de escopo ecológico.

A obtenção dos registros primários ocorre, a priori, por meio de coletas ou amostragens das espécies (p.ex., visual, acústico ou por registros indiretos) em comunidades ecológicas onde elas habitam. Dependendo do habitat e da história natural de cada grupo-alvo estudado, diferentes métodos e horários de coleta são necessários (Papavero 1994, Alves *et al.* 2021, Appel *et al.* 2021). Após a coleta, os registros, individualizados ou em lotes (amostras), devem ser obrigatoriamente etiquetados com informações sobre o(s) organismo(s) coletado(s): localização (p.ex., país, estado, cidade e coordenadas geográficas), data de coleta, metodologia utilizada, bem como o nome do(s) coletor(es) (Almeida *et al.* 2012). A identificação taxonômica prévia do material é importante, porém não obrigatória nessa etapa do processo. Em seguida, os registros precisam ser armazenados de forma adequada em recipientes com líquido fixador (por exemplo, álcool ou formol) ou em caixas secas contendo repelentes (p.ex., cânfora ou naftalina) e enviados para laboratórios ou coleções biológicas, onde serão triados, montados e identificados (Papavero 1994).

Como toda a biodiversidade na Terra ainda não é conhecida em suas múltiplas facetas do conhecimento, nos últimos anos foram identificadas e evidenciadas diversas lacunas no conhecimento da diversidade biológica. Tais lacunas conhecidas como “*shortfalls*”, em inglês, chamadas a partir de agora de lacunas de conhecimento, referem-se à falta de conhecimento ou ao conhecimento desigual da biodiversidade no tempo e espaço, no que tange as diferentes facetas da biodiversidade (Hortal *et al.* 2015, Ficetola *et al.* 2019, Lee *et al.* 2020, Lopes-Lima 2021). Contudo, além das lacunas existem limitações que são obstáculos, restrições ou desafios que impedem o conhecimento completo da biodiversidade. Devido a existência das lacunas de conhecimento e limitações, vão existir também os vieses de conhecimento da biodiversidade, que se referem a um conhecimento desigual da biodiversidade em relação a sua totalidade.

Com o avanço de ferramentas de bioinformática e ecoinformática pode-se facilmente construir bancos de dados robustos com milhares de registros primários presentes em artigos e coleções biológicas que estão disponibilizados em plataformas digitais (Sousa-Baena *et al.* 2014), como por

exemplo GBIF (<https://www.gbif.org/>), OBIS (<https://obis.org/>) e iNaturalist (<https://www.inaturalist.org/>) que podem servir como base para preencher e/ou apontar as lacunas, as limitações e os vieses existentes no conhecimento da diversidade biológica. Mas com o advento dos bancos de dados, surgiram também novos problemas, como a falta de refinamento das informações dos registros, a partir de agora chamados de déficits do conhecimento, e os erros nas informações, a partir de agora chamados de ruídos do conhecimento.

Neste texto, iremos descrever de forma sistematizada as lacunas, limitações, e vieses do conhecimento da biodiversidade, problemas que estão diretamente associados à obtenção e coleta de informações das espécies na natureza. Seguidamente, delimitaremos os tipos de déficits e ruídos, que são problemas associados aos bancos de dados da biodiversidade (Figura 1). Isso se deve ao fato de que as informações presentes nos registros primários e nos grandes bancos de dados frequentemente são usadas de forma inadequada.



**Figura 1.** Sistematização das lacunas (1), limitações (2) e vieses (3) do conhecimento da biodiversidade, problemas diretamente associados à obtenção e coleta de dados. Em seguida, sistematizamos os déficits (4) e ruídos (5), problemas associados aos bancos de dados da biodiversidade.

**Figure 1.** Systematization of the gaps (1), limitations (2), and biases (3) in biodiversity knowledge, issues directly associated with data acquisition and collection. Next, we systematize the deficits (4) and noises (5), issues associated with biodiversity databases.

### *As Lacunas*

Atualmente existem dez lacunas de conhecimento da biodiversidade descritas na literatura, embora futuramente ainda possam ser descritas novas (veja Hortal *et al.* 2015, Ficetola *et al.* 2019, Lee *et al.* 2020, Lopes-Lima 2021), enquanto algumas já existentes podem vir a ser sinonimizadas. Elas surgiram devido à ausência de correspondência entre o que se espera ter de informação total na natureza e o que é/foi realmente coletado e revelado pela ciência. A lacuna mais fundamental de todas é a lacuna Linneana (LL), que representa nosso desconhecimento sobre a identidade taxonômica das espécies (Cardoso *et al.* 2011, Hortal *et al.* 2015). Estima-se que cerca de 86% das espécies terrestres e 91% das espécies marinhas sejam ainda desconhecidas para a ciência (Mora *et al.* 2011). Ainda assim, estima-se que alguns táxons, como vertebrados, por exemplo, possuem quase que a totalidade das suas espécies descritas. Por outro lado, alguns grupos, como artrópodes, ainda possuem uma pequena porção das suas espécies descritas, possuindo assim alta LL (Mora *et al.* 2011, Stork *et al.* 2018). A LL engloba também inúmeros novos morfotipos que estão depositados em coleções, porém, ainda não foram formalmente descritos (Hortal *et al.* 2015, Kemp 2015). Atualmente espécies crípticas estão sendo distinguidas em táxons distintos por meio de novas metodologias/critérios (p.ex., DNA barcoding, bioacústica e cariótipo) (veja discussão em Freeman & Pennell, 2021). A LL não existe só na natureza, mas também nos bancos de dados de registros primários, tendo em vista que nem todas as espécies atualmente descritas podem constar nessas plataformas.

Assumindo que espécies existam (veja discussão em Coyne & Orr 2004), a LL deve ser considerada a mais basal, sendo a primeira fonte de (des)conhecimento da biodiversidade. As demais lacunas são, portanto, subordinadas a essa lacuna, já que dependem do conhecimento taxonômico (descrição formal de uma espécie). Em suma, não se pode discutir padrões espaciais da biodiversidade se não sabemos com quais espécies estamos lidando. Embora a LL se demonstre importante, sabe-se

hoje que o conhecimento espacial da biodiversidade, baseado na taxonomia das espécies descritas atualmente, é enviesado pela acessibilidade - locais próximos a cidades, estradas e rios são mais amostrados (Meyer *et al.* 2015, Zizka *et al.* 2020). Questões socioculturais também são direcionadores de viés de amostragem, tendo em vista que existem táxons e regiões mais “carismáticas” a serem estudados. Aves, por exemplo, é o táxon mais estudado quando comparado com outras Classes de vertebrados e invertebrados (Troudet *et al.* 2017). Por outro lado, fragmentos naturais maiores e mais conectados entre si são mais amostrados que outros dentro da Mata Atlântica (Sobral-Souza *et al.* 2021).

A lacuna Wallaceana (LW) representa o (des)conhecimento acerca da distribuição espacial das espécies (Cardoso *et al.* 2011, Hortal *et al.* 2015), que depende a priori do conhecimento taxonômico (LL). Para a maior parte das espécies já descritas, existem poucas ou, se não uma única localidade conhecida (localidade tipo), o que torna o conhecimento espacial da maioria das espécies tendencioso e ao mesmo tempo incerto para inúmeras regiões (Cardoso *et al.* 2011). Adicionalmente, considerando os dados de cobertura de distribuição para as principais espécies de animais terrestres e marinhos, sabemos que menos de 7% da biota terrestre e apenas 5% da biota marinha possuem informações sobre a distribuição espacial. Essas informações se concentram principalmente próximas a fatores de acessibilidade, como estradas, zonas costeiras e rotas marítimas (Hughes *et al.* 2021). Tal padrão ocorre muitas vezes devido à grande parte do conhecimento wallaceano estar pouco refinado, onde podemos ter uma quantidade substancial de registros primários com informações de localidade incertas, como por exemplo, a nível de município, estado ou província, e até mesmo país, que são desconsiderados nos processos de filtragem que utilizam apenas dados com informações seguras de coordenadas geográficas (veja Meyer *et al.* 2016, Ronquillo *et al.* 2023).

A lacuna Prestoniana (LP) representa o (des)conhecimento sobre a dinâmica populacional das espécies ao longo do espaço e do tempo (Cardoso *et al.* 2011). Da mesma forma, não se pode inferir se não sabemos quais são as espécies (LL). Para acessar tal lacuna é necessário um método de

coleta (ou observação) representativo para o grupo. Além disso as informações de abundância amostradas em um determinado local muitas vezes podem ser perdidas devido o mal aproveitamento dos dados catalogados (Cardoso *et al.* 2011) e escala de amostragem (p.ex., no caso de metapopulações e populações de distribuição irregular) (Freckleton & Watkinson 2002). Adicionalmente, vale ressaltar que a dinâmica populacional das espécies, por vezes, depende de metodologia que visa a não amostragem de espécimes, em processos de marcação e recaptura, por exemplo (Sobral-Souza *et al.* 2015). Dessa forma, esses dados muitas vezes não podem ser acessados várias vezes, em diferentes períodos de tempo por diferentes pesquisadores. Somados a isso, a dinâmica populacional é um processo local que varia ao longo das subpopulações das espécies, aumentando ainda mais a LP.

As lacunas Darwiniana (LD) e Raunkiaeriana (LR) refletem nosso desconhecimento da evolução das espécies (Diniz-Filho *et al.* 2013) e das características funcionais das espécies e suas funções ecológicas (Hortal *et al.* 2015), respectivamente. Estas lacunas também podem ser consideradas de forma subordinada à LL, e também a LW, visto que se desconhecemos as espécies em um determinado espaço geográfico, não podemos inferir relações evolutivas (posicionamento filogenético) e nem quais as características fenotípicas relacionadas as suas funcionalidades evoluíram em um dado local (embora a proposição original da LD tenha sido feita em um contexto de substituir medidas de diversidade a partir da riqueza de espécies por métricas filogenéticas, que seriam mais robustas à incertezas taxonômicas; Diniz-Filho *et al.* 2013). De modo antagônico, a falta de conhecimento a respeito da história evolutiva e características dos grupos (LD) também prejudica a compreensão da distribuição espacial atual das espécies, estando assim, diretamente relacionada com a LW (Hortal *et al.* 2015, Rudbeck *et al.* 2022, Diniz-Filho *et al.* 2023). Isso ressalta que as lacunas do conhecimento da biodiversidade são influenciadas não somente por escalas ecológicas, mas também por escalas de dimensões evolutivas (Whittaker *et al.* 2005). Contudo, mesmo existindo essas discussões, novas iniciativas biogeográficas estão surgindo como forma de interligar o

conhecimento entre LL, LD e LW, gerando conhecimento sobre padrões biogeográficos (veja Diniz-Filho *et al.* 2023).

Além das lacunas preditas, temos as lacunas Hutchinsoniana (LH) e Eltoniana (LE), que indicam nosso desconhecimento sobre a tolerância ambiental das espécies às condições abióticas e o desconhecimento das interações biológicas, respectivamente (Hortal *et al.* 2015, mas veja também Rosado *et al.* (2016)). A compreensão da LH pode auxiliar na compreensão de outras lacunas, pois algumas dinâmicas ambientais podem ser capazes de regular o tamanho das populações e interações das espécies (veja a discussão em Hortal *et al.* 2015). Em escalas amplas, a LE pode ser considerada um ruído para a construção de modelos correlativos de nicho (veja Soberón & Nakamura 2009). No entanto, o uso de traços funcionais tem demonstrado que a compreensão dos padrões locais pode ser determinante nos padrões macroecológicos e biogeográficos (Rosado *et al.* 2016). Por fim, também temos as lacunas Parkerianas (LPa), Ostromianas (LO) e Racovitzanas (LRc), formalizadas mais recentemente, que ressaltam, respectivamente, nosso (des)conhecimento sobre a história natural, a eficácia das medidas de conservação e a exploração e o mapeamento das espécies em todos os ambientes terrestres e aquáticos (Ficetola *et al.* 2019, Lee *et al.* 2020, Lopes-Lima 2021) (Tabela 1).

**Tabela 1.** Síntese sobre as principais lacunas do conhecimento da biodiversidade (referências marcadas com (\*) são as originais de cada lacuna).

**Table 1.** Synthesis on the main knowledge gaps of biodiversity (references marked with (\*) are the original sources of each shortfall).

LACUNA	FACETA DA BIODIVERSIDADE	BREVE DEFINIÇÃO	REFERÊNCIAS SUGERIDAS
Linneana	Espécies	Desconhecimento da identidade taxonômica das espécies;	(Raven & Wilson 1992*, Cardoso <i>et al.</i> 2011, Hortal <i>et al.</i> 2015)
Wallaceana	Distribuição geográfica	Desconhecimento da distribuição geográfica das espécies;	(Lomolino 2004*, Cardoso <i>et al.</i> 2011, Hortal <i>et al.</i> 2015)
Prestoniana	Populações	Desconhecimento da abundância e dinâmica populacional das espécies ao longo do espaço e do tempo;	(Cardoso <i>et al.</i> 2011*, Hortal <i>et al.</i> 2015)
Darwiniana	Evolução	Desconhecimento sobre a evolução das espécies;	(Diniz-Filho <i>et al.</i> 2013*, Hortal <i>et al.</i> 2015)

LACUNA	FACETA DA BIODIVERSIDADE	BREVE DEFINIÇÃO	REFERÊNCIAS SUGERIDAS
Raunkiaeriana	Características funcionais e funções ecológicas	Desconhecimento sobre as características funcionais das espécies e de suas funções ecológicas;	(Hortal <i>et al.</i> 2015*, mas veja Rosado <i>et al.</i> 2016)
Hutchinsoniana	Tolerâncias Abióticas	Desconhecimento sobre às tolerâncias ambientais das espécies à condições abióticas;	(Cardoso <i>et al.</i> 2011*, Hortal <i>et al.</i> 2015)
Eltoniana	Interações Ecológicas	Desconhecimento das interações biológicas das espécies e quais seus efeitos sobre a sobrevivência e o fitness;	(Hortal <i>et al.</i> 2015*, mas veja Rosado <i>et al.</i> 2016)
Parkeriana	História Natural	Desconhecimento sobre história natural das espécies;	(Lee <i>et al.</i> 2020*)
Ostromiana	Medidas de Conservação	Desconhecimento sobre a eficácia das medidas de conservação;	(Lopes-Lima 2021*)
Racovitzana	Exploração geográfica	Desconhecimento da biodiversidade devido a áreas não exploradas e não mapeadas;	(Ficetola <i>et al.</i> 2019*)

As lacunas no conhecimento da biodiversidade se tornaram um dos grandes desafios para as pesquisas biológicas em todo o globo, visto que impactam diretamente a interpretação e qualidade dos padrões de biodiversidade em diversas escalas ecológicas, em instâncias como a precisão, generalização e realismo dos dados (Hortal *et al.* 2015). Estes desafios podem ser ainda maiores quando considerados os efeitos hierárquicos anteriormente mencionados, uma vez que, a depender das áreas (configuração da paisagem), períodos e escalas estudadas, as lacunas do conhecimento da biodiversidade podem ter influência umas sobre as outras, correlacionando-se de forma sinérgica (Diniz-Filho *et al.* 2023).

#### *As Limitações*

As limitações são obstáculos, restrições e/ou desafios que impedem o conhecimento completo da biodiversidade. As limitações podem ser responsáveis pela formação das lacunas de conhecimento. Aqui, vamos listar as cinco principais limitações, embora existam outras, como a falta de fotografias das espécies na natureza (Mesaglio *et al.* 2023) ou até mesmo a dificuldade de acessar a série tipo das

espécies depositadas em coleções estrangeiras por pesquisadores das regiões tropicais (Nakamura *et al.* em preparação).

A primeira refere-se às limitações de exploração, que ressalta a falta de inventários da biodiversidade em alguns ambientes naturais, resultando no desconhecimento quase que total da biodiversidade em muitos lugares do planeta (p.ex., cavernas, fundo dos oceanos, manchas de florestas e desertos) (Ficetola *et al.* 2019). Certamente existem ambientes mais “atraentes” (florestas tropicais, recifes de corais) que são mais amostrados, seja pela sua beleza cênica ou pela elevada biodiversidade. Por outro lado, ambientes menos “atraentes”, como brejos, alagados e mangues, são menos amostrados sendo, portanto, menos conhecidos.

Outra limitação é a detectabilidade. Muitas espécies, e/ou assembleias, são menos detectáveis do que outras fazendo com que algumas espécies sejam subamostradas, como, por exemplo, espécies raras, com abundância reduzida e/ou espécies migratórias (Ladle & Hortal 2013). Somados a isso, existem diferentes estratégias e finalidades na forma de inventariar espécies em ambientes naturais (Ahrends *et al.* 2011). As técnicas utilizadas para amostragens da biodiversidade podem ter atratividade específica para alguns táxons, em detrimento de outros que estão no mesmo ambiente. Isso resulta na hiper amostragem de alguns táxons como resultado da técnica de amostragem específica e não pela realidade das abundâncias relativas das espécies na comunidade local (Bach *et al.* 2023).

As limitações de similaridade ambiental também são inerentes ao conhecimento da biodiversidade. O conhecimento da distribuição das espécies diminui com o aumento da distância (espacial e/ou climática) em relação à ocorrência conhecida da espécie. Isso acontece devido à existência de uma variabilidade na composição de espécies ao longo do espaço geográfico e ambiental, como, por exemplo, em função de gradientes de habitat, clima e vegetação (Ladle & Hortal *et al.* 2013, Tessarolo *et al.* 2021). Isso deixa claro que as limitações de similaridade ambiental e a limitação de exploração são intimamente relacionadas. Embora isso seja bastante óbvio, é necessário

considerar que, dependendo do tipo de extrapolação feita com os dados existentes e da falta de registros primários, o padrão real de agregação de ocorrências próximo ao local de amostragem (baixa heterogeneidade de registros e alta correlação espacial), tem sido desconsiderado na maioria dos estudos como no caso do uso de modelos de distribuição de espécies (SDMs - veja a problematização levantada em Lobo *et al.* 2018) ou construídos a partir de ocorrência enviesadas no espaço geográfico e ambiental (veja Gutierrez-Velez & Wiese 2020, Sobral-Souza *et al.* 2021), podendo resultar em interpretações irrealísticas de padrões e processos.

As limitações temporais dos bancos de dados podem ser explicadas pelo dinamismo natural (p.ex., distúrbios estocásticos, extinções e invasões biológicas) e antrópico (p.ex., perda de habitat, poluição e mudanças climáticas) nos sistemas naturais que altera ao longo do tempo a distribuição geográfica dos registros primários, o que acontece devido a diversos processos em diferentes escalas ecológicas (Escribano *et al.* 2016, Tessarolo *et al.* 2017). Um registro originalmente descrito para um ambiente florestal, por exemplo, pode não ser mais válido atualmente por conta da perda do habitat florestal daquela localidade após a amostragem da espécie. Outra fonte de limitação temporal é que a nomenclatura taxonômica das espécies pode sofrer mudanças devido a novos atos nomenclaturais e novas classificações taxonômicas que ocorreram após o registro. Isso gera históricos taxonômicos complexos como sinonímias e homonímias que não podem ser desconsiderados na curadoria dos bancos de dados (Tessarolo *et al.* 2017, Stropp *et al.* 2022). Além disso, com o tempo, também pode ocorrer a perda dos metadados e das informações básicas sobre as coletas de dados, além do intemperismo nos registros armazenados em coleções biológicas. Tais perdas temporais ocasionam o desaparecimento de informações vitais sobre os registros primários, como a data de coleta, coordenadas, identificação do autor e do coletor, ou até mesmo partes dos espécimes depositados (Tessarolo *et al.* 2017). Muitas vezes isso pode depender mais das condições de armazenamento dos dados do que do tempo (Greshko 2018). Todas as limitações temporais aqui listadas tornam as informações conhecidas em um banco de dados obsoletas ao longo do tempo, afetando diretamente o

conhecimento da distribuição das entidades taxonômicas (Tessarolo *et al.* 2017; Stropp *et al.* 2022), bem como o estudo e acesso aos espécimes físicos estudados pelos taxonomistas (Tessarolo *et al.* 2017). Isso resulta diretamente na perda de informações preciosas sobre a biodiversidade, transformando informações conhecidas em desconhecidas.

A última limitação refere-se à limitação de disponibilidade. As coleções biológicas de diferentes partes do mundo ainda não são de domínio público e/ou não estão digitalizadas e disponíveis em repositórios digitalmente acessíveis (Krishtalka & Humphrey 2000, Nelson & Ellis 2018, Bakker *et al.* 2020). Além disso, muitos dos dados disponíveis na literatura taxonômica científica e literatura cinza (dados científicos não publicados, incluindo dissertações, teses e relatórios) também não estão digitalizados e disponíveis (Heidorn 2008, Sikes *et al.* 2016). A falta das grandes coleções biológicas nos repositórios digitais é creditada, principalmente, à falta de mão de obra especializada (e remunerada) para o preenchimento do banco de dados, além da quantidade exorbitante de material ainda a ser digitalizado. Muitas das espécies atualmente descritas pela ciência podem também não possuir registros de sua descrição e/ou distribuição disponíveis digitalmente, uma deficiência que normalmente ocorre nos repositórios digitais da biodiversidade (Figura 2).



**Figura 2.** Síntese sobre as Limitações, Déficits, Vieses e Ruídos do conhecimento da biodiversidade (baseado em Scott & Hallam 2003, Hey *et al.* 2003, Murphey *et al.* 2004, Hortal *et al.* 2008, Boakes *et al.* 2010, Ahrends *et al.* 2011, Rocchini *et al.* 2011, Ledle & Hortal 2013, Meyer *et al.* 2015, Escribano *et al.* 2016, Meyer *et al.* 2016, Tessarolo *et al.* 2017, Troudet *et al.* 2017, Ficetola *et al.* 2019, Siracusa *et al.* 2020, Hughes *et al.* 2021, Tessarolo *et al.* 2021).

**Figure 2.** *Synthesis on the limitations, gaps, biases, and noises of biodiversity knowledge (based on Scott & Hallam 2003, Hey et al. 2003, Murphey et al. 2004, Hortal et al. 2008, Boakes et al. 2010, Ahrends et al. 2011, Rocchini et al. 2011, Ledle & Hortal 2013, Meyer et al. 2015, Escribano et al. 2016, Meyer et al. 2016, Tessarolo et al. 2017, Troudet et al. 2017, Ficitola et al. 2019, Siracusa et al. 2020, Hughes et al. 2021, Tessarolo et al. 2021).*

Embora o grau de impacto dessas limitações possa variar dependendo do tipo de estudo e enfoque da pesquisa, a maioria dos levantamentos de dados de biodiversidade vão sofrer com algumas dessas limitações. Assim, mesmo bancos de dados com milhões de registros primários, provenientes de diversos estudos, vão herdar as limitações inerentes aos dados originais, e, portanto, sofrer das mesmas limitações em maior ou menor grau, o que pode resultar na impossibilidade de responder algumas perguntas biogeográficas (Sobral-Souza *et al.* 2021, Ronquilo *et al.* 2023). Por exemplo, a falta de informações sobre a data de coleta de determinados registros pode impossibilitar qualquer interpretação associada à cobertura temporal. Da mesma forma, a inclusão apenas de registros que possuem espécimes depositados em coleções biológicas pode negligenciar o papel da ciência cidadã na construção do conhecimento ao longo do espaço e do tempo.

### *Os Vieses*

As limitações dos dados geram as lacunas no conhecimento. Por sua vez, as lacunas podem se propagar para os bancos de dados em forma de vieses. Aqui, definimos vieses como o conhecimento desigual de parte dos dados em relação à totalidade dos dados existentes. Os vieses podem ocorrer na hiper ou sub amostragem de táxons, regiões geográficas e períodos temporais estudados. Os vieses normalmente são resultados da heterogeneidade no esforço dos inventários de coletas das espécies, compilação e estudo dos registros primários ao longo do planeta (Hughes *et al.* 2021), e/ou também em função do padrão heterogêneo de distribuição da biodiversidade. Muitos estudos têm demonstrado que fatores de acessibilidade - locais próximos de vias de acesso (estradas e rios), centros de pesquisa, unidades de conservação e fragmentos florestais grandes e conectados são mais amostrados, gerando um conhecimento enviesado para estes locais (Meyer *et al.* 2015, Zizka *et al.* 2020, Sobral-Souza *et*

*al.* 2021). Além disso, a acessibilidade e construção do conhecimento da biodiversidade são creditados também ao colonialismo científico ocorrente em cada país, níveis de riqueza, educação e estabilidade política (veja Boakes *et al.* 2010; Meyer *et al.* 2015; Raja *et al.* 2022).

Historicamente as coleções biológicas dos países europeus e da América do Norte se originaram e cresceram conforme as nações imperiais estenderam seus domínios coloniais e adquiriram uma variedade de recursos de suas colônias, o que inclui direta e indiretamente a coleta de exemplares de animais e plantas (Zizka *et al.* 2021). Tal fato é conhecido como colonialismo científico clássico. Na modernidade, esse colonialismo ocorre na forma de uma ciência de "para-quebras" onde pesquisadores de países desenvolvidos e estruturados aparecem em países menos desenvolvidos para coletar dados e desenvolver suas pesquisas, sem trazer retorno científico e tecnológico para a comunidade local. Além disso, o material biológico coletado, em muitos casos, é depositado fora de seu país de origem (Raja *et al.* 2022). Nesse sentido, como os períodos de colonização e expedições foram diferentes ao longo do planeta, o colonialismo científico foi/é mais forte em algumas regiões, gerando vieses espaciais, taxonômicos e temporais (Zizka *et al.* 2021). Além disso, a disponibilização dos registros primários, mesmo estando depositados em museus, pode ocorrer de forma diferente devido a questões ligadas à macroeconomia (Meyer *et al.* 2015) e sigilo nos dados.

Os vieses podem ser divididos em pelo menos três categorias. Os vieses taxonômicos indicam que alguns táxons são mais estudados do que outros em consequência de maior financiamento, à melhor facilidade e rapidez de amostragem, maior número de especialistas, maior diversidade de espécies dentro dos grupos taxonômicos (Troudet *et al.* 2017, Siracusa *et al.* 2020) e/ou maior necessidade de conhecimentos sobre certos táxons visando políticas públicas para conservação e/ou de serviços ecossistêmicos essenciais para a manutenção do bem-estar humano. Os vieses espaciais, por sua vez, são relacionados a amostragem da biodiversidade de forma desigual ao longo do espaço geográfico. Isso resulta em lugares mais amostrados do que outros, devido a fatores como

acessibilidade, condições socioeconômicas, de infraestrutura e de condições paisagísticas (Hortal *et al.* 2008, Meyer *et al.* 2015, Meyer *et al.* 2016, Sobral-Souza *et al.* 2021). Os vieses temporais descrevem que existem determinados períodos temporais que são mais estudados do que outros, devido questões ligadas ao esforço de amostragem, sazonalidade climática, grupos taxonômicos estudados e ao colonialismo científico (clássico e moderno) (Boakes *et al.* 2010, Stropp *et al.* 2016, Meyer *et al.* 2016, Ronquillo *et al.* 2020) (Figura 2).

### *Os Déficits*

Os déficits - a diferença entre o que é esperado em termos de conhecimento da biodiversidade e o que realmente está disponível nos dados - podem estar presentes nas identificações taxonômicas, nas informações de localização espacial e temporal dos registros primários, bem como em todos os outros tipos de informações além da tríade taxonomia, espaço e tempo (p.ex., informações acústicas, de traço funcional ou genéticas).

Embora existam alguns métodos para descrever o déficit taxonômico dentro dos inventários (LW) (Soberón *et al.* 2007) e acerca da quantidade de espécies/gênero/famílias a serem descobertos (LL) (Mora *et al.* 2011, Freitas *et al.* 2021), um dos mais simples é representar a proporção de identificações em níveis taxonômicos supraespecífico (gênero, família e Ordem) em relação a todos os registros existentes (Castro-Souza *et al.* submetido). Quanto maior a proporção de registros supraespecífico no banco de dados, maior é o déficit taxonômico. O ideal é que o banco de dados contenha todos os registros à nível específico. Contudo, a exclusão dos dados supraespecíficos impossibilita o cálculo do déficit taxonômico e negligencia a incerteza taxonômica dos dados e dos padrões de biodiversidade. Tal problemática nasce dentro do sistema de classificação linneano que possui níveis mais “refinados” (p.ex., categorias de espécies e subespécies), enquanto outros são mais “amplos” (p.ex., categorias de Ordens, Classes e Filos). Para buscar padrões biogeográficos e

estratégias de conservação biogeográfica, muitas vezes somos indagados a resolver questões que utilizam apenas partículas de conhecimento mais “refinadas” (Hortal *et al.* 2015).

Os déficits espaciais podem ser representados pelos registros que possuem indicação equivocadas, ou faltantes, sobre o local amostrado. Consequentemente o déficit espacial não permite apontar com acurácia geográfica onde o exemplar foi obtido. Em muitos casos os registros possuem localidades descritivas em níveis de município, estado ou país, ou informações ausentes (Rocchini *et al.* 2011, Ronquillo *et al.* 2020). Alguns outros, sendo representados principalmente por registros antigos, colocam pontes, comunidades tradicionais e fazendas como locais de coleta. Contudo, muitos locais não existem mais nos dias atuais, sendo necessário o estudo de documentos históricos, como o Guia Postal-Telegráfico do país, compilados sobre localidades descritas na literatura científica (veja Stephens & Traylor 1983) ou trabalhos que abordem a vida e história dos grandes coletores (veja Papavero 1971, 1973, Lamas 1980, Horn *et al.* 1990a, b, Evenhuis 1997). Essas deficiências nos dados primários podem muitas vezes resultar na não utilização do registro nas análises biogeográficas, transformando dados já escassos em mais escassos ainda.

Por fim, os déficits temporais podem ser representados pelos dados que contém informações incompletas a respeito de quando um espécime foi registrado, não permitindo identificar a data de sua coleta (Figura 2), ou até mesmo alguns dados mais refinados, como horário da observação (p.ex., estudos de horário de atividade e forrageamento) e/ou estações em que os dados foram coletados (p.ex., seca e chuvosa). Embora aqui nós tenhamos separados os déficits das lacunas, limitações e vieses, eles estão interligados e devem ser estudados e sumarizados de forma indissociável e não isolada.

### *Os Ruídos*

Além das problemáticas mencionadas, existem ainda ruídos que podem estar presentes nos bancos de dados que sintetizam o conhecimento da biodiversidade. Ruídos são os “erros” que afetam

a qualidade taxonômica, espacial e temporal dos dados e que podem surgir em diferentes fases da curadoria, digitalização ou coleta dos dados (Meyer *et al.* 2016). O ruído taxonômico se apresenta a partir da existência de nomes das espécies incorretos devido a rearranjos taxonômicos que ocorreram no passado e não foram atualizados nos registros ou bases de dados (Hey *et al.* 2003, Meyer *et al.* 2016). Os erros de digitação e digitalização ao longo da construção do banco de dados também podem fazer com que novas entidades taxonômicas, não existentes no mundo real, sejam criadas (mesmo não reconhecidas pelo código de nomenclatura).

O ruído de identificação surge quando as identificações taxonômicas não foram feitas por especialistas do grupo e, conseqüentemente, estão equivocadas (Scott & Hallam 2003, Meyer *et al.* 2016). Esse problema é complexo, uma vez que existem poucos especialistas para a maioria dos táxons e muitas das identificações em bancos de dados podem não conter a autoria de quem realizou ou revisou as identificações (Meyer *et al.* 2016). Além disso, as identificações conduzidas por especialistas podem estar erradas, considerando que o especialista pode estar iniciando sua carreira, ou no caso de serem espécies crípticas. Nesse sentido, iniciativas como o iNaturalist representam uma possível solução, pois permitem conhecer o grau de confiança da identificação do registro através de um consenso da sua comunidade de especialistas, naturalistas e fotógrafos. Desta forma, iniciativas como esta poderiam ser adaptadas e incorporadas em outras fontes ou até mesmo nas coleções biológicas, desde que possuam imagens ou alguma evidência da espécie acessível por outros pesquisadores, como é o caso de sinais acústicos e sequências genéticas de DNA mitocondrial. No caso das coleções, mais de um especialista poderia identificar a mesma espécie, em busca de um consenso.

O ruído de coordenadas ocorre quando as coordenadas geográficas e as localidades descritivas dos registros foram medidas ou denominadas de forma incorreta, devido a problemas com o uso de aparelhos de registros, como GPS e celulares, e principalmente devido ao uso incorreto de *datum* e projeção geográfica, bem como descrições errôneas sobre o local onde as espécies foram observadas

(Murphey *et al.* 2004, Meyer *et al.* 2016), ou até mesmo transcritas de forma errônea das etiquetas originais. Esse ruído pode ser mais comum em dados antigos, onde não existiam aparelhos de GPS para determinar a localização exata da coleta. Além disso, muitos aparelhos de GPS podem possuir uma “*Selective Availability*” (SA) para degradar intencionalmente a precisão do sinal para usuários civis (veja Pace *et al.* 1995). Por fim, existe também o ruído de data, que acontece quando as datas de coleta estão incorretas, seja por terem sido marcadas e/ou transcritas de forma errônea, ou até mesmo pela imprecisão fornecida pelo coletor (Figura 2).

Além disso, os déficits e ruídos supracitados podem ser responsáveis pela formação de muitos vieses, tendo em vista que ao descartar/filtrar dados problemáticos, podemos estar favorecendo determinados táxons, localidades, períodos temporais e/ou até autores específicos (veja Ronquillo *et al.* 2023).

## DISCUSSÃO

### *Desafiando a Ignorância do Conhecimento*

Se as lacunas, limitações, vieses, déficits e ruídos fossem desconsiderados, estaríamos assumindo que toda a biodiversidade na Terra foi amostrada e compreendida e as informações contidas nos bancos de dados estão corretas. Dificilmente isso é verdade. Portanto, destacamos que a chave é estar sempre atento a pelo menos dois tipos de problemas inerentes aos bancos de dados: a curadoria dos bancos de dados e a ignorância existente no conhecimento que eles representam. Levá-los em consideração ao avaliar quaisquer padrões ecológicos e/ou realizar análises fundamentadas no conhecimento digitalmente acessível é importante (Figura 3).



**Figura 3.** Perguntas que podem ajudar a identificar incompletudes do conhecimento taxonômico, espacial e temporal associadas aos registros primários, considerando o momento de sua descoberta, formação de vieses, ruídos e déficits. Aqui, não consideramos as lacunas e limitações, tendo em vista que elas são intrínsecas ao conhecimento da biodiversidade.

**Figure 3.** Questions that can help identify incompleteness in taxonomic, spatial, and temporal knowledge associated with primary records, considering the moment of their discovery, formation of biases, noises, and deficits. Here, we do not consider the gaps and limitations, given that they are intrinsic to biodiversity knowledge.

A curadoria do banco de dados é um processo demorado, complexo e extremamente essencial. Saber quais dados serão filtrados (excluídos) e quais permanecerão nas análises é fundamental. Contudo, excluir dados primários pode aumentar os vieses e lacunas; mantê-los pode aumentar os ruídos e déficits. A decisão dos processos de filtragem dos registros primários é diretamente dependente dos objetivos e hipóteses do trabalho que utilizará tais dados. Logo, o mesmo banco de dados original (pré-filtragem) poderá ser usado de forma diferente, com diferentes filtragens, em

diferentes trabalhos. Sugerimos a confecção de metadados da tabela final (filtrada) que contenha os processos detalhados (com as explicações) das filtrações realizadas para que se saiba as decisões tomadas ao longo da curadoria dos dados (veja por exemplo, Ronquillo *et al.* 2020). Indicamos também que cada filtro executado dê origem a nova tabela. Assim, todos os passos ficarão registrados, permitindo a utilização do mesmo banco de dados por diferentes pesquisadores.

Uma solução proposta para compreender a ignorância existente no conhecimento e integrar as incertezas das informações dos registros primários nos trabalhos é construir mapas de ignorância (Boggs 1949, Rocchini *et al.* 2011, Ladle & Hortal 2013, Tessarolo *et al.* 2021). Esses mapas possibilitam a compreensão do que realmente sabemos e o que de fato não sabemos sobre o conhecimento da biodiversidade (Boggs 1949). Como na maioria das vezes os dados de biodiversidade apresentam algum tipo de inconsistência (Meyer *et al.* 2016), mapeá-las pode ser a base de um conhecimento mais sólido nos dias de hoje. Esses mapas podem também basear decisões acerca de investimentos para inventários e quais os táxons que devem ser priorizados. Entender os vieses taxonômicos e espaciais, em conjunto, por exemplo, pode salientar quais regiões geográficas e quais táxons estão superamostrados e quais possuem regiões e táxons subamostrados. Isso permite ordenar e priorizar locais e táxons para amostragens futuras (Sobral-Souza *et al.* 2021).

Existe uma variedade de métodos que nos permitem compreender parte da ignorância dos dados da biodiversidade, tanto em termos taxonômicos, espaciais e/ou temporais (veja Soberón *et al.* 2007, Sousa-Baena *et al.* 2014, Stropp *et al.* 2016, Meyer *et al.* 2015, Meyer *et al.* 2016, Lobo *et al.* 2018, Cornwell *et al.* 2019, Zizka *et al.* 2020, Freitas *et al.* 2021, Tessarolo *et al.* 2021). Nesse sentido, podemos delimitar nossa problemática científica, como, por exemplo, qual lacuna, qual limitação, qual viés, qual déficit, ou qual ruído da biodiversidade ou biogeografia desejamos investigar. Compilamos todo o conhecimento existente (registros primários) e realizamos uma curadoria nestes dados (filtração) para poder construir mapas que permitirão representar essas diferentes questões, seja de forma isolada ou combinada em um único mapa de ignorância (veja por exemplo, Boggs 1949,

Meyer *et al.* 2016, Stropp *et al.* 2016, Tessarolo *et al.* 2021). Desta forma, compreender, mapear e inserir a ignorância existente nos bancos de dados que sumarizam o conhecimento da biodiversidade dentro de pesquisas ecológicas pode ser algo extremamente valioso, contribuindo para o uso crítico dos bancos de dados e a realização de pesquisas ainda mais sólidas que considerem as incompletudes associadas aos dados.

### ***Desafios Futuros***

Uma coleta de dados de biodiversidade eficiente, que permita diminuir a(s) lacuna(s), é um dos desafios que surgem. Para isso, torna-se necessário o uso de protocolos de coleta sistematizados (p.ex., abrangendo aspectos taxonômicos, espaciais e temporais) que podem variar conforme o grupo estudado. Nesse sentido, apenas o uso do conhecimento digital acessível (Sousa-Baena *et al.* 2014) pode não ser suficiente para o(s) grupo(s) ou região considerada(s) como modelo, tendo em vista que grande parte das coleções biológicas de diferentes partes do mundo não são de domínio público e/ou não estão digitalizadas e disponíveis online (Krishtalka & Humphrey 2000; Nelson & Ellis 2018; Bakker *et al.* 2020). Adicionalmente, os dados disponíveis na literatura taxonômica científica e literatura cinza não necessariamente significam que toda distribuição relatada nessas fontes esteja digitalizada em bancos de dados online (veja Heidorn 2008; Sikes *et al.* 2016). No entanto, como a literatura cinza não passa por uma revisão por pares, é comum que as listas de espécies contenham identificações equivocadas, tenham utilizado metodologias inadequadas para coleta dos dados e/ou tenham sido conduzidas por um profissional inexperiente para identificar espécies. Nesse contexto, a inclusão de dados da literatura cinza poderia trazer mais limitações e ruídos do que soluções, embora possa existir casos que contenha informações de boa qualidade. Assim, torna-se necessária a realização de novas coletas de dados primários em campo ou visitas a coleções biológicas ou acervos de materiais ainda não digitalizados. Isso pode, conseqüentemente, requerer financiamento de projeto, e dependendo do país ou instituição, pode tornar-se inviável.

Também é crucial conseguir compreender de forma precisa “*que espécie é essa?*”. O taxonomista é o principal profissional responsável por responder essa questão, a partir da descrição, nomeação e elaboração de material científico necessários para identificação de cada uma das unidades taxonômicas (nossos registros primários). Como exemplo prático, a taxonomia desempenha um papel fundamental na identificação de pragas agrícolas e espécies exóticas, evitando gastos e esforços inúteis. Além disso, é crucial na medicina, onde a identificação correta de organismos, como larvas de moscas, é vital para tratamentos utilizados na chamada terapia larval (Masiero *et al.* 2015). Ainda, a correta identificação de um organismo, como por exemplo um inseto, pode auxiliar os estudos periciais de casos envolvendo morte violenta, onde a partir da biologia do inseto estudado, pode-se determinar o intervalo pós-morte (Benecke 2001, Löbl *et al.* 2023).

No entanto, mesmo com sua notória importância no campo das ciências naturais, a taxonomia tem sido historicamente negligenciada, contando com um número cada vez menor de taxonomistas no mercado de trabalho (Wheeler 2014, Löbl *et al.* 2023), baixo impacto de suas publicações (Krell 2012), baixa perspectivas de emprego (Disney 1998, Jaspars 1998, Pearson *et al.* 2011, Salvador *et al.* 2022) e estímulos econômicos pífios para formação de novos taxônomos (Wheeler 1995, Disney 1998, Bortolus 2008, Werner 2009, Boero 2010, Salvador *et al.* 2022) frente as áreas de tecnologia e inovação.

A formação de um taxonomista, assim como de outros profissionais que atuam na biodiversidade, demanda tempo e dedicação, além de ocorrer em uma velocidade muito menor do que a capacidade dos problemas a serem resolvidos (Fontaine *et al.* 2012). Não que em outras áreas seja diferente (p.ex., genética molecular, estatística e ecoinformática), contudo, no caso da taxonomia, muitas vezes esse processo não conta com uma estrutura tecnológica mais avançada que pode ajudar a poupar um pouco de tempo. Essa incompatibilidade entre a velocidade que a sociedade contemporânea quer para as respostas e o tempo que as evidências existentes na biodiversidade demoram para aparecer, tornam essa formação cada vez mais desafiadora. Em um contexto de

taxonomia integrativa (Schlick-Steiner 2010), os taxonomistas “modernos” necessitam, muitas vezes, possuir conhecimentos abrangentes de uma vasta literatura taxonômica (e histórica) em diferentes idiomas, e, além disso, dominar métodos de análises de filogenética, microscopia, e, especialmente no contexto atual, compreender as aplicações de marcadores moleculares em sistemática (Wägele 2011). Mesmo dando o devido destaque ao taxonomista, o ecólogo, geneticista, e outros pesquisadores da biodiversidade, são limitados pelo tempo e precisam estabelecer estratégias eficientes para conciliar as complexas dificuldades e responder as perguntas científicas. Nesse quesito, conseguir inovações nas linhas de pesquisa em biodiversidade parece ser um caminho.

Com o advento da era da bioinformática e ecoinformática, a formação de novos taxonomistas e cientistas de base tem se tornado ainda mais comprometida. Atualmente vivemos em um contexto científico no qual se a informação existe, ela deve ser disponibilizada para a sociedade (Boero 2010). Isso se traduz no enviesamento dos recursos financeiros para iniciativas que objetivam construir, armazenar e gerenciar conjuntos de dados digitalmente acessíveis que sumarizam diferentes facetas da biodiversidade, seja em escala local, regional ou global (Bingham *et al.* 2017). Além disso, iniciativas deste tipo, como por exemplo, “*Catalogue of Life*” (<https://www.catalogueoflife.org/>) e “*Global Biodiversity Information Facility*” (<https://www.gbif.org/>) dão a impressão equivocada de que a taxonomia recebe um apoio eficiente e suficiente (Löbl *et al.* 2023). Essa “tendência” parece ter alocado parte dos recursos disponíveis para a conservação da biodiversidade para o financiamento de propostas que objetivam o surgimento de novas tecnologias, como a bioinformática (Löbl *et al.* 2023). Em contrapartida, os recursos destinados aos estudos taxonômicos, de história natural e de ciência de base, foram diminuindo (Boero 2010) mantendo as lacunas no conhecimento. Logo, a integração entre as áreas da biodiversidade é necessária como forma de (i) demonstrar quais as lacunas de conhecimento existentes e (ii) preenchê-las. Sem a integração entre as áreas, e com o isolamento delas, a inovação nas pesquisas ficará cada vez mais distante.

Outro desafio que surge é a curadoria contínua dos bancos de dados, que nesse caso inclui tanto as coleções biológicas físicas quanto os repositórios online. Assim, torna-se de extrema urgência a obtenção de subsídios financeiros, tanto para manutenção da infraestrutura física das coleções, mas também a fim de organizar e manter todos os espécimes em bom estado, o que se traduz na contratação de técnicos capacitados tanto em taxonomia quanto em bioinformática. Além disso, também é necessária uma infraestrutura de informática para a mobilização de dados em coleções biológicas de forma aberta e colaborativa (Krishtalka & Humphrey 2000, Nelson & Ellis 2018, Johnson *et al.* 2023), seja por meio de políticas locais de cada país ou acordos globais, como a Convenção sobre Diversidade Biológica (CDB) (<https://www.cbd.int/convention/>). Isso é especialmente importante considerando que acidentes em coleções biológicas resultam em perdas inestimáveis (Phillips, 2010; Phillips, 2018), principalmente quando não temos informações digitalizadas.

## CONCLUSÕES

Por fim, reforçamos que novas coletas de dados são importantes para obtenção de mais informações e preenchimento das lacunas. O taxonomista é importante para descrever e detalhar informações sobre as espécies (p.ex., distribuição geográfica, história natural, habitat, entre outras). Os ecólogos são importantes para entender os processos locais e regionais que mantem e estruturam a biodiversidade. Esses fatos remontam a necessidade da interação entre as áreas. A curadoria dos dados e a bioinformática também são importantes para sumarizar e disponibilizar essa informação em grandes bancos de dados que sejam úteis para projetos que objetivem as tomadas de decisão e a descrição de padrões em larga escala. Contudo, deve-se ter em mente que os bancos de dados construídos, por serem baseados nos dados primários, sempre possuem lacunas, limitações, vieses, déficits e ruídos. Estas deficiências não podem ser negligenciadas e devem ser estudadas e discutidas entre taxonomistas, ecólogos, e diferentes pesquisadores da biodiversidade para que os padrões, processos e tomadas de decisões sejam trabalhadas em conjunto e com excelência.

## **AGRADECIMENTOS**

Somos gratos à equipe do Laboratório de Macroecologia e Conservação da Biodiversidade (MacrEco) da Universidade Federal de Mato Grosso (UFMT) pelas valiosas discussões. Agradecemos a FAPEMAT (projeto FAPEMAT-PRO.000274/2023) e a IABS/CECAV (edital 01/2023 - TCCE Vale 01/2022) pelo financiamento. RACS, NB e AHK agradecem a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelas bolsas de doutorado concedidas e ao programa de Pós-graduação em Ecologia e Conservação da Biodiversidade (PPGECB). VCS agradece ao Programa de Pós-Doutorado da University of Pretoria (África do Sul). RACS, JAFDF, GT e TSS agradecem ao Instituto Nacional de Ciência & Tecnologia (INCT) em Ecologia, Evolução e Conservação da Biodiversidade (EECBio), sediado na Universidade Federal de Goiás (UFG), pelas preciosas discussões que contribuíram para o desenvolvimento deste artigo.

## **REFERÊNCIAS**

- Ahrends, A., Rahbek, C., Bulling, MT., Burgess, ND., Platts, PJ., Lovett, JC., Kindemba, VW., Owen, N., Sallu, AN., Marshall, AR., Mhoru, BE., Fanning, E., & Marchant, R. 2011. Conservation and the botanist effect. *Biological Conservation*, 144(1), 131–140. DOI: 10.1016/j.biocon.2010.08.008
- Almeida, LM., Ribeiro-Costa, CS., & Marinoni, L. 2012. Coleta, montagem, preservação e métodos para estudo [pp. 175-190]. In: Rafael, JA., Melo, GAR., Carvalho CJB., Casari, AS, & Constantino, R (Eds). *Insetos do Brasil: diversidade e taxonomia*, Holos Editora, Ribeirão Preto (SP), 810 p.
- Alves, CBM., Pompeu, PS., Mazzoni, R., & Brito, MFG. 2021. Avanços em métodos de coleta de peixes e caracterização de habitat de riachos tropicais. *Oecologia Australis*, 25(2), 246–265. DOI: 10.4257/oeco.2021.2502.03

- Appel, G., Capaverde, UD., de Oliveira, LQ., Pereira, LGA., Tavares, VC., López-Baucells, A., Magnusson WE., Baccaro FB., & Bobrowiec, PE. 2021. Use of complementary methods to sample bats in the Amazon. *Acta Chiropterologica*, 23(2), 499–511. DOI: 10.3161/15081109ACC2021.23.2.017
- Disney, H. 1998. Rescue plan needed for taxonomy. *Nature*, 394(6689), 120–120. DOI: 10.1038/28027
- Freckleton, R.P., & Watkinson, A.R. 2002. Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology*, 90(3), 419–434. DOI: 10.1046/j.1365-2745.2002.00692.x
- Ficetola, GF., Canedoli, C., & Stoch, F. 2019. The Racovitza impediment and the hidden biodiversity of unexplored environments. *Conservation Biology*, 33(1), 214–216. DOI: 10.1111/cobi.13179
- Freitas, TMDS., Stropp, J., Calegari, BB., Calatayud, J., De Marco Jr, P., Montag, LFDA., & Hortal, J. 2021. Quantifying shortfalls in the knowledge on Neotropical Auchenipteridae fishes. *Fish and Fisheries*, 22(1), 87–104. DOI: 10.1111/faf.12507
- Freeman, B.G., & Pennell, M.W. 2021. The latitudinal taxonomy gradient. *Trends in Ecology & Evolution*, 36(9), 778–786. DOI: 10.1016/j.tree.2021.05.003
- Bach, A., Mateus, LA., Peres, C. A., Haugaasen, T., Louzada, J., Hawes, JE., Azevedo, RA., Lucena, E., Ferreira, JVA., & Vaz-de-Mello, FZ. 2023. Bait attractiveness changes community metrics in dung beetles (Coleoptera: Scarabaeidae: Scarabaeinae). *Ecology and Evolution*, 13(4), e9975. DOI: 10.1002/ece3.9975
- Bakker, FT., Antonelli, A., Clarke, JA., Cook, JA., Edwards, SV., Ericson, PGP., Faurby, S., Ferrand, N., Gelang, M., Gillespie, RG., Irestedt, M., Lundin, K., Larsson, E., Matos-Maraví, P., Müller, J., von Proschwitz, T., Roderick, GK., Schliep, A., Wahlberg, N., Wiedenhoeft, J., & Källersjö, M. 2020. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ*, 8, e8225. DOI: 10.7717/peerj.8225

- Benecke, M. 2001. A brief history of forensic entomology. *Forensic Science International*, 120, 2–114.
- Bingham, HC., Doudin, M., Weatherdon, LV., Despot-Belmonte, K., Wetzell, FT., Groom, Q., Lewis, E., Regan, E., Appeltans, W., Güntsch, A., Mergen, P., Agosti, D., Penev, L., Hoffmann, A., Saarenmaa, H., Geller, G., Kim, K., Kim, H., Archambeau, AS., Häuser, C., Schmeller, DS., Geijzendorffer, I., Camacho, Guerra, AGC., Robertson, T., Runnel, V., Valland, N., & Martin, CS. 2017. The biodiversity informatics landscape: elements, connections and opportunities. *Research Ideas and Outcomes*, 3, e14059. DOI: 10.3897/rio.3.e14059
- Boakes, EH., McGowan, PJ., Fuller, RA., Chang-qing, D., Clark, NE., O'Connor, K., & Mace, GM. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS biology*, 8(6), e1000385. DOI: 10.1371/journal.pbio.1000385
- Boero, F. 2010. The study of species in the era of biodiversity: a tale of stupidity. *Diversity*, 2(1), 115–126. DOI: 10.3390/d2010115
- Boggs, SW. 1949. An atlas of ignorance: a needed stimulus to honest thinking and hard work. *Proceedings of the American Philosophical Society*, 93(3), 253–258.
- Bortolus, A. 2008. Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *AMBIO: A journal of the human environment*, 37(2), 114–118. DOI: 10.1579/0044-7447(2008)37[114:ECITBS]2.0.CO;2
- Cardoso, P., Erwin, TL., Borges, PA., & New, TR. 2011. The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, 144(11), 2647–2655. DOI: 10.1016/j.biocon.2011.07.024
- Cornwell, WK., Pearse, WD., Dalrymple, RL., & Zanne, AE. 2019. What we (don't) know about global plant diversity. *Ecography*, 42(11), 1819–1831. DOI: 10.1111/ecog.04481
- Coyne JA, & Orr HA. 2004. Species: Reality and Concepts. In: J. A Coyne & H. A. Orr (Eds.), *Speciation*. pp. 9–54. Sunderland, MA: Sinauer Associates.

- Diniz-Filho, JAF., Loyola, RD., Raia, P., Mooers, AO., & Bini, LM. 2013. Darwinian shortfalls in biodiversity conservation. *Trends in Ecology & Evolution*, 28(12), 689–695. DOI: <https://doi.org/10.1016/j.tree.2013.09.003>
- Diniz-Filho, JAF., Jardim, L., Guedes, JJM., Meyer, L., Stropp, J., Frateles, LEF., Pinto, RB., Lohmann, L. G., Tessarolo, G., Carvalho, CJB., Ladle, RJ., & Hortal, J. 2023. Macroecological links between the Linnean, Wallacean, and Darwinian shortfalls. *Frontiers of Biogeography*, 15(2), e59566. DOI:10.21425/F5FBG59566
- Disney, RHL. 1998. Insect biodiversity and the demise of alpha taxonomy. *Antenna*, 23, 84–88.
- Escribano, N., Ariño, AH., & Galicia, D. 2016. Biodiversity data obsolescence and land uses changes. *PeerJ*, 4, e2743. DOI: [10.7717/peerj.2743](https://doi.org/10.7717/peerj.2743)
- Evenhuis, NL. 1997. *Litteratura taxonomica dipterorum (1758–1930)*. Volume II. L–Z. Backhuys, Leiden, pp. 427–871.
- Fontaine, B., Perrard, A., & Bouchet, P. 2012. 21 years of shelf life between discovery and description of new species. *Current Biology*, 22(22), R943– R944. DOI: [10.1016/j.cub.2012.10.029](https://doi.org/10.1016/j.cub.2012.10.029)
- Freitas, TMDS., Stropp, J., Calegari, BB., Calatayud, J., De Marco Jr, P., Montag, LFDA., & Hortal, J. 2021. Quantifying shortfalls in the knowledge on Neotropical Auchenipteridae fishes. *Fish and Fisheries*, 22(1), 87-104. DOI: [10.1111/faf.12507](https://doi.org/10.1111/faf.12507)
- Greshko, M. 2018. Fire Devastates Brazil’s Oldest Science Museum. *National Geographic* (diponível: [www.nationalgeographic.com/science/article/news-museunacional-fire-rio-de-janeiro-natural-history](http://www.nationalgeographic.com/science/article/news-museunacional-fire-rio-de-janeiro-natural-history)).
- Gutierrez-Velez, VH., & Wiese, D. 2020. Sampling bias mitigation for species occurrence modeling using machine learning methods. *Ecological Informatics*, 58, 101091. DOI: [10.1016/j.ecoinf.2020.101091](https://doi.org/10.1016/j.ecoinf.2020.101091)
- Heidorn, PB. 2008. Shedding light on the dark data in the long tail of science. *Library Trends*, 57(2), 280–299. DOI: [10.1353/lib.0.0036](https://doi.org/10.1353/lib.0.0036)

- Hey, J., Waples, RS., Arnold, ML., Butlin, RK. & Harrison, RG. 2003. Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology & Evolution*, 18(11), 597–603. DOI: 10.1016/j.tree.2003.08.014
- Horn, W., Kahle, I., Friese, G. & Gaedike, R. 1990a. *Collectiones entomologicae. Ein Kompendium über den Verbleib entomologischer Sammlungen der Welt bis 1960. Teil I: A bis K.* Berlin (Akademie der Landwirtschaftswissenschaften der Deutschen Demokratischen Republik), pp. 1–220.
- Horn, W., Kahle, I., Friese, G. & Gaedike, R. 1990b. *Collectiones entomologicae. Ein Kompendium über den Verbleib entomologischer Sammlungen der Welt bis 1960. Teil II: L bis Z.* Berlin (Akademie der Landwirtschaftswissenschaften der Deutschen Demokratischen Republik), pp. 221–573.
- Hortal, J., Jiménez-Valverde, A., Gómez, JF., Lobo, JM., & Baselga, A. 2008. Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos*, 117(6), 847–858. DOI: 10.1111/j.0030-1299.2008.16434.x
- Hortal, J., de Bello, F., Diniz-Filho, JAF., Lewinsohn, TM., Lobo, JM., & Ladle, RJ. 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 523–549. DOI: 10.1146/annurev-ecolsys-112414-054400
- Hughes, AC., Orr, MC., Ma, K., Costello, MJ., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. 2021. Sampling biases shape our view of the natural world. *Ecography*, 44, 1–14.
- Jaspars, M. 1998. Tough time for taxonomy. *Nature*, 394(6692), 413–413. DOI: 10.1038/28723
- Johnson, KR., Owens, IF., & Global Collection Group. 2023. A global approach for natural history museum collections. *Science*, 379(6638), 1192–1194. DOI: 10.1126/science.adf6434
- Krell, F-T. 2012. The Journal Impact Factor as a performance indicator. *European Science Editing*, 38, 3–6.

- Krishtalka, L., & Humphrey, PS. 2000. Can natural history museums capture the future?. *BioScience*, 50(7), 611–617. DOI: 10.1641/0006-3568(2000)050[0611:CNHMCT]2.0.CO;2
- Lees, AC., Rosenberg, KV., Ruiz-Gutierrez, V., Marsden, S., Schulenberg, TS., & Rodewald, AD. 2020. A roadmap to identifying and filling shortfalls in Neotropical ornithology. *The Auk*, 137(4), ukaa048. DOI: 10.1093/auk/ukaa048
- Ladle, RJ., & Hortal, J. 2013. Mapping species distributions: living with uncertainty. *Frontiers of Biogeography*, 5(1), 8–9. DOI: 10.21425/F55112942
- Lamas, G. 1980. Introducción a la Historia de la Entomología en el Perú. II. Período de los viajeros, Colectores y Estudiosos Especializado. *Revista Peruana de Entomología*, 23(1), 25–31.
- Lobo, JM., Hortal, J., Yela, JL., Millán, A., Sánchez-Fernández, D., García-Roselló, E., González-Dacosta, J., Heine, J., González-Vilas, L., & Guisande, C. 2018. KnowBR: An application to map the geographical variation of survey effort and identify well-surveyed areas from biodiversity databases. *Ecological Indicators*, 91, 241–248. DOI: 10.1016/j.ecolind.2018.03.077
- Löbl, I., Klausnitzer, B., Hartmann, M., & Krel, F-T. 2023. The silent extinction of the species and taxonomists - An appeal to Science policymakers and legislators. *Diversity*, 15(10): 1–17. DOI: 10.3390/d15101053
- Kemp, C. 2015. The endangered dead. *Nature* 518(7539), 292–294. DOI: 10.1038/518292a
- Lomolino MV. 2004. Conservation biogeography. In: M.V. Lomolino & L.R. (Eds.) Heaney *Frontiers of Biogeography: New Directions in the Geography of Nature*. pp. 293–96. Sunderland, MA: Sinauer.
- Lopes-Lima, M., Riccardi, N., Urbanska, M., Köhler, F., Vinarski, M., Bogan, AE., & Sousa, R. 2021. Major shortfalls impairing knowledge and conservation of freshwater molluscs. *Hydrobiologia*, 848(12–13), 2831–2867. DOI: doi.org/10.1007/s10750-021-04622-w

- Masiero, FS., Martins, DS., & Thyssen, PJ. 2015. Terapia Larval e a aplicação de larvas para cicatrização: revisão e estado da arte no Brasil e no mundo. *Revista Thema*, 12(1), 4–14.
- Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. 2015. Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, 6(1), 1–8. DOI: 10.1038/ncomms9221
- Meyer, C., Weigelt, P., & Kreft, H. 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. *Ecology Letters*. 19(8), 992–1006. DOI: 10.1111/ele.12624
- Mora, C., Tittensor, DP., Adl, S., Simpson, AG., & Worm, B. 2011. How many species are there on Earth and in the ocean?. *PLoS Biology*, 9(8), e1001127. DOI: 10.1371/journal.pbio.1001127
- Murphey, PC., Guralnick, RP., Glaubitz, R., Neufeld, D., & Ryan, JA. 2004. Georeferencing of museum collections: A review of problems and automated tools, and the methodology developed by the Mountain and Plains Spatio-Temporal Database-Informatics Initiative (Mapstedi). *PhyloInformatics*, 3, 1–29. DOI: 10.5281/zenodo.59792
- Nelson, G., & Ellis, S. 2018. The history and impact of digitization and digital data mobilization on biodiversity research. *Philosophical Transactions of the Royal Society B*, 374, 20170391. DOI: 10.1098/rstb.2017.0391
- Pace, S., Frost, G., Lachow, I., Frelinger, D., Fossum, D., Wassem, D., & Pinto, MM. 1995. The global positioning system: assessing national policies. Santa Monica, CA: Rand: p. 368.
- Papavero, N. 1971. Essays on the history of Neotropical dipterology, with special reference to collectors (1750–1905). Vol. I. São Paulo: Museu de Zoologia, Universidade de São Paulo: p. 216.
- Papavero, N. 1973. Essays on the history of Neotropical dipterology, with special reference to collectors (1750–1905). Vol. II. São Paulo: Museu de Zoologia, Universidade de São Paulo: p. 217–426.
- Papavero, N. 1994. Fundamentos práticos de taxonomia zoológica. Sao Paulo: Editora Unesp. p. 285.

- Pearson, DL., Hamilton, AL., & Erwin, TL. 2011. Recovery plan for the endangered taxonomy profession. *BioScience*, 61(1), 58–63. DOI:10.1525/bio.2011.61.1.11
- Phillips T. 2010. Sao Paulo fire destroys one of the largest collections of dead snakes. *The Guardian*. <https://www.theguardian.com/world/2010/may/16/firedestroys-snakecollection>.
- Phillips T. 2018. Brazil National Museum: as much as 90% of collection destroyed in fire. *The Guardian*. <https://www.theguardian.com/world/2018/sep/04/brazilnationalmuseum-fire-collection-destroyed-not-insured>.
- Raven, PH., & Wilson, EO. 1992. A fifty-year plan for biodiversity surveys. *Science*, 258(5085), 1099–1100. DOI: 10.1126/science.258.5085.1099
- Raja, NB., Dunne, EM., Matiwane, A., Khan, TM., Nätcher, PS., Ghilardi, AM., & Chattopadhyay, D. 2022. Colonial history and global economics distort our understanding of deep-time biodiversity. *Nature ecology & evolution*, 6(2), 145–154. DOI: <https://doi.org/10.1038/s41559-021-01608-8>.
- Rocchini, D., Hortal, J., Lengyel, S., Lobo, JM., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. 2011. Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, 35(2), 211–226. DOI: 10.1177/0309133311399491
- Ronquillo, C., Alves-Martins, F., Mazimpaka, V., Sobral-Souza, T., Vilela-Silva, B., Medina, N. G., & Hortal, J. 2020. Assessing spatial and temporal biases and gaps in the publicly available distributional information of Iberian mosses. *Biodiversity Data Journal*, 8, e53474. DOI: 10.3897/BDJ.8.e53474
- Ronquillo, C., Stropp, J., Medina, NG., & Hortal, J. 2023. Exploring the impact of data curation criteria on the observed geographical distribution of mosses. *Ecology and Evolution*, 13(12), e10786. DOI: 10.1002/ece3.10786

- Rosado, BH., Figueiredo, MS., de Mattos, EA., & Grelle, CE. 2016. Eltonian shortfall due to the Grinnellian view: functional ecology between the mismatch of niche concepts. *Ecography*, 39(11), 1034–1041. DOI: 10.1111/ecog.01678
- Rudbeck, AV., Sun, M., Tietje, M., Gallagher, RV., Govaerts, R., Smith, SA., Svenning, JC., & Eiserhardt, WL. 2022. The Darwinian shortfall in plants: phylogenetic knowledge is driven by range size. *Ecography*, 2022: e06142. DOI: 10.1111/ecog.06142
- Salvador, RB., Cavallari, DC., Rands, D., & Tomotani, BM. 2022. Publication practice in Taxonomy: Global inequalities and potential bias against negative results. *Plos one*, 17(6), e0269246. DOI: 10.1371/journal.pone.0269246
- Schlick-Steiner, BC., Steiner, FM., Seifert, B., Stauffer, C., Christian, E., Cozier, RH. 2010. Integrative taxonomy: A multisource approach to exploring biodiversity. *Annual Review of Entomology* 55, 421–438. DOI: 10.1146/annurev-ento-112408-085432
- Scott, WA., & Hallam, CJ. 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. *Plant Ecology*, 165(1), 101–115. DOI: 10.1023/A:1021441331839
- Sikes, DS., Copas, K., Hirsch, T., Longino, JT., & Schigel, D. 2016. On natural history collections, digitized and not: a response to Ferro and Flick. *ZooKeys*, 618, 145–158. DOI: 10.3897/zookeys.618.9986
- Siracusa, PC., Gadelha, LM., & Ziviani, A. 2020. New perspectives on analysing data from biological collections based on social network analytics. *Scientific reports*, 10(1), 1–10. DOI: 10.1038/s41598-020-60134-y
- Soberón, J., & Peterson, T. 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1444), 689–698. DOI: 10.1098/rstb.2003.1439

- Soberón, J., Jiménez, R., Golubov, J., & Koleff, P. 2007. Assessing completeness of biodiversity databases at different spatial scales. *Ecography*, 30(1), 152–160. DOI: 10.1111/j.0906-7590.2007.04627.x
- Soberón, J., & Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, 106(supplement\_2), 19644–19650.
- Sobral-Souza, T., Bastos Francini, R., Guimarães, M., & Benson, WW. 2015. Short-term dynamics reveals seasonality in a subtropical *Heliconius* butterfly. *Journal of Insects*, 761058, 1–5. DOI: 10.1155/2015/761058
- Sobral-Souza, T., Stropp, J., Santos, JP., Prasniewski, VM., Szinwelski, N., Vilela, B., ... & Hortal, J. 2021. Knowledge gaps hamper understanding the relationship between fragmentation and biodiversity loss: the case of Atlantic Forest fruit-feeding butterflies. *PeerJ*, 9, e11673. DOI: 10.7717/peerj.11673
- Sousa-Baena, MS., Garcia, LC., & Peterson, AT. 2014. Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. *Diversity and Distributions*, 20(4), 369–381. DOI: 10.1111/ddi.12136
- Stephens, L., & Traylor, MA. 1983. *Ornithological gazetteer of Peru*. 2ed. Obtainable from Bird Dept., Museum of Comparative Zoology: Harvard University, vi: p. 273.
- Stork, NE. 2018. How many species of insects and other terrestrial arthropods are there on earth? *Annual Review of Entomology*, 63, 31–45. DOI: 10.1146/annurev-ento-020117-043348
- Stropp, J., Ladle, RJ., Malhado, ACM., Hortal, J., Gaffuri, J., Temperley, WH., Skøien, JO., Mayaux, P. 2016. Mapping ignorance: 300 years of collecting flowering plants in Africa. *Global Ecology and Biogeography*, 25(9), 1085–1096. DOI: 10.1111/geb.12468

- Stropp, J., Ladle, R.J., Emilio, T., Lessa, T., & Hortal, J. 2022. Taxonomic uncertainty and the challenge of estimating global species richness. *Journal of Biogeography*, 49, 1654–1656. DOI: 10.1111/jbi.14463
- Tessarolo, G., Ladle, R., Rangel, T., & Hortal, J. 2017. Temporal degradation of data limits biodiversity research. *Ecology and Evolution*, 7(17), 6863–6870. DOI: 10.1002/ece3.3259
- Tessarolo, G., Ladle, R.J., Lobo, J.M., Rangel, T.F., & Hortal, J. 2021. Using maps of biogeographical ignorance to reveal the uncertainty in distributional data hidden in species distribution models. *Ecography*, 44(12), 1743–1755. DOI: 10.1111/ecog.05793
- TrouDET, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. 2017. Taxonomic bias in biodiversity data and societal preferences. *Scientific reports*, 7(1), 1–14. DOI: 10.1038/s41598-017-09084-6
- Wägele, H., Klussmann-Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A., & Wägele, J.W. 2011. The taxonomist-an endangered race. A practical proposal for its survival. *Frontiers in zoology*, 8(1), 1–7. DOI: 10.1186/1742-9994-8-25
- Werner, Y.L. 2009. The aspiration to be good is bad: The “Impact Factor” hurts both science and society. *The International Journal of Science and Society*, 1, 99–105. DOI: 10.18848/1836-6236/CGP/v01i01/51264
- Wheeler, Q. 2014. Are reports of the death of taxonomy an exaggeration?. *New Phytologist*, 201(2), 370–371. DOI: 10.1111/nph.12612
- Wheeler, Q.D. 1995. The “Old Systematics”: Classification and phylogeny. In: *Biology, Phylogeny and Classification of Coleoptera: Papers celebrating the 80th Birthday of Roy A. Crowson*; Pakaluk, J., Slipiński, S.A. (Eds.); pp. 31–62. Muzeum i Institut Zoologii: Warszawa, Poland.
- Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M., & Willis, K.J. 2005. Conservation Biogeography: assessment and prospect. *Diversity and Distributions*, 11: 3-23. DOI: 10.1111/j.1366-9516.2005.00143.x

Zizka, A., Antonelli, A., & Silvestro, D. 2020. Sampbias, a method for quantifying geographic sampling biases in species distribution data. *Ecography*, 44(1), 25–32. DOI: 10.1111/ecog.05102

Zizka, A., Rydén, O., Edler, D., Klein, J., Perrigo, A., Silvestro, D., ... & Antonelli, A. 2021. Bio-Dem, a tool to explore the relationship between biodiversity data availability and socio-political conditions in time and space. *Journal of Biogeography*, 48(11), 2715–2726. DOI: 10.1111/jbi.14256

*Submitted: 17/10/2023*

*Accepted: 18/04/2024*

*Published online: 26 April 2024*

*Editor Associado: Camila dos Santos de Barros*

Ahead of print