

Chapter 1

Primer on Ontologies

Janna Hastings

Abstract

As molecular biology has increasingly become a data-intensive discipline, ontologies have emerged as an essential computational tool to assist in the organisation, description and analysis of data. Ontologies describe and classify the entities of interest in a scientific domain in a computationally accessible fashion such that algorithms and tools can be developed around them. The technology that underlies ontologies has its roots in logic-based artificial intelligence, allowing for sophisticated automated inference and error detection. This chapter presents a general introduction to modern computational ontologies as they are used in biology.

Key words Ontology, Knowledge representation, Bioinformatics, Artificial intelligence

1 Introduction

Examining aspects of the world to determine the nature of the entities that exist and their causal networks is at the heart of many scientific endeavours, including the modern biological sciences. Advances in technology have made it possible to perform large-scale high-throughput experiments, yielding results for thousands of genes or gene products in single experiments. The data from these experiments are growing in public repositories [1], and in many cases the bottleneck has moved from the generation of these data to the analysis thereof [2]. In addition to the sheer volume of data, as the focus has moved to the investigation of systems as a whole and their perturbations [3], it has become increasingly necessary to integrate data from a variety of disparate technologies, experiments, labs and even across disciplines. Natural language data description is not sufficient to ensure smooth data integration, as natural language allows for multiple words to mean the same thing, and single words to mean multiple things. There are many cases where the meaning of a natural language description is not fully unambiguous. Ontologies have emerged as a key technology going beyond natural language in addressing these challenges.

The most successful biological ontology (bio-ontology) is the Gene Ontology (GO) [4], which is the subject of this volume.

Ontologies are computational structures that describe the entities and relationships of a domain of interest in a structured computable format, which allows for their use in multiple applications [5, 6]. At the heart of any ontology is a set of entities, also called classes, which are arranged into a hierarchy from the general to the specific. Additional information may be captured such as domain-relevant relationships between entities or even complex logical axioms. These entities that are contained in ontologies are then available for use as hubs around which data can be organised, indexed, aggregated and interpreted, across multiple different services, databases and applications [7].

2 Elements of Ontologies

Ontologies consist of several distinct elements, including classes, metadata, relationships, formats and axioms.

2.1 Classes

The class is the basic unit within an ontology, representing a type of thing in a domain of interest, for example *carboxylic acid*, *heart*, *melanoma* and *apoptosis*. Typically, classes are associated with a unique identifier within the ontology's namespace, for example (respectively) CHEBI:33575, FMA:7088, DOID:1909 and GO:0006915. Such identifiers are semantics free (they do not contain a reference to the class name or definition) in order to promote stability even as scientific knowledge and the accompanying ontology representation evolve. Ontology providers commit to maintaining identifiers for the long term, so that if they are used in annotations or other application contexts the user can rely on their resolution. In some cases as the ontology evolves, multiple entries may become merged into one, but in these cases alternate identifiers are still maintained as secondary identifiers. When a class is deemed to no longer be needed within the ontology it may be marked as obsolete, which then indicates that the ID should not be used in further annotations, although it is preserved for historical reasons. Obsolete classes may contain metadata pointing to one or more alternative classes that should be used instead.

2.2 Metadata

Classes are usually associated with annotated textual information—metadata. The metadata associated with classes may include any associated secondary (alternate) identifiers and flags to indicate whether the class has been marked as obsolete. It may also include one or more synonyms; for example the synonyms of *apoptotic process* (a class in the GO) include *cell suicide*, *programmed cell death* and *apoptosis*. It further may include cross references to that class in alternative databases and web resources. For example, many Chemical Entities of Biological Interest (ChEBI) [8] entries

contain cross references to the KEGG resource [9], which represents those chemicals in the context of the biological pathways they participate in. Textual comments and examples of intended usage may be annotated. It is very important that each class include a clear definition, which provides enough information to pinpoint the meaning of the class and suggest its appropriate use—sufficiently distinguishing different classes in an ontology so that a user can determine which is the best to use for annotation. The definition of apoptosis offered by the Gene Ontology is as follows:

A programmed cell death process which begins when a cell receives an internal (e.g. DNA damage) or external signal (e.g. an extracellular death ligand), and proceeds through a series of biochemical events (signaling pathway phase) which trigger an execution phase. The execution phase is the last step of an apoptotic process, and is typically characterized by rounding-up of the cell, retraction of pseudopodes, reduction of cellular volume (pyknosis), chromatin condensation, nuclear fragmentation (karyorrhexis), plasma membrane blebbing and fragmentation of the cell into apoptotic bodies. When the execution phase is completed, the cell has died.

2.3 Relations

Classes are arranged in a hierarchy from the general (high in the hierarchy) to the specific (low in the hierarchy). For example, in ChEBI *carboxylic acid* is classified as a *carbon oxoacid*, which in turn is classified as an *oxoacid*, which in turn is classified as a *hydroxide*, and so on up to the root *chemical entity*, which is the most general term in the structure-based classification branch of the ontology.

Despite the hierarchical organisation, most ontologies are not simple trees. Rather, they are structured as *directed acyclic graphs*. This is because it is possible for classes to have multiple parents in the classification hierarchy, and furthermore ontologies include additional types of relationships between entities other than hierarchical classification (which itself is represented by *is_a* relations). All relations are directed and care must be taken by the ontology editors to ensure that the overall structure of the ontology does not contain cycles, as illustrated in Fig. 1.

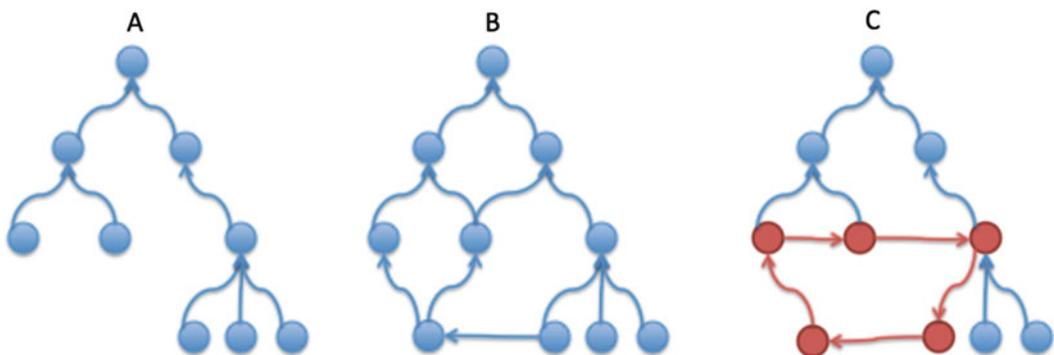


Fig. 1 (a) A simple hierarchical tree, (b) a directed, acyclic graph, (c) a graph that contains a cycle, indicated in red

Table 1
A selection of relationship types commonly used in bio-ontologies

| Relationship type | Informal meaning | Examples |
|-------------------|---|--|
| part_of | The standard relation of parthood. | A brain is part_of a body. |
| derives_from | Derivation holds between distinct entities when one succeeds the other across a temporal divide in such a way that a biologically significant portion of the matter of the earlier entity is inherited by the latter. | A zygote derives_from a sperm and an ovum. |
| has_participant | A relation that links processes to the entities that participate in them. | An apoptotic process has_participant a cell. |
| has_function | A relation that links material entities to their functions, e.g. the biological functions of macromolecules. | An enzyme has_function to catalyse a specific reaction type. |

A common relationship type used in multiple ontologies is **part_of** or **has_part**, representing composition or constitution. For example, in the Foundational Model of Anatomy (FMA) [10], *heart has_part aortic valve*. The Relationship Ontology (RO) defines several relationship types that are commonly used across multiple bio-ontologies [11], a selection of which is shown in Table 1.

In addition, specific ontologies may also include additional relationships that are particular to their domain. For example, GO includes biological process-specific relations such as **regulates**, while ChEBI includes chemistry-specific relationships such as **is_tautomer_of** and **is_enantiomer_of**.

The specification for a relationship type in an ontology includes a unique identifier, name and classification hierarchy, as for classes, as well as a specification whether the relationship is reflexive (i.e. A **rel** B if and only if B **rel** A) and/or transitive (if A **rel** B and B **rel** C then A **rel** C), and the name of the inverse relationship type if it exists. The same metadata as is associated with the classes in the ontology may also be associated with relationship types: alternative identifiers, synonyms, a definition and comments, and a flag to indicate if the relationship is obsolete.

2.4 Formats

Typically, ontologies are stored in files conforming to a specific file format, although there are exceptions that are stored in custom-built infrastructures. Ontologies may be represented in different underlying ontology languages, and historically there has been an evolution of the capability of ontology languages towards greater logical expressivity and complexity, which is mirrored by the advances in computational capacity (hardware) and tools. Biological ontologies such as the GO have historically been represented in the

human-readable Open Biomedical Ontologies (OBO) language,¹ which was designed specifically for the structure and metadata content associated with bio-ontologies, but in recent years there has been a move towards the Semantic Web standard Web Ontology Language (OWL)² largely due to the latter's adoption within a wider community and expansive tool support. Within OWL, specific standardised annotations are used to encode the metadata content of bio-ontologies as OWL annotations. However, the distinction has become cosmetic to some extent, as tools have been created which are able to interconvert between these languages [12], provided that certain constraints are adhered to.

2.5 Axioms

Within logic-based languages such as OWL, statements in ontologies have a definite logical meaning within a set-based logical theory. Classes have instances as members, and logical axioms define constraints on class definitions that apply to all class members. For example, the statement *carboxylic acid is_a carbon oxoacid* has the logical meaning that all instances of carboxylic acid are also instances of carbon oxoacid:

$$\forall x: \text{CarboxylicAcid}(x) \rightarrow \text{CarbonOxoacid}(x)$$

The logical languages underlying ontology technology are collectively called Description Logics [13]—in the plural because there are different variants with different levels of complexity. Some of the different ingredients of logical axioms that are available in the OWL language—quantification, cardinality, logical connectives and negation, disjointness and class equivalence—are explained in Table 2.

Like the carboxylic acid example above, each of these axiom types can be expressed as a logical statement. With these axioms, logic-based ontology reasoners are able to check for errors in an ontology. For example, if a class relation is quantified with ‘only’ such as the hydrocarbon example given in the table, which in logical language means

$$\forall x \forall y: \text{Hydrocarbon}(x) \wedge \text{hasPart}(x,y) \leftrightarrow \text{Hydrogen}(y) \vee \text{Carbon}(y)$$

and then if a subclass of hydrocarbon in the ontology has a **has_part** relation with a target other than a hydrogen or a carbon (e.g. an oxygen):

$$\text{Hydrocarbon}(a) \wedge \text{hasPart}(a,b) \wedge \text{Oxygen}(b)$$

that class will be detected as inconsistent and flagged as such by the reasoner.

¹ <http://www.cs.man.ac.uk/~horrocks/obo/>

² <http://www.w3.org/TR/owl2-overview/>

Table 2
Logical constructs available in the OWL language

| Language component | Informal meaning | Examples |
|--|---|--|
| Quantification: universal (only) or existential (some) | When specifying relationships between classes, it is necessary to specify a constraint on how the relationship should be interpreted: universal quantification means that for all relationships of that type the target has to belong to the specified class, while existential quantification means that at least one member of the target class must participate in a relationship of that type | <i>molecule has_part some atom</i> <i>hydrocarbon has_part only (hydrogen or carbon)</i> |
| Cardinality: exact, minimum or maximum | It is possible to specify the number of relationships with a given type and target that a class must participate in, or a minimum or maximum number thereof. | <i>human has_part exactly 2 leg</i> |
| Logical connectives: intersection (and) or union (or) | It is possible to build complex expressions by joining together parts using the standard logical connectives and, or. | <i>vitamin B equivalentTo (thiamin or riboflavin or niacin or pantothenic acid or pyridoxine or folic acid or vitamin B12)</i> |
| Negation (not) | In addition to building complex expressions using the logical connectives, it is possible to compose negations. | <i>tailless equivalentTo not (has_part some tail)</i> |
| Disjointness of classes | It is possible to specify that classes should not share any members. | <i>organic disjointFrom inorganic</i> |
| Equivalence of classes | It is possible to specify that two classes—or class expressions—are logically equivalent, and that they must by definition thus share all their members. | <i>melanoma equivalentTo (skin cancer and develops_from some melanocyte)</i> |

The end result—an ontology which combines terminological knowledge with complex domain knowledge captured in logical form—is thus amenable to various sophisticated tools which are able to use the captured knowledge to check for errors, derive inferences and support analyses.

3 Tools

Developing a complex computational knowledge base such as a bio-ontology (for example, the Gene Ontology includes 43,980 classes) requires tool support at multiple levels to assist the human knowledge engineers (curators) with their monumental task. For editing ontologies, a commonly used freely available platform is Protégé [14]. Protégé allows the editing of all aspects of an

ontology including classes and relationships, logical axioms (in the OWL language) and metadata. Protégé furthermore includes built-in support for the execution of automated reasoners to check for logical errors and for ontology visualisation using various different algorithms. Examples of reasoners that can be used within Protégé are Hermit [15] and Fact++ [16]. For the rapid editing and construction of ontologies, various utilities are available, such as the creation of a large number of classes in a single ‘wizard’ step. The software is open source and has a pluggable architecture, which allows for custom modular extensions. Protégé is able to open both OBO and OWL files, but it is designed primarily for the OWL language. An alternative editor specific to the OBO language is OBO-Edit [17]. Relative to Protégé, OBO-Edit offers more sophisticated metadata searching and a more intuitive user interface.

To browse, search and navigate within a wide variety of bio-ontologies without installing any software or downloading any files, the BioPortal web platform provides an indispensable resource [18] that is especially important when using terminology from multiple ontologies. Additional browsing interfaces for multiple ontologies include the OLS [19] and OntoBee [20]. Most ontologies are also supported by one or more browsing interfaces specific to that single ontology, and for the Gene Ontology the most commonly used interfaces are AmiGO [21] and QuickGO [22].

Large-scale ontologies such as the GO and ChEBI are often additionally supported by custom-built software tailored to their specific use case, for example embedding the capability to create species-specific ‘slims’ (subsets of terms of the greatest interest within the ontology for a specific scenario) for the GO, or cheminformatics support for ChEBI. As ontologies are shared across communities of users, an important part of the tool support profile is tools for the community to provide feedback and to submit additional entries to the ontology.

4 Applications

The purposes that are supported by modern bio-ontologies are diverse. The most straightforward application of ontologies is to support the structured annotation of data in a database. Here, ontologies are used to provide unique, stable identifiers—associated to a controlled vocabulary—around which experimental data or manually captured reference information can be gathered [23]. An ontology annotation links a database entry or experimental result to an ontology class identifier, which, being independent of the single database or resource being annotated, is able to be shared across multiple contexts. Without such shared identifiers for biological entities, discrepant ways of referring to entities tend to accumulate—different key words, or synonyms, or variants of

identifying labels—which significantly hinders reuse and integration of the relevant data in different contexts.

Secondly, ontologies can serve as a rich source of vocabulary for a domain of interest, providing a dictionary of names, synonyms and interrelationships, thereby facilitating text mining (the automated discovery of knowledge from text) [24], intelligent searching (such as automatic query expansion and synonym searching, an example is described in [25]) and unambiguous identification. When used in multiple independent contexts, such a common vocabulary can become additionally powerful. For example, uniting the representation of biological entities across different model organisms allows common annotations to be aggregated across species [26], which facilitates the translation of results from one organism into another in a fashion essential for the modern accumulation of knowledge in molecular biology. The use of a shared ontology also allows the comparison and translation entities from one discipline to another such as between biology and chemistry [27], enabling interdisciplinary tools that would be impossible computationally without a unified reference vocabulary.

While the above applications would be possible even if ontologies consisted only of controlled vocabularies (standardised sets of vocabulary terms), the real power of ontologies comes with their hierarchical organisation and use of formal inter-entity relationships. Through the hierarchy of the ontology, it is possible to annotate data to the most specific applicable term but then to examine large-scale data in aggregate for patterns at the higher level categories. By centralising the hierarchical organisation in an application-independent ontology, different sources of data can be aggregated to converge as evidence for the same class-level inferences, and complex statistical tools can be built around knowledge bases of ontologies combined with their annotations, which check for over-representation or under-representation of given classes in the context of a given dataset relative to the background of everything that is known [28] (for more information *see* Chap. 13 [29]). The knowledge-based relationships captured in the ontology can be used to assign quantitative measures of similarity between entities that would otherwise lack a quantifiable comparative metric [30] (for more information *see* Chap. 12 [31]). And the relationships between entities can be used to power sophisticated knowledge-based reasoning, such as the inference of which organs, tissues and cells belong to in anatomical contexts [32].

With all these applications in mind, it is no wonder that the number and scope of bio-ontologies have been proliferating over the last decades. The OBO Foundry is a community organisation that offers a web portal in which participating ontologies are listed [33]. The web portal currently lists 137 ontologies, excluding

obsolete records. Each of these ontologies has biological relevance and has agreed to abide by several community principles, including providing the ontology under an open license. Examples of these ontologies include ChEBI, the FMA, the Disease Ontology [34] and of course the Gene Ontology which is the topic of this book. In the context of the OBO Foundry, different ontologies are now becoming interrelated through inter-ontology relationships [35], and where there are overlaps in content they are being resolved through community workshops.

5 Limitations

Ontologies are a powerful technology for encoding domain knowledge in computable form in order to drive a multitude of different applications. However, they are not one-stop solutions for all knowledge representation requirements. There are certain limitations to the type of knowledge they can encode and the ways that applications can make use of that encoded knowledge.

Firstly, it is important to bear in mind that ontologies are based on logic. They are good at representing statements that are either true or false (categorical), but they cannot elegantly represent knowledge that is vague, statistical or conditional [36]. Classes that derive their meaning from comparison to a dynamic or conditional group (e.g. *the shortest person in the room*, which may vary widely) are also not possible to represent well within ontologies. It can be difficult to adequately capture knowledge about change over time at the class level, i.e. classes in which the members participate in relationships at one time and not at another, as including a temporal index for each relation would require ternary relations which neither the OBO nor the OWL language support.

Furthermore, although the underlying technology for representation and automated reasoning has advanced a lot in recent years, there are still pragmatic limits to ensure the scalability of the reasoning tools. For this reason, higher order logical statements, non-binary relationships and other complex logical constructs cannot yet be represented and reasoned with in most of the modern ontology languages.

Acknowledgements

The author was supported by the European Molecular Biology Laboratory (EMBL). Open Access charges were funded by the University College London Library, the Swiss Institute of Bioinformatics, the Agassiz Foundation, and the Foundation for the University of Lausanne.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

1. Marx V (2013) Biology: the big challenges of big data. *Nature* 498:255–260
2. Holzinger A, Dehmer M, Jurisica I (2014) Knowledge discovery and interactive data mining in bioinformatics – state-of-the-art, future challenges and research directions. *BMC Bioinformatics* 15(Suppl 6):I1
3. Pálsson BO (2015) *Systems biology: constraint-based reconstruction and analysis*. Cambridge University Press, Cambridge
4. Ashburner M, Ball CA, Blake JA et al (2000) Gene ontology: a tool for the unification of biology. *Nat Genet* 25:25–29
5. Stevens R, Goble CA, Bechhofer S (2000) Ontology-based knowledge representation for bioinformatics. *Brief Bioinform* 1(4):398–414
6. Bodenreider O, Stevens R (2006) Bio-ontologies: current trends and future directions. *Brief Bioinform* 7(3):256–274
7. Hoehdorf R, Schofield PN, Gkoutos GV (2015) The role of ontologies in biological and biomedical research: a functional perspective. *Brief Bioinform* (Advance Access) doi:10.1093/bib/bbv011
8. Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, Turner S, Swainston N, Mendes P, Steinbeck C (2015) ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res* (advance online access). doi:10.1093/nar/gkv1031
9. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42:D199–D205
10. Golbreich C, Grosjean J, Darmoni SJ (2013) The foundational model of anatomy in OWL 2 and its use. *Artif Intell Med* 57(2): 119–132
11. Smith B, Ceusters W, Klagges B, Köhler J, Kumar A, Lomax J, Mungall C, Neuhaus F, Rector AL, Rosse C (2005) Relations in biomedical ontologies. *Genome Biol* 6:R46
12. Tirmizi SH, Aitken S, Moreira DA, Mungall C, Sequeda J, Shah NH, Miranker DP (2011) Mapping between the OBO and OWL ontology languages. *J Biomed Semantics* 2 (Suppl 1):S3
13. Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider PF (2007) *The description logic handbook: theory, implementation and applications*, 2nd edn. Cambridge University Press, Cambridge
14. Protégé ontology editor. <http://protege.stanford.edu/>. Last Accessed Nov 2015
15. Shearer R, Motik B, Horrocks I (2008) HermiT: a highly-efficient OWL reasoner. In Proceedings of the 5th international workshop on owl: experiences and directions, Karlsruhe, Germany, 26–27 October 2008
16. Tsarkov D, Horrocks I (2006) Fact++ description logic reasoner: system description. In Proceedings of the third international joint conference on automated reasoning (IJCAR), pp 292–297
17. Day-Richter J, Harris M, Haendel M, The Gene Ontology OBO-Edit Working Group, Lewis S (2007) OBO-Edit—an ontology editor for biologists. *Bioinformatics* 23(16): 2198–2200

18. Noy NF, Shah NH, Whetzel PL, Dai B et al (2009) BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res* 37(Suppl 2):W170–W173
19. Côté RG, Jones P, Apweiler R, Hermjakob H (2006) The Ontology Lookup Service, a lightweight cross-platform tool for controlled vocabulary queries. *BMC Bioinformatics* 7:97
20. Xiang Z, Mungall C, Ruttenberg A, He Y (2011) OntoBee: a linked data server and browser for ontology terms. In Proceedings of the 2nd international conference on biomedical ontologies (ICBO), 28–30 July, Buffalo, NY, USA, pp 279–281
21. Carbon S, Ireland A, Mungall C, Shu S, Marshall B, Lewis S, The Amigo Hub and the Web Presence Working Group (2008) AmiGO: online access to ontology and annotation data. *Bioinformatics* 25(2):288–289
22. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R (2009) QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* 25(22):3045–3046
23. Blake J, Bult C (2006) Beyond the data deluge: data integration and bio-ontologies. *J Biomed Inform* 39(3):314–320
24. Rebholz-Schuhmann D, Oellrich A, Hoehndorf R (2012) Text-mining solutions for biomedical research: enabling integrative biology. *Nat Rev Genet* 13:829–839
25. Imam, F, Larson, S, Bandrowski, A, Grethe, J, Gupta A, Martone MA (2012) Maturation of neuroscience information framework: an ontology driven information system for neuroscience. In Proceedings of the formal ontologies in information systems conference, Frontiers in artificial intelligence and applications, vol 239, pp 15–28
26. Huntley RP, Sawford T, Mutowo-Meullenet P, Shypitsyna A, Bonilla C, Martin MJ, O'Donovan C (2015) The GOA Database: gene ontology annotation updates for 2015. *Nucleic Acids Res* 43(Database issue):D1057–D1063
27. Hill DP, Adams N, Bada M, Batchelor C et al (2013) Dovetailing biology and chemistry: integrating the Gene Ontology with the ChEBI chemical ontology. *BMC Genomics* 14:513
28. Tipney H, Hunter L (2010) An introduction to effective use of enrichment analysis software. *Hum Genomics* 4(3):202–206
29. Bauer S (2016) Gene-category analysis. In: Dessimoz C, Škunca N (eds) *The gene ontology handbook. Methods in molecular biology*, vol 1446. Humana Press. Chapter 13
30. Pesquita C, Faria D, Falcao AO, Lord P, Couto FM (2009) Semantic similarity in biomedical ontologies. *PLoS Comput Biol* 5(7):e1000443
31. Pesquita C (2016) Semantic similarity in the gene ontology. In: Dessimoz C, Škunca N (eds) *The gene ontology handbook. Methods in molecular biology*, vol 1446. Humana Press. Chapter 12
32. Osumi-Sutherland D, Reeve S, Mungall CJ, Neuhaus F, Ruttenberg A, Jefferis GS, Armstrong JD (2012) A strategy for building neuroanatomy ontologies. *Bioinformatics* 28(9):1262–1269
33. Smith B, Ashburner M, Rosse C, Bard J et al (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol* 25: 1251–1255
34. Kibbe WA, Arze C, Felix V, Mitraka E et al (2015) Disease ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Res* 43:D1071–D1078
35. Mungall CJ, Bada M, Berardini TZ, Deegan J, Ireland A, Harris MA, Hill DP, Lomax J (2011) Cross-product extensions of the gene ontology. *J Biomed Inform* 44(1):80–86
36. Schulz S, Stenzhorn H, Boeker M, Smith B (2009) Strengths and limitations of formal ontologies in the biomedical domain. *Rev Electron Comun Inf Inov Saude* 3(1):31–45