Received: 26 April 2022 | Revised: 12 April 2023 | Accepted: 3 May 2023

DOI: 10.1111/jbi.14631

METHOD

Journal of Biogeography WILEY

BIOREALM—An ontology of comparative biogeography: New insights into the semantics of biodiversity conservation

Visotheary Ung[1] | Pier Luigi Buttigieg[2]

[1]UMR 7205 ISYEB, CNRS, MNHN, SU, UA, Paris, France

[2]GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Kiel, Germany

Correspondence
Visotheary Ung, UMR 7205 ISYEB, Centre National Recherche Scientifique, MNHN, SU, UA, EPHE, 57 rue Cuvier CP39 75005 Paris, France.
Email: visotheary.ung@mnhn.fr

Handling Editor: Lynne (GE) R Parenti

Abstract

Aim: We aimed to apply ontological techniques to address semantic ambiguities in protected area and conservation informatics. By doing so, we aimed to create a coherent, machine-actionable semantic representation of the biogeographic areas (which often overlap protected areas) to support more efficient and standardized informatics, supporting research and decision-making. We present BIOREALM, the first informatic ontology for comparative biogeography.

Location: Global.

Taxon: Any taxon can be integrated in BIOREALM.

Methods: We convert a cladogram of biogeographic areas—generated by a process known as bioregionalization—into a series of ontological classes. Areas of endemism are treated as formal objects related by hierarchical relationships and constrained by a condition of monophyly. We use semantic web approaches to extend the Environment Ontology (ENVO) with classes for (often semantically confounded) biogeographic entities, including biogeographic areas, areas of endemism and endemic areas. We applied this approach to a bioregionalization of Australia as a case study. In all, 20 subregions which are part of the Austral Bioregionalisation Atlas have been selected for the study and integrated in BIOREALM.

Results: We have created an ontology—formatted in the Web Ontology Language and adhering to the practices of the Open Biomedical and Biological Ontology Foundry—which provides a rigorous, extensible and machine-actionable framework that can improve biogeographic analyses and interoperability between systems. One main class and 20 individuals per class were implemented.

Main Conclusions: BIOREALM encodes a model-theoretic view of endemism using semantic web approaches, offering new avenues to express and analyse biogeographic units. This approach offers a means to identify monophyletic biogeographic areas for conservation, based on specific combinations of monophyletic endemic taxa. Such an ontology provides knowledge representation solutions which supports interoperability along the FAIR (Findable, Accessible, Interoperable, Reusable) principles, thus fostering more consistent ecological informatics.

KEYWORDS
biogeographic areas, comparative biogeography, conservation, controlled vocabularies, monophyletic endemics, ontology, semantics

1 | INTRODUCTION

We are facing the sixth wave of extinction (Ceballos et al., 2015, 2020; Cowie et al., 2022; Kolbert, 2014; Régnier et al., 2015 among others), a major ecological disaster and loss of biodiversity largely precipitated by human activities, ecological stressors, habitat destruction, habitat fragmentation, global biodiversity loss (Butchart et al., 2010; Cardillo & Pratt, 2013; Gonzalez-Orozco et al., 2016; Myers et al., 2000; Sachs et al., 2009) and climate change (IPCC AR5, 2014; Scriven, Hodgson, et al., 2015; Thornton et al., 2014). In this dire scenario, the main question for policymakers remains: how can we prevent this from worsening? What areas and ecosystems are the most urgent to protect? Which indicators should guide us in our decisions?

Conservation biology aims to identify ecosystems most worthy of protection due to their unique value (Dudley & Stolton, 2008), biodiversity hotspots are recognized all over the globe (Bellwood et al., 2012; Myers et al., 2000; Seligmann et al., 2007; Woodruff, 2010). However, the social cost of maintaining protected status competes with conservation regimes, and such regimes are no guarantee that biodiversity decline will be halted or reversed (Craigie et al., 2010; Dudley et al., 2014; Orme et al., 2005), often due to poor staffing (Gill et al., 2017). This has been observed both in terrestrial (Lele et al., 2010; Scriven, Woodall, et al., 2015) and marine environments (Camillo & Peter, 2011; Giakoumi et al., 2018; Rife et al., 2013). At present, the selection and delimitation of protected areas is still an intuitive and often subjective process. One outcome of this subjectivity is that the very terminology used to reference such areas contains semantic, which hinder comparisons. One key concept which falls victim to semantic ambiguity is that of endemism and the definition of 'areas of endemism' (Andersen, 1994; Crother & Murray, 2011; Harold & Mooi, 1994; Morrone, 1994 among others).

Crother and Murray in 2011 define an area of endemism as:

> An area of endemism is a spatially and temporally bounded geographical area with species. Neither species alone nor geographical areas alone are sufficient for diagnosis. Species and the areas they are distributed in constitute a single unit. The space and time boundaries of these areas are flexible, as is their existence, because the species and the geographical areas themselves are interactors, involved in processes that can eradicate, shrink, or expand the areas of endemism. They are also reducible in that they are nested within one another. Operationally, an area that contains at least one unique species, or a unique combination of species, is an area of endemism, and the biogeographical use of multiple areas of endemism is scale-specific. Crother & Murray, 2011: 1013

Furthermore, Orme et al., 2005 stressed the fact that 'global hotspots of species richness are not necessarily congruent with endemism or threat' (Orme et al., 2005). Simultaneously, there is no consensus about what a protected area should be, with each international organization having its own definition of a protected area. According to the IUCN, a protected area is '*an area of land and/or sea especially dedicated to the protection and maintenance of biological diversity, and of natural and associated cultural resources, and managed through legal or other effective means*', whereas for the Convention on Biological Diversity, a protected area is '*a geographically defined area which is designated or regulated and managed to achieve specific conservation objectives*'. Biodiversity hotspots have been defined as a combination of 'areas with exceptional concentrations of endemic species and undergoing exceptional loss of habitat' based on the analysis of about 800 scientific articles and according to more than 100 scientists with significant experience in the concerned countries (Myers et al., 2000). The authors stated that 'the hotspots boundaries have been determined by 'biological commonalities''. Each of the areas featured a separate biota or community of species that fit together as a 'biogeographic unit'. Finally, the many attempts to define protected areas (e.g. Kerr, 1997; Bonn et al., 2002; Dudley, 2008; Dudley & Stolton, 2008; Gillepsie, 2009; Solton & Dudley, 2010 among others) have not led to a global consensus or semantic. Crother and Murray (2011: 1013) have anticipated the importance of a deeper understanding of areas of endemism with regard to conservation and biodiversity (Crother & Murray, 2011:1013). Thus, we propose that the application of ontological techniques, that is definition of terms and their relationships to the ambiguities in protected area semantics is a timely contribution to conservation.

We suggest that a coherent semantic representation of the understandings of biogeographic areas and protected areas will provide more efficient and standardized informatics, supporting research and decision-making processes.

2 | MATERIALS AND METHODS

To avoid any ambiguities vocabulary used below, we first provide some terminological clarifications. BIOREALM is an ontology in the sense of being a computable and logically coherent representation of entities; it does not attempt to provide a complete metaphysical account of its subject matter as one would expect in Ontology, as defined by Guarino et al. (1995). In other words, we do not address the metaphysics of biogeographic areas and areas of endemism (see Crother & Murray, 2011), but we offer a formalization of knowledge for comparative biogeography. We use semantic web standards, Resource Description Framework and the Web Ontology Language (OWL) published by the World Wide Web Consortium (www.w3c.org) to address this need. Generic ontological classes are used to define categories and individuals which belong to these classes are understood as instances of the class. All entities (classes, individuals and relations) are associated with a unique identifier (a Uniform Resource Identifier, URI).

Here we present an example of the implementation of this model. We applied these methods to a study and particularly, its associated dataset, published by Murphy et al., 2019.

BIOREALM is expresses in OWL following the technical guidelines of the OBO foundry and Library. BIOREALM was developed as an extension of the Environment ontology (ENVO, http://purl.obolibrary.org/obo/envo.owl) (Buttigieg et al., 2013), contending with often confounded biogeographic entities including biogeographic areas, for example, areas of endemism and endemic areas as well as their relationships. We used the OBO foundry tools (https://obofoundry.org/docs/OntologyTools) to generate BIOREALM. The Ontology Development Toolkit (ODK) was used to generate the development and release environment for the ontology, and Protégé was used to manually edit and add content to the ontology. BIOREALM is still in development (https://github.com/EnvironmentOntology/BIOREALM).

Through the standardized routines provided by the ODK, BIOREALM imports the following ontologies: BFO (https://obofoundry.org/ontology/bfo.owl), Ecocore (https://obofoundry.org/ontology/ecocore.owl), PATO (https://obofoundry.org/ontology/pato.owl), PCO (https://obofoundry.org/ontology/pco.owl), RO (https://obofoundry.org/ontology/ro.owl) and NCBITaxon (https://obofoundry.org/ontology/ncbitaxon.owl) to ensure full interoperability with those ontologies and provide a basis to import multiple domain ontologies into ontologies supporting comparative biogeography.

3 | RESULTS

In the following sections, ontology classes (or ontology 'terms'), are written in italics and are marked with the namespace associated with the ontology they come from. For example: 'ENVO: *ecosystem*' is a class from the Environment Ontology (which has the ENVO namespace). Retrieving this class from ENVO, one would discover the OBO Foundry Uniform Resource Identifier, 'ENVO:01001110' and the full URIs of the form: http://purl.obolibrary.org/obo/ENVO_01001110.

3.1 | Semantics of biogeographic terms

BIOREALM—as an extension of ENVO—inherits ENVO's structure and hierarchy (see Figure 1a). Its main class—BIOREALM: '*ecosystem with monophyletic endemics*'—was created as a subclass of ENVO's

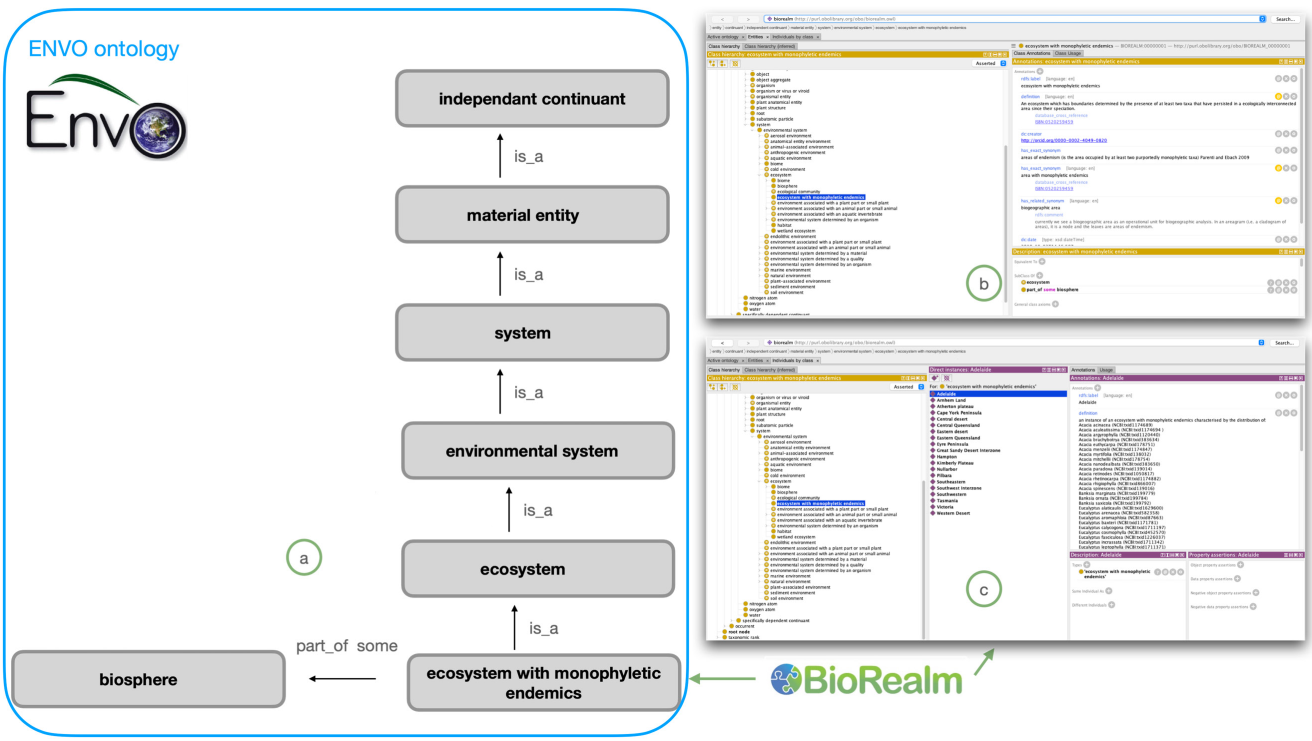

FIGURE 1 BIOREALM in ENVO ontology expressed in the software Protégé. (a) Subclasses in ENVO's hierarchy and BIOREALM main class, (b) Entities and class hierarchy, (c) Individuals by class or direct instances

class *ecosystem* (an environmental system which includes both living and non-living components).

The immediate semantic neighbourhood of BIOREALM's main class is:

1. biome [ENVO_00000428],
2. biosphere [ENVO_01000817],
3. ecological community [PCO_0000002],
4. **ecosystem with monophyletic endemics [BIOREALM_00000001]**,
5. environment associated with a plant part or small plant [ENVO_01001057],
6. environment associated with an animal part or small animal [ENVO_01001055],
7. environment associated with an aquatic invertebrate [ENVO_01001176]
8. environmental system determined by an organism [ENVO_01001000]
9. habitat [ENVO_01000739]
10. wetland ecosystem [ENVO_01001209]

BIOREALM: *ecosystem with monophyletic endemics* is defined as:

> an ecosystem which has boundaries determined by the presence of at least two taxa that have persisted in an ecologically interconnected area since their speciation.

To account for linguistic variation, we added string literals as the values of the OBO Foundry's annotation properties has_exact_synonym and has_related_synonym (http://www.geneontology.org/formats/oboInOwl#hasExactSynonym, http://www.geneontology.org/formats/oboInOwl#hasRelatedSynonym, respectively).

We have added two exact synonyms for BIOREALM's main class:

1. '*area of endemism* (an area occupied by at least two purportedly monophyletic taxa, Parenti & Ebach, 2009)' and
2. '*an area with monophyletic endemics*' and one related synonym: '*biogeographic area*' which is more general.

We added one related synonym: '*biogeographic area*' which is a more general and semantically ambiguous class.

3.2 | Individuals by class in BIOREALM

We defined 20 individuals (i.e. instances of a class) corresponding to the 20 Australian sub-regions described in Ebach et al. (2015). Each individual was asserted to be part of the ENVO: *Australasia ecoregion*. We added exact and related synonyms as described above, as well as URLs to related resources using the annotation database_cross_reference (http://www.geneontology.org/formats/oboInOwl#hasDbXref) (e.g. the World Wide Fund (www.wwf.org), Ecoregions® (https://ecoregions.appspot.com/) and the Interim Biogeographic Regionalisation for Australia v7 https://data.gov.au/data/dataset/interim-biogeographic-regionalisation-for-australia-ibra-version-7-regions). In each instance, any monophyletic endemic organism noted is linked to its taxon profile in the NCBI taxonomy database. A table summarizing the entire compendium of linked entities is provided as a supplementary file and is referenced in this document when the number of species is too numerous to be fully detailed in the text.

3.2.1 | Adelaide (BIOREALM_00000028)

This biogeographic area is defined as 'an instance of ENVO: ecosystem which contains endemic populations of 14 species of *Acacia*, 3 species of *Banksia* and 20 species of eucalyps' (see Table S1 in Supplementary information for full details).

Adelaide (BIOREALM_00000028) has one exact synonym which is 'Adelaide Cracraft 1991' from the Austral Bioregionalisation Atlas (https://aba.myspecies.info/nontaxonomy/term/28) and five related synonyms: 'Southeastern Australia (AA1207)' from WWF, 'Naracoorte woodlands' from Ecoregions® and 'Murray Darling Depression (MDD)', 'Naracoorte coastal Plain (NCP)', 'Victorian Midlands (VIM)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.2 | Arnhem Land (BIOREALM_00000029)

This biogeographic area is defined as 'an instance of ENVO: ecosystem which contains endemic populations 23 species of *Acacia* and 26 species of eucalyps' (see Supp Table S1 in Supplementary information for full details). Arnhem land (BIOREALM_00000029) has an exact synonym: 'Arnhem Land Cracaft 1991' (https://aba.myspecies.info/nontaxonomy/term/22) and 10 related synonyms: 'Northern Australia (AA0701') from WWF, 'Arnhem Land tropical savanna' from Ecoregions® and 'Darwin coastal (DAC)', 'Daly Basin (DAB)', 'Pine creek (PCK)', 'Arnhem Plateau (ARP)', 'Arnhem coastal (ARC)', 'Central Arnhem (CEA)', 'Gulf Fall' and 'Uplands (GFU)', 'Gulf coastal (GUC)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.3 | Atherton Plateau (BIOREALM_00000024)

This biogeographic area is defined as 'an instance of ENVO: ecosystem, which contains endemic populations of *Acacia spirorbis* (NCBI:txid1280825), *Banksia plagiocarpa* (NCBI:txid199787) and *Eucalyptus paedoglauca* (NCBI:txid1711454)'. Atherton Plateau (BIOREALM_00000024) has an exact synonym: 'Atherton plateau Cracaft 1991' (https://aba.myspecies.info/nontaxonomy/term/16) and three related synonyms: 'Northeastern Australia (AA0705)' from WWF, 'Queensland tropical rain forests' from Ecoregions® and 'Wet Tropics (WET)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.4 | Cape York Peninsula (BIOREALM_00000025)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110), which contains endemic populations of 2 species of *Banksia*, 9 species of *Acacia* and 28 species of eucalyps' (see Table S1 in Supplementary information for full details). Cape York Peninsula has (BIOREALM_00000025) an exact synonym: 'Cape York Peninsula Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/17) and three related synonyms: 'Northeastern Australia (AA0703)' from WWF, 'Cape York Peninsula tropical savanna' from Ecoregions® and 'Cape York Peninsula' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.5 | Central desert (BIOREALM_00000005)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 28 species of *Acacia* and 13 species of eucalyps' (see Table S1 in Supplementary information for full details). Central desert (BIOREALM_00000005) has an exact synonym: 'Central Desert Ebach et al., 2015' (https://aba.myspecies.info/nontaxonomy/term/87) and four related synonyms: 'Northeastern Australia (AA0707)' from WWF, 'Mitchell Grass Downs' from Ecoregions® and 'Mitchell Grass Downs (MGD)' and 'Desert Uplands (DEU)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.6 | Central Queensland (BIOREALM_00000023)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 3 species of *Acacia* (*Acacia argyrodendron* (NCBI:txid1174723), *Acacia faucium* (NCBI:txid1174792), *Acacia tephrina* (NCBI:txid1173666) and 9 species of eucalyps (*Corymbia blakei* (NCBI:txid1711099), *Corymbia brachycarpa* (NCBI:txid368742), *Corymbia lamprophylla* (NCBI:txid1711113), *Corymbia plena* (NCBI:txid1711120), *Eucalyptus exilipes* (NCBI:txid1711293), *Eucalyptus farinosa* (NCBI:txid1711296), *Eucalyptus quadricostata* (NCBI:txid1711497), *Eucalyptus similis* (NCBI:txid666056), *Eucalyptus whitei* (NCBI:txid1711591))'. Central Queensland (BIOREALM_00000023) has an exact synonym: 'Central Queensland Ebach et al., 2015' (https://aba.myspecies.info/nontaxonomy/term/88) and four related synonyms: 'Northeastern Australia (AA0707)' from WWF, 'Mitchell Grass Downs' from Ecoregions® and 'Mitchell Grass Downs (MGD)' and 'Desert Uplands (DEU)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.7 | Eastern desert (BIOREALM_00000007)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 80 species of *Acacia* and 47 species of eucalyps' (see Table S1 in Supplementary information for full details). Eastern Desert (BIOREALM_00000007) has an exact synonym: 'Eastern Desert Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/23) and 20 related synonyms: 'Eastern Central Australia (AA0802 and AA1308)' from WWF, 'Central Ranges xeric scrub', 'Simpson desert', 'Great Victoria desert', 'Tirari-Sturt stony desert', 'Eastern Australia mulga shrublands', 'Southeast Australia temperate savanna' and 'Flinders-Lofty montane woodlands' from Ecoregions® and 'Channel country (CHC)', 'Simpson Strzelecki Dunefields (SSD)', 'MacDonnell Ranges (MAC)', 'Finke (FIN)', 'Sturt Plateau (STP)', 'Gawler (GAW)', 'FLB Flinders Lofty Block', 'Broken Hill complex (BHC)', 'Mulga Lands (MUL)', 'Cobar Peneplain (COP)' and 'Darling Riverine Plains (DRP)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.8 | Eastern Queensland (BIOREALM_00000008)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 2 species of *Banksia*, 41 species of *Acacia* and 47 species of eucalyps' (see Table S1 in Supplementary information for full details). Eastern Queensland (BIOREALM_00000008) has an exact synonym: 'Eastern Queensland Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/10) and six related synonyms: 'Eastern Australia (AA0702)' and 'Northeastern Australia (AA0117)' from WWF, 'Brigalow tropical savanna', 'Eastern Australian temperate forests' from Ecoregions® and 'Central Mackay Coast (CMC)' and 'Brigalow Betl (BBN)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.9 | Eyre Peninsula (BIOREALM_00000019)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 15 species of *Acacia* and 14 species of eucalyps' (see Table S1 in Supplementary information for full details). Eyre Peninsula (BIOREALM_00000019) has an exact synonym: 'Eyre Peninsula Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/29) and three related synonyms: 'Southern central Australia (AA1203)' from WWF, 'Eyre and Yorke mallee' from Ecoregions® and 'Eyre Yorke Block (EYB)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.10 | Great Sandy desert interzone (BIOREALM_00000010)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 20 species of *Acacia* and 5 species of eucalyps' (see Table S1 in Supplementary information for full details). Great Sandy Desert interzone (BIOREALM_00000010) has an exact synonym: 'Great

Sandy Desert Interzone Ebach et al., 2015' (https://aba.myspecies.info/nontaxonomy/term/14) and three related synonyms: 'Northwestern Australia (AA1304)' from WWF, 'Great Sandy-Tanami desert' from Ecoregions® and 'Great Sandy Desert (GSD)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.11 | Hampton (BIOREALM_00000017)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 8 species of *Banksia*, 9 species of *Acacia* and 40 species of eucalyps' (see Table S1 in Supplementary information for full details). Hampton (BIOREALM_00000017) has an exact synonym: 'Hampton Ladiges et al. 2005' (https://aba.myspecies.info/nontaxonomy/term/30) and two related synonyms: 'Hampton mallee and woodlands' from Ecoregions® and 'Hampton (HAM)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.12 | Kimberly plateau (BIOREALM_00000011)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 12 species of *Acacia* and 13 species of eucalyps' (see Table S1 in Supplementary information for full details). Kimberly plateau (BIOREALM_00000011) has an exact synonym: 'Kimberley plateau Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/19) and six related synonyms: 'Northwestern Australia (AA0706)' from WWF, 'Kimberly tropical savanna' from Ecoregions® and 'Dampierland (DAL)', 'Central Kimberley (CEK)', 'Northern Kimberley (NOK)' and 'Victoria Bonaparte (VIB)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.13 | Nullarbor (BIOREALM_00000018)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of only one species: *Eucalyptus yalatensis* (NCBI:txid1711599)'. Nullabor (BIOREALM_00000018) has an exact synonym: 'Nullabor Ebach et al., 2015' (https://aba.myspecies.info/nontaxonomy/term/89) and two related synonyms: 'Southern Australia (AA1306)' from WWF, 'Nullarbor Plains xeric shrublands' from Ecoregions® and 'Nullarbor (NUL)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.14 | Pilbara (BIOREALM_00000013)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of *Eucalyptus prominens* (NCBI:txid1711486), *Eucalyptus ultima* (NCBI:txid1711575), *Acacia alexandri* (NCBI:txid334172), *Acacia gregorii* (NCBI:txid1174805)'. Pilbara (BIOREALM_00000013) has an exact synonym: 'Pilbara Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/25) and three related synonyms: 'Western Australia (AA1307)' from WWF, 'Pilbara shrublands' from Ecoregions® and 'Pilbarra (PIL)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.15 | Southeastern (BIOREALM_00000006)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 10 species of *Banksia*, 80 species of *Acacia* and 130 species of eucalyps' (see Table S1 in Supplementary information for full details). Southeastern has (BIOREALM_00000006) an exact synonym: 'Southeastern NSW Crisp et al. 2015' (https://aba.myspecies.info/nontaxonomy/term/18) and four related synonyms: 'East Coast of Australia (AA0402)' from WWF, 'Eastern Australian temperate forests' from Ecoregions® and 'NSW North Coast (NNC)' and 'Sydney Basin (SYB)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.16 | Southwest interzone (BIOREALM_00000016)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 7 species of *Banksia*, 80 species of *Acacia* and 130 species of eucalyps" (see Table S1 in Supplementary information for full details). Southeastern has (BIOREALM_00000016) an exact synonym: 'Southeastern NSW Crisp et al. 2015' (https://aba.myspecies.info/nontaxonomy/term/18) and three related synonyms: 'Western Australia (AA1201)' from WWF, 'Coolgardie woodlands' from Ecoregions® and 'Coolgardie (COO)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.17 | Southwestern (BIOREALM_00000015)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 121 species of *Banksia*, 68 species of *Acacia* and 95 species of eucalyps' (see Table S1 in Supplementary information for full details). Southwestern has (BIOREALM_00000015) an exact synonym: 'Southwestern Ebach et al., 2015' (https://aba.myspecies.info/nontaxonomy/term/86) and 13 related synonyms: 'South western Australia (AA1209)' and 'Southwestern coast of Australia (AA1205)' from WWF, 'Southwest Australia savanna', 'Esperance mallee', 'Southwest Australia woodlands' and 'Jarrah-Karri forest and shrublands' from Ecoregions® and 'Geraldton Sandplains (GES)', 'Avon

Wheatbelt (AVW)', 'Mallee (MAL)', 'Esperance plains (ESP)', 'Swan Coastal Plain (SCP)', 'Jarrah Forrest (JAF)' and 'Warren (WAR)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.18 | Tasmania (BIOREALM_00000026)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 3 species of *Acacia* and 13 species of eucalyps' (see Table S1 in Supplementary information for full details). Tasmania (BIOREALM_00000026) has an exact synonym: 'Tasmania Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/21) and 11 related synonyms: 'Tasmania South of Australia (AA0411)' from WWF; 'Tasmanian temperate rain forests', 'Tasmanian Central Highland forests', 'Tasmanian temperate forests' from Ecoregions® and 'Tasmanian Northern Slopes (TNS)', 'Tasmanian Central Highlands (TCH)', 'Tasmanian West (TWE)', 'Ben Lomond (BEL)', 'Tasmanian Northern Midlands (TNM)', 'Tasmanian South East (TSE)' and 'Tasmanian Southern Ranges (TSR)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.19 | Victoria (BIOREALM_00000021)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations of 1 species of *Banksia*, 4 species of *Acacia* and 19 species of eucalyps' (see Table S1 in Supplementary information for full details). Victoria (BIOREALM_00000021) has an exact synonym: 'Victoria Crisp et al. 1995' (https://aba.myspecies.info/nontaxonomy/term/20) and 11 related synonyms: 'Southeastern of Australia (AA0409)', 'Eastern part of Southern Coast of Australia (AA1208)' and 'Southeastern Australia (AA1001)' from WWF, 'Southeast Australia temperate forests', 'Eastern Australian temperate forests' and 'Eastern Australian temperate forests' from Ecoregions® and 'Southeastern Highlands (SEH)', 'Southeast Corner (SEC)', 'Australian Alps (AUA)', 'Victorian Midlands (VIM)', 'Southern Volcanic Plain (SVP)' and 'Southeast Coastal Plain (SCP)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

3.2.20 | Western desert (BIOREALM_00000014)

This biogeographic area is defined as 'an instance of ENVO: ecosystem (ENVO:01001110) which contains endemic populations 20 species of *Acacia* and 13 species of eucalyps' (see Table S1 in Supplementary information for full details). Western desert (BIOREALM_00000014) has an exact synonym: 'Western desert Cracraft 1991' (https://aba.myspecies.info/nontaxonomy/term/24) and 14 related synonyms: 'Western central Australia (AA1303)', 'Western Australia (AA1310)', 'Western coast of Australia (AA1301)' and 'Northwestern Australia (AA1304)' from WWF, 'Great Sandy-Tanami desert', 'Gibson desert', 'Carnarvon xeric shrublands' and 'Western Australian Mulga shrublands' from Ecoregions® and 'Carnavon (CAR)', 'Little Sandy Desert (LSD)', 'Gibson Desert (GID)', 'Central Ranges (VIM)', 'Great Victoria Desert (GVD)' and 'Murchison (MUR)' from the Interim Biogeographic Regionalisation for Australia v7 (IBRA).

4 | DISCUSSION

4.1 | Theoretical approach

The attributes we used to assign entities to classes are derived from comparative biogeography, which aims to discover and classify biogeographic patterns via congruence analysis, a process known as 'bioregionalization' (Ebach & Parenti, 2015; Michaux & Ung, 2021; Morrone, 2018; Ung et al., 2016; Ung, 2018). We translated the constraints used to construct bioregionalization models to the semantic/logical axioms linking classes in BIOREALM.

In biology, Elisabeth Lloyd, John Beatty and Paul Thompson (for example) presented semantic models for population genetics and Mendelian theory (Beatty, 1980, 1987; Lloyd, 1986, 1987, 1988; Thompson, 1983, 1986, 1987). Their approaches have been taken up by Gildenhuys (2013). Thompson (1986) viewed evolutionary theory as a family of interacting theories. He stated that '*the semantic conception of theories provides a framework within which a formalization of evolutionary theory understood as a family of interacting theories can be given [...]*'. Just as Thompson demonstrated that the theory of heredity interacts with the theory of natural selection, we emphasize that comparative biogeography interacts with systematic phylogenetics. Thus, a semantic model accounting for comparative biogeography is actually domain independent. As a consequence, such a model also accounts for phylogenetics.

In 1994, Anderson stated that:

> There are three major difficulties encountered by those dealing with the phenomenon of endemism: a semantic problem, the absence of a clear conceptual framework, and an analytical problem. Sydney Anderson, 1994: 451.

We propose a conceptual framework that will help to solve the analytical problem, that is defining (in the sense of delimiting or diagnosing) biogeographic areas that should be targets for conservation efforts through an ontology which addresses the semantic question.

We developed our ontological classes from the categorization of biogeographic units both by traditional phylogenetic trees and areagrams, which are widely used in phylogenetics and biogeography, respectively. The nested groups provided by phylogenetic trees account for relationships between monophyletic taxa. In these trees, taxa are formal objects linked by hierarchical relationships and constrained by the condition of monophyly, and the trees are constructed from observations performed in areas of endemism delimited a priori (Crother & Murray, 2011). Similarly, within comparative

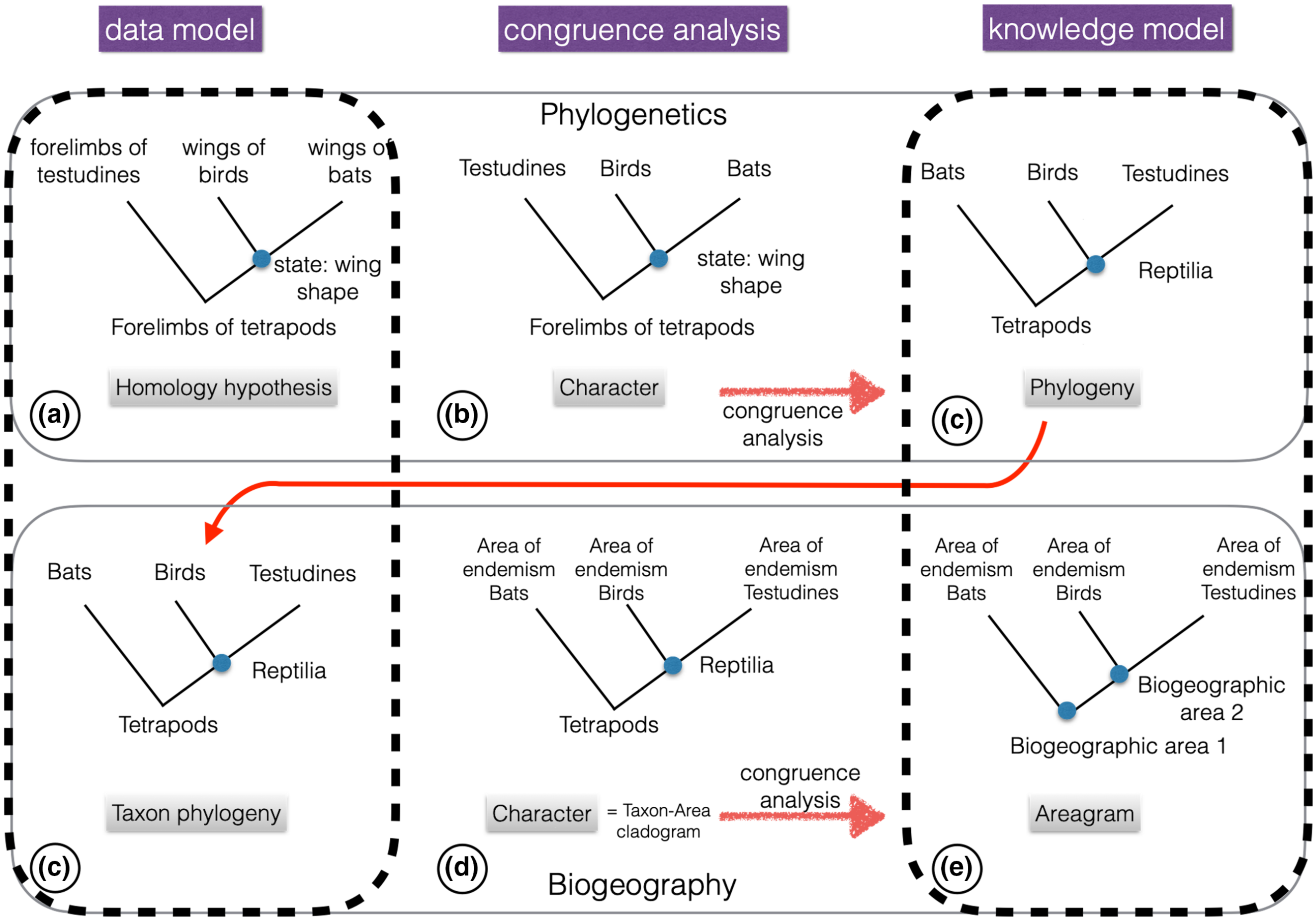

FIGURE 2 Analogy between Phylogenetics (a–c) and Biogeography (c–e) (modified from Ung, 2018). (a) homology hypothesis grouping morpho-anatomical features. (b) a character (mixed structure) grouping taxa bearing the features. (c) a phylogeny: (result of the congruence analysis). (c) a phylogeny standing for a biogeographic homology hypothesis about distributions of taxa. (d) a biogeographic character (mixed structure) grouping the areas of endemism (X, Y, Z) in which the taxa occur. (e) an areagram.

biogeography, areagrams model similarity and groupings of biogeographic areas, deduced from the inclusive groups yielded by cladograms of taxa in areas of endemism. Figure 2 illustrates the analogy between Phylogenetics and Biogeography.

The semantic model we describe here is a nested model for which a state of a character or an area of endemism is a particular case (i.e. a terminal) of a taxon or of a biogeographic area respectively. Biogeographic areas are thus deduced from the congruence analysis, and not only the endemism criterion. This approach brings new insights to conservation: the monophyletic biogeographic areas are clustered on the basis of the presence of specific combination of monophyletic endemic taxa that can be interpreted as a new biodiversity indicator for conservation.

In this semantic context, our contribution goes further and includes a reflection on the relational ontology of taxa and areas of endemism; this issue is essential for defining the term 'biodiversity' (taxic and biogeographic) more accurately. Consequently, this ontological reflection influences the way the methods of discovery of biodiversity and the conservation of biodiversity are implemented. This representation allows linking the data model and the knowledge model extracted from the congruence analysis. The model of embedded classes gives a relational ontology for areas of endemism and as such can be useful in the process of delimiting protected areas (i.e. areas of conservation). The associated methodology allows for a process-free analysis and gives new clues stemming from a mathematical representation of knowledge and a calculation procedure. Protected areas would be those with the highest number of taxa retrieved analytically rather than those based on simple enumeration.

Nelson and Platnick wrote in 1981: '*knowledge of general characters is useful for another reason: it is the basis for expectations about order in nature*' (Nelson & Platnick, 1981). They gave the example of the axial skeleton character in lampreys, sharks and perches. Considering this character, sharks and perches are grouped together on the basis of the presence of a vertebral column. They concluded that this pattern 'will prove true also for all other general characters that might subsequently be discovered'. Thus, we would expect that a general pattern of areas would prove true not only for the distributions of taxa used to reconstruct the general pattern, but also for other taxa that might subsequently be discovered. This means that delimiting protected areas based on particular taxa distributions could contribute to the protection of other unexpected taxa. Hence, due to its phylogenetic dimension, the conservation of taxic and biogeographic diversity contains a predictive dimension.

Faithfulness to a hierarchical model throughout the steps that relate the knowledge level (i.e. the conceptual level) to the empirical level (i.e. observations, measures) is important for internal consistency. Philosophers who formalized the semantic conception of theories stressed the importance of implementing isomorphic models (e.g. French, 2003; Frigg, 2002; Van Fraassen, 1994). We claim that the hierarchical model should be implemented at the character level, and since this model accounts for the interdisciplinarity of theories, the implementation should be the same for phylogenetics and biogeography.

4.2 | Technical implementation

Above, we have described how BIOREALM provides a semantically precise account of often ambiguous terminology in biogeography. Using the well-established best practices developed by the OBO Foundry and Library, BIOREALM constitutes a FAIR semantic resource and a highly expressive, machine-actionable means to represent knowledge in comparative biogeography. BIOREALM thus allows a platform to represent terminological and conceptual frameworks for comparative biogeographic analysis and natural classification of biogeographic areas (e.g. aligned to the bioregionalisation concept). This unification of model theoretic views of endemism with the semantic web offers new possibilities to communicate about the conservation of biogeographic units. Building on the work reported here, we will continue to develop BIOREALM with the purpose of integrating further biogeographic regionalisations such as the biogeographic regionalization of the Neotropical region (Morrone, 2014).

Importantly, BIOREALM does not claim to be an authoritative source of terminology: its content simply uses the conventions of knowledge representation and the semantic web to reduce ambiguity in complex terminology while providing a FAIR resource. As such, the resource can accommodate alternative classifications, clearly disentangling points of semantic overlap and divergence. The resource is thus open to contributions from the wider community, through GitHub's social coding interfaces (e.g. issues, forks and pull requests).

5 | CONCLUSION AND OUTLOOK

Biology (as a whole) is transitioning towards a data-intensive and being fully integrative discipline in need of greater use of mathematical and theoretical models (Murray, 2000) to advance research in common frameworks. In this context, new conservation policies arising through comparative biogeography also require structured, unambiguous and traceable interaction with theoretical constructs. BIOREALM provides one resource to support this transition: a standardized and semantically controlled representation of regions with unique combinations of monophyletic endemics, and with high conservation value.

In BIOREALM, we have created a FAIR and computationally actionable ontology as an implementation of discussions on the metaphysical nature of areas of endemism for comparative biogeography (Crother & Murray, 2011, 2013). Using a model of embedded (i.e. hierarchical) classes structured by the monophyly constraint, we reconcile elements of Crother and Murray's metaphysical Ontology on areas of endemism through formalization via logics and mathematics (Guarino et al., 1995).

More practically, BIOREALM provides support to semantically describe and align data and databases in an open data and open model following the FAIR principles (Wilkinson et al., 2016). Further developments will focus on linking descriptions of biomes and other ecological units related to those biogeographic areas to extend and enhance both BIOREALM and ENVO. Moreover, we seek to apply these ontologies to augment software such as the Global Biotic Interactions platform (https://www.globalbioticinteractions.org/) and more generic use of semantic web resources to annotate, mobilize and analyse data. The use of ontologies sensu *largo* is part of a global solution, not a solution in itself but helps reaching the goal: facilitating clear data and information flow for enhanced decision-making processes to better conserve the biodiversity and the planet.

AUTHOR CONTRIBUTIONS

V.U. and P.L.B. conceived the ideas. P.L.B. installed the informatic environment. V.U. integrated the data and wrote the manuscript. Both authors reviewed the manuscript.

ACKNOWLEDGEMENTS

The authors would like to thank Guillaume Lecointre and Christopher Hoagstrom for reviewing and commenting a previous version of this article and Malte Ebach and David Williams for fruitful discussions and Bernard Michaux for his revision of this manuscript. The authors would like to thank Lynne Parenti, two anonymous reviewers and Brian Crother for their great revisions and suggestions. All comments greatly helped to improve this manuscript. Finally, the authors would like to thank Aurélien Miralles for the design of BIOREALM logo. No specific permits were needed to carry out the work.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY STATEMENT

BIOREALM is accessible at https://github.com/EnvironmentOntology/BIOREALM.

ORCID

Visotheary Ung https://orcid.org/0000-0002-4049-0820

REFERENCES

Andersen, S. (1994). Area and endemism. *The Quarterly Review of Biology*, *69*(4), 451–471.

Beatty, J. (1980). What's wrong with the received view of evolutionary theory? *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *2*, 397–426.

Beatty, J. (1987). On behalf of the semantic view. *Biology and Philosophy*, *2*, 17–23.

Bellwood, D. R., Renema, W., & Rosen, B. R. (2012). Biodiversity hotspots, evolution and coral reef biogeography: A review (Chapter 9. In D. J. Gower, K. G. Johnson, J. E. Richardson, B. R. Rosen, L. Rüber, & S. T. Williams (Eds.), *Biotic evolution and environmental change in Southeast Asia* (pp. 216–245). Cambridge University Press.

Bonn, A., Rodrigues, A. S. L., & Gaston, K. (2002). Threatened and endemic species: Are they good indicators of patterns of biodiversity on a national scale? *Ecology Letters*, *5*, 733–741.

Butchart, S. H. M., Walpole, M., Collen, B., Van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. N., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., … Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*, *328*, 1164–1168.

Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J., & Lewis, S. E. (2013). The environment ontology: Contextualising biological and biomedical entities. *Journal of Biomedical Semantics*, *4*, 43.

Camillo, M., & Peter, S. (2011). Ongoing global biodiversity loss and the need to move beyond protected areas: A review of the technical and practical shortcomings of protected areas on land and sea. *Marine Ecology Progress*, *434*, 251–266.

Cardillo, M., & Pratt, R. (2013). Evolution of a hotspot genus: Geographic variation in speciation and extinction rates in banksia (Proteaceae). *BMC Evolutionary Biology*, *13*, 155.

Ceballos, G., Ehrlich, P., Barnosky, A., García, A., Pringle, R., & Palmer, T. (2015). Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*(5), e1400253.

Ceballos, G., Ehrlich, P. R., & Raven, P. H. (2020). Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proceedings of the National Academy of Sciences USA*, *117*, 13596–13602.

Cowie, R. H., Bouchet, P., & Fontaine, B. (2022). The sixth mass extinction: Fact, fiction or speculation? *Biological Review*, *97*(2), 640–663.

Craigie, I., Baillie, J. M., Balmford, A., Carbone, C., Collen, B., Green, R., & Hutton, J. (2010). Large mammal population declines in Africa's protected areas. *Biological Conservation*, *143*(9), 2221–2228.

Crother, B., & Murray, C. (2011). Ontology of areas of endemism. *Journal of Biogeography*, *38*, 1009–1015.

Crother, B., & Murray, C. (2013). Parsimony analysis of endemism under the "areas of endemism as individuals" thesis. *Cladistics*, *29*(6), 571–573.

Dudley, N. (2008). *Guidelines for Applying Protected Area Management Categories*. IUCN.

Dudley, N., Groves, C., Redford, K., & Stolton, S. (2014). Where now for protected areas? Setting the stage for the 2014 world parks congress. *Oryx*, *48*(4), 496–503.

Dudley, N., & Stolton, S. (2008). *Defining protected areas: An international conference in Almeria, Spain* (p. 220). IUCN.

Ebach, M. C., Gonzalez-Orozco, C. E., Miller, J. T., & Murphy, D. J. (2015). A revised area taxonomy of phytogeographical regions within the Australian Bioregionalisation atlas. *Phytotaxa*, *208*, 261–277.

Ebach, M. C., & Parenti, L. R. (2015). The dichotomy of the modern bioregionalization revival. *Journal of Biogeography*, *42*, 1801–1808.

French, S. R. D. (2003). A model-theoretic account of representation (or, I Don't know much about art…but I know it involves isomorphism). *Philosophy of Science*, *70*(5), 1472–1483.

Frigg, R. (2002). Models and representation: Why structures are not enough. In *Measurement in physics and economics discussion paper series DP MEAS 25/02*. London School of Economics.

Giakoumi, S., McGowan, J., Mills, M., Beger, M., Bustamante, R. H., Charles, A., Christie, P., Fox, M., Garcia-Borboroglu, P., Gelcich, S., Guidetti, P., Mackelworth, P., Maina, J. M., McCook, L., Micheli, F., Morgan, L. E., Mumby, P. J., Reyes, L. M., White, A., … Possingham, H. P. (2018). Revisiting "success" and "failure" of marine protected areas: A conservation scientist perspective. *Frontiers in Marine Science*, *5*, 223.

Gildenhuys, P. (2013). Classical population genetics and the semantic approach to scientific theories. *Synthese*, *190*, 273–291.

Gill, D. A., Mascia, M. B., Ahmadia, G. N., Glew, L., Lester, S. E., Barnes, M., Craigie, I., Darling, E. S., Free, C. M., Geldmann, J., Holst, S., Jensen, O. P., White, A. T., Basurto, X., Coad, L., Gates, R. D., Guannel, G., Mumby, P. J., Thomas, H., … Fox, H. E. (2017). Capacity shortfalls hinder the performance of marine protected areas globally. *Nature*, *543*, 665–669.

Gillepsie, A. (2009). Defining internationally protected areas. *Journal of International Wildlife Law and Policy*, *12*, 229–247.

González-Orozco, C. E., Pollock, L. J., Thornhill, A. H., Mishler, B. D., Knerr, N., Laffan, S. W., Miller, J. T., Rosauer, D. F., Faith, D. P., Nipperess, D. A., Kujala, H., Linke, S., Butt, N., Külheim, C., Crisp, M. D., & Gruber, B. (2016). Phylogenetic approaches reveal biodiversity threats under climate change. *Nature Climate Change*, *6*, 1110–1114.

Guarino, N., Uniti, S., & Giaretta, P. (1995). Ontologies and knowledge bases. *Towards a terminological, clarification*. IOS Press.

Harold, A., & Mooi, R. D. (1994). Areas of endemism: Definition and recognition criteria. *Systematic Biology*, *43*(2), 261–266.

IPCC, Climate Change. (2014). Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change [Core Writing Team. In R. K. Pachauri & L. A. Meyer (Eds.), (p. 151). IPCC.

Kerr, J. T. (1997). Species richness, endemism, and the choice of areas for conservation. *Conservation Biology*, *11*(5), 1094–1100.

Kolbert, E. (2014). *The Sixth Extinction* (p. 336). An Unnatural History.

Lele, S., Wilshusen, P., Brockington, D., Seidler, R., & Bawa, K. (2010). Beyond exclusion: Alternative approaches to biodiversity conservation in the developing tropics. *Current Opinion in Environmental Sustainability*, *2*, 94–100.

Lloyd, E. A. (1986). Thinking about models in evolutionary theory. *Philosophica*, *37*, 87–100.

Lloyd, E. A. (1987). Confirmation of ecological and Evolutionary models. *Biology and Philosophy*, *2*, 277–293.

Lloyd, E. A. (1988). *The structure and confirmation of evolutionary theory*. Princeton University Press.

Michaux, B., & Ung, V. (2021). Biotectonics of Sulawesi: Principles, methodology, and area relationships. *Zootaxa*, *5068*, 451–484.

Morrone, J. J. (1994). On the identification of areas of endemism. *Systematic Biology*, *43*(3), 438–441.

Morrone, J. J. (2014). Biogeographical regionalisation of the Neotropical region. *Zootaxa*, *3782*, 1.

Morrone, J. J. (2018). The spectre of biogeographical regionalization. *Journal of Biogeography*, *45*(2), 282–288.

Murphy, D. J., Ebach, M. C., Miller, J. T., Laffan, S. W., Cassis, G., Ung, V., Thornhill, A. H., Kerr, N., & Tursky, M. L. (2019). Do phytogeographic patterns reveal biomes or biotic regions? *Cladistics*, *35*, 654–670.

Murray, J. D. (2000). Pattern formation in integrative biology—A marriage of theory and experiment. *Comptes Rendus de l'Académie Des Sciences de Paris, Life Sciences*, *323*, 5–14.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Nelson, G. J., & Platnick, N. L. (1981). *Systematics and Biogeography: Cladistics and Vicariance*. Columbia Univ Press.

Orme, C. D. L., Davies, R. G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V. A., Webster, A. J., Ding, T. S., Rasmussen, P. C., Ridgely, R. S., Stattersfield, A. J., Bennett, P. M., Blackburn, T. M., Gaston, K.

J., & Owens, I. P. F. (2005). Global hotspots of species richness are not congruent with endemism or threat. *Nature*, *436*, 1016–1019.
Parenti, L. R., & Ebach, M. C. (2009). *Comparative Biogeography: Discovering and Classifying Biogeographical Patterns of a Dynamic Earth*. University of California Press.
Régnier, C., Achaz, G., Lambert, A., Cowie, R., Bouchet, P., & Fontaine, B. (2015). Mass extinction in poorly known taxa. *Proceedings of the National Academy of Sciences*, *112*(25), 7761–7766.
Rife, A. N., Erisman, B., Sanchez, A., & Aburto-Oropeza, O. (2013). When good intentions are not enough … insights on networks of “paper park” marine protected areas. *Conservation Letters*, *6*, 200–212.
Sachs, J. D., Baillie, J. E. M., Sutherland, W. J., Armsworth, P. R., Ash, N., Beddington, J., Blackburn, T. M., Collen, B., Gardiner, B., Gaston, K. J., Godfray, H. C. J., Green, R. E., Harvey, P. H., House, B., Knapp, S., Kümpel, N. F., Macdonald, D. W., Mace, G. M., Mallet, J., … Jones, K. E. (2009). Biodiversity conservation and the millennium development goals. *Science*, *325*, 1502–1503.
Scriven, J. J., Woodall, L. C., Tinsley, M. C., Knight, M. E., Williams, P. H., Carolan, J. C., Brown, M. J. F., & Goulson, D. (2015). Revealing the hidden niches of cryptic bumblebees in Great Britain: Implications for conservation. *Biological Conservation*, *182*, 126–133.
Scriven, S. A., Hodgson, J. A., Colin, J., McClean, C. J., & Hill, J. K. (2015). Protected areas in Borneo may fail to conserve tropical forest biodiversity under climate change. *Biological Conservation*, *184*, 414–423.
Seligmann, P., Mittermeir, R. A., Da Fonseca, A. B., Gascon, C., Crone, N., Da Silva, J. M. C., Famolare, L., … Beehler, B. (2007). *Centers for Biodiversity Conservation*. CBC Ressources.
Solton, S., & Dudley, N. (2010). Arguments for protected areas: Multiple benefits for conservation and use. Earthscan ltd, Dunstan house, 14a St cross street, London EC1N 8XA, UK.
Thompson, P. (1983). The structure of evolutionary theory: A semantic approach. *Studies in History and Philosophy of Science*, *14*, 215–229.
Thompson, P. (1986). The interaction of theories and the semantic conception of evolutionary theory. *Philosophica*, *37*, 73–86.
Thompson, P. (1987). A defense of the semantic conception of evolutionary theory. *Biology and Philosophy*, *2*, 26–32.
Thornton, P. K., Ericksen, P. J., Herrero, M., & Challinor, A. J. (2014). Climate variability and vulnerability to climate change: A review. *Global Change Biology*, *20*, 3313–3328.
Ung, V. (2018). The need for standard protocols in bioregionalisation: Comments on "the spectre of biogeographical regionalization" by Morrone (2018). *Zootaxa*, *4532*(2), 296–300.
Ung, V., Zaragüeta i Bagils, R., & Williams, D. M. (2016). Comparative biogeography of Southeast Asia and the West Pacific region. *Biological Journal of the Linnean Society*, *117*, 372–385.
Van Fraassen, B. (1994). Interpretation of science: Science as interpretation. In J. Hilgevoord (Ed.), *Physics and Our View of the World* (pp. 169–187). Cambridge University Press.
Wilkinson, M., Dumontier, M., Aalbersberg, I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., … Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*, 160018.
Woodruff, D. S. (2010). Biogeography and conservation in Southeast Asia: How 2.7 million years of repeated environmental fluctuations affect today's patterns and future of the remaining refugial–phase biodiversity. *Biodiversity and Conservation*, *19*, 919–941.

BIOSKETCH

Visotheary Ung is a research engineer at CNRS (Centre National de la Recherche Scientifique). She is a project manager and develops research in comparative biogeography with a particular focus on Southeast Asia region. She has extended her research to ontologies with the goal of providing new insights to comparative biogeography by inter-connecting data and benefit from cross-fertilization of domains. She works towards developing new solutions to better impact and foster biodiversity conservation.

Pier Luigi Buttigieg is a Digital Knowledge Steward and Senior Data Scientist working in the Earth and Environment branch of the Helmholtz Metadata Collaboration. He develops advanced, operational grade, and internationally adopted knowledge representation (KR) technologies for ecology, planetary science and Sustainable Development. He applies these solutions in mobilizing and analysing complex and high-dimensional data sets, while leading community efforts to enhance global interoperability, particularly in the marine life sciences. His goal is to use data science to better bridge scientific data, information and knowledge with operations, applications, policy and society.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ung, V., & Buttigieg, P. L. (2023). BIOREALM—An ontology of comparative biogeography: New insights into the semantics of biodiversity conservation. *Journal of Biogeography*, *50*, 1576–1586. https://doi.org/10.1111/jbi.14631