ICBO: International Conference on Biomedical Ontologies · Buffalo, NY, USA
Working with Multiple Biomedical Ontologies Workshop, July 26, 2011

Bridging Multiple Ontologies: Representation of the Liver Immune Response

Anna Maria Masci[1], Jeffrey Roach[2], Bernard de Bono[3], Pierre Grenon[3], Lindsay Cowell[4]

[1]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, USA
[2]Research Computing Center, University of North Carolina, Chapel Hill, NC, USA
[3]European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, UK
[4]Department of Clinical Sciences, University of Texas Southwestern Medical Center at Dallas, TX, USA

Abstract Chronic liver injury resulting in cirrhosis is the seventh leading cause of disease-related death in the USA. The liver immune response plays a key role in promoting, attenuating, and eventually resolving this severe condition. Although the liver immune response has been extensively investigated by conventional and *omic* approaches, gaps in knowledge still persist, and there is an urgent need for tools to facilitate the integration and analysis of these large, heterogeneous data sets. While ontologies are now accepted as an essential tool for data integration, no currently available ontology includes a representation of immunological reactions in the context of the liver. To address this need, we propose the Liver Immunology Ontology (LIO). LIO is being developed within the Open Biomedical Ontologies (OBO) Foundry framework, importing and linking relevant portions of orthogonal reference ontologies. LIO, is a novel tool for comprehensive analysis of liver immunology data sets, providing a valuable resource for the liver disease research community.

Keywords: Ontology, OBO Foundry, Immunology, Liver diseases, OWL.

1 Introduction

Over the last decades, an increasing amount of evidence has accumulated supporting a pivotal role for the liver immune system in the pathogenesis of severe acute and chronic hepatic disorders, including fibrosis and its final clinical stage, cirrhosis [1]. Critical to uncover the precise mechanisms of this pathogenesis is understanding the unique features of immune response regulation within the liver, in particular through interaction between parenchymal (hepatocytes) and non parenchymal cell types (liver sinusoidal cells, Kupffer cells, hepatic stellate cells, dendritic cells, T cells, natural killer cells, biliar epithelial cells) [2]. A considerable amount of conventional and *omics* findings has been generated from these cell types, but a coherent view of the structural and functional connections between them will require the integration and joint analysis of distinct and heterogeneous data sets.

The annotation of multiple bodies of data using common controlled vocabularies or 'ontologies' represents a powerful tool for integrating heterogeneous data sets [3]. Human readability together with machine accessibility make ontologies the perfect tool for organizing, retrieving, and integrating biological data. Although there are well known reference ontologies providing terms relevant to portions of the liver immunology domain, none of them provide a coherent, integrated representation of liver physiology and pathology. Current representations cover only narrow sub domains and are restricted to the assertion of classification and parthood relations within these sub domains. For example, the Foundational Model of Anatomy (FMA) covers normal anatomical entities [4]; the Cell Ontology (CL) covers normal cell types [5-7]; the Gene Ontology (GO) includes a small number of normal developmental and metabolic liver processes [8]; and the Disease Ontology (DO) includes a classification of liver diseases [9]. No currently available ontology includes a representation of immunological reactions in the context of the liver, of the pathological entities relevant to liver disease,

and of the relationships between the various relevant types of entities. To address this gap we are developing the Liver Immunology Ontology (LIO), an application ontology developed within the framework of the Open Biomedical Ontologies (OBO) Foundry [3].

In this paper, we briefly present the approach taken to developing LIO as a product of the selection, combination, and application-specific expansion of relevant communal reference ontologies. We give examples based on the current state of development and discuss briefly our more ambitious aims.

2 Methods

The development of LIO relies heavily on importing terms from existing reference ontologies in order to reduce duplication of effort, better ensure interoperability with other resources, and adhere to the OBO Foundry principles of ontology development best practice [3]. New terms are defined as needed and submitted to the appropriate reference ontology or maintained locally. To provide a coherent, integrated representation of liver physiology and pathology, including of the structural and functional connections between the various relevant entity types, LIO enriches the imported hierarchies by asserting additional relations, in particular between terms imported from distinct reference ontologies.

To date, the development of LIO has focused:

i) the types of cell found in the hepatic acinus,

ii) the types of molecule expressed by such cells in the hepatic acinus microenvironment,

iii) the processes in which these cells and molecules participate, including,

iv) the outcome of these processes.

To bring together the relevant ontological classes, LIO integrates the high-level classes of five OBO Foundry reference ontologies under the framework of the Basic Formal Ontology.

- The Foundational Model of Anatomy (FMA) (4): *anatomical entities*
- The Cell Ontology (CL) [7, 10]: *cell types*
- The Protein Ontology (PRO) [11]: *proteins*
- Gene Ontology Cellular Component (GO:CC) [8]: *cellular components*
- Gene Ontology Biological Process Ontology (GO: BP) [8]: *processes*

An extensive review of the primary literature was used to generate a list of all relevant entities. The National Center for Biomedical Ontology Bioportal (http://bioportal.bioontology.org/) and the Ontology Lookup Service (http://www.ebi.ac.uk/ ontology-lookup/) were then used to identify corresponding terms in each of the reference ontologies. When relevant terms were found, they were imported into LIO. When needed terms were not found, specific additions were made to LIO[1].

For each new term, LIO includes a formal definition and a Pub Med identifier supporting the definition. Where appropriate, terms and their definitions were submitted to the relevant reference ontology. The reference ontologies that serve as a source of terms for LIO all use the *is_a* relation as defined in the OBO Foundry Relation Ontology (RO) [12] to form their hierarchy, greatly facilitating the integration of their terms into a single, coherent hierarchy. To capture the structural and functional connections between the various entity types, the integrated LIO hierarchy is enriched with additional relations from the RO and its proposed extension (http://www.obofoundry.org/ro/), as described below.

3 Results

Using the approach described above, we are developing an ontology of the immune response induced during liver injury. Immune responses are in general highly context dependent and cannot be described without specifying a well-defined environment. The liver immune response in particular is characterized by local regulation mediated through interaction between parenchymal and non parenchymal cells. Thus, the focus of LIO is inclusion of the structural and functional relationships between these cell types and their components. This is being accomplished, as described above, by

[1] LIO is being encoded using the Web Ontology Language (OWL) (http://www.w3.org/TR/owl-features/) and Protégé (http://protege.stanford.edu/).

enriching a basic hierarchy, created by importing portions of reference ontologies, with a complex set of intra- and inter-ontological relations. The basic hierarchy of independent continuants is created through importing terms from the FMA, CL, GO CC, and PRO. The hierarchy is enriched to specify the relevant structural features of the hepatic acinus by asserting additional relations as described below. In the example triples that follow, ontology terms are shown in normal font. The source ontology is indicated as a subscript. RO and proposed RO relations are shown in italics.

Some of the anatomical structures that are relevant to the domain of LIO belong to a reference ontology of anatomy. For example, we can find in the FMA terms for the *space of Disse* (FMA ID:63162), *hepatic lamina* (FMA ID:14655) and *endothelium of hepathic sinusoids* (FMA ID:63137). To thoroughly characterize those aspects of the hepatic acinus relevant for describing liver immune responses, additional information is needed. For example, the FMA does not account for the fact that the space of Disse is an anatomical space between the hepatocyte lamina and the endothelium of hepatic sinusoid. Using *adjacent_to* relation from RO, we can capture some aspects of the spatial configuration:

Space of Disse$_{FMA}$ *adjacent_to* Hepatic Lamina$_{FMA}$

Space of Disse$_{FMA}$ *adjacent_to* Endothelium of Hepatic Sinusoids$_{FMA}$

LIO requires more than anatomy. In particular, further details of the hepatic acinus microenvironment can be represented by specifying the cellular and molecular components of the environment. This is done by relating anatomical entities (from the FMA) and cellular and molecular entities (from CL and PRO). For example, the space of Disse is characterized by the presence of hepatic stellate cells (cells playing a key role in liver injury, repair and inflammation):

Hepatic Stellate Cell$_{CL}$ *contained_in* Space of Disse$_{FMA}$

The cell types that participate in and regulate liver immune responses are characterized in LIO by specifying their molecular and cellular component parts. In general, the *has_part* relation is used. Cell-surface molecules, however, are specified using the *has_plasma_membrane_part* relation, which is defined as a sub-relation of the *has_part* relation. Examples include:

Platelet$_{CL}$ *has_part* platelet alpha granule$_{GOCC}$

Kupffer cell$_{CL}$ *has_plasma_membrane_part* CCR2$_{PRO}$

Platelet alpha granule$_{GOCC}$ *has_part* VLA-2$_{PRO}$

In addition to enriching the set of relations between independent continuants in LIO, we enrich the relations between independent continuants and the processes in which they participate. This is critical for capturing the functional relationships at play in a liver immune response. To enable precise characterization of the processes, they are defined as liver-specific processes. They are then related to general terms in GO BP using the *is_a* or *part_of* relation. The processes are then characterized using the *has_participant*, *has_output* and *occurs_in* relations, which are used to specify the independent continuants that are participants in a process, downstream processes that are triggered a process, and the locations in which processes occur, respectively. For example, apoptotic cell clearance is one of the main mechanisms involved in maintaining the homeostasis of the liver immune response. In the liver, Kupffer cell are the cell types involved in this process. This information is captured in LIO using the following set of triples:

Liver apoptotic cell clearance$_{LIO}$ *is_a* apoptotic cell clearance$_{GOBP}$

Liver apoptotic cell clearance$_{LIO}$ *has_participant* Kupffer cell$_{CL}$

Liver apoptotic cell clearance$_{LIO}$ *occurs_in* Space of Disse$_{FMA}$

Kupffer cell cytokine secretion$_{GOBP}$ *has_output* IFN gamma$_{PRO}$

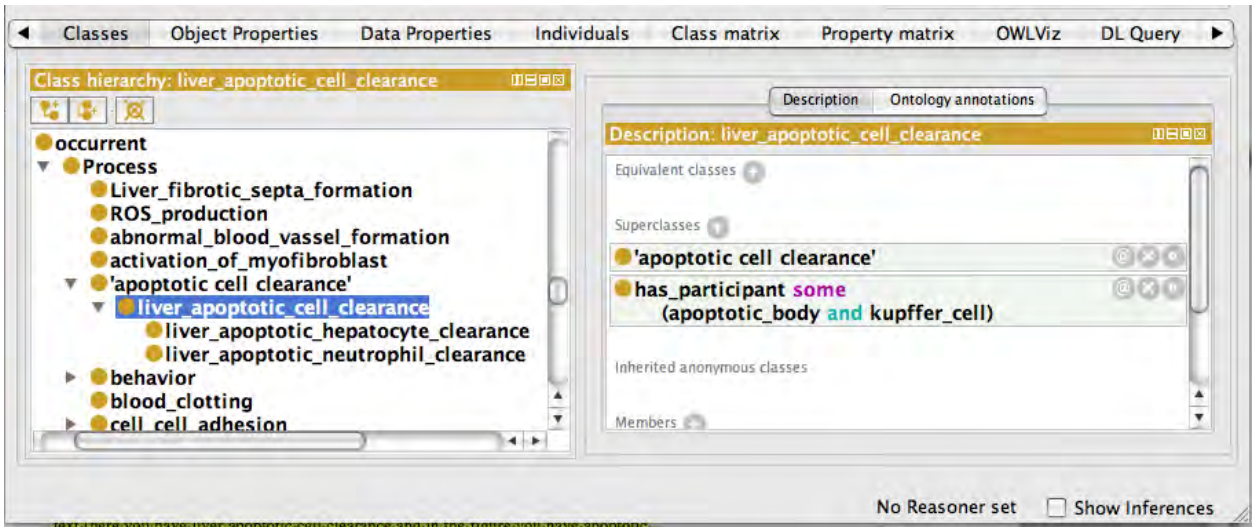

Figure 1. LIO Ontology. A screenshot of the ontology editor Protégé 4 showing the liver apoptotic cell clearance.

The resulting representation creates a bridge of knowledge that spans distinct reference ontologies, ontological types, and multiple levels of granularity.

Assertions in LIO are curated on the basis of the relevant primary literature and supported by associated PMIDs and the evidence code from the Evidence Code Ontology (http://www.obofoundry.org/cgi-bin/detail.cgi?id=evidence_code).

4 Discussion

By representing an immune response in the context of a specific organ, the LIO aims at expanding the resources available for knowledge representation in the domain of liver immunology. Our general approach can be summarized in three steps:

i) selection of relevant ontology terms from communal reference ontologies to bootstrap our application ontology,

ii) expansion to fill domain specific gaps after selection,

iii) combination of ontology terms obtained through selection and expansion to capture a more detailed description of the phenomena and processes that may require crossing multiple domains and levels of granularity.

The examples presented above are relatively simple illustrations of this approach.

They are examples of representation of ground facts. Our aim is to be able to use and refine the sort of knowledge so recorded to define more dynamically the further refinements in the LIO taxonomy. In particular, this is to support the definition of special subclasses of basic classes of liver physiological and immunological processes of LIO in application contexts; for example, kinds of connected pathologies under specific circumstances. Figure 2 illustrates such a use case in which LIO would allow capturing the difference in the physio-pathological responses to different concentrations of the bacterial lipopolysaccharide (LPS) in liver. The result is an explicit and specific linkage between physiological and pathological mechanism.

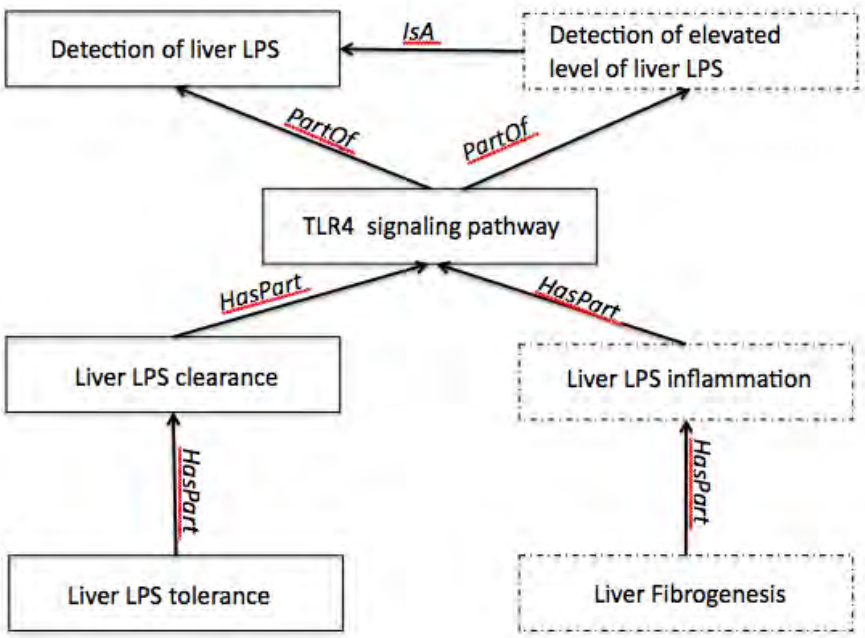

Figure 2. Example of physiological and pathological processes representation in LIO. Shown in the rectangular are processes involved in physiological and pathological conditions (*solid and dotted line, respectively*).
The arrows between boxes represent which entities are linked together. In italic are the specific relations used.

We believe that LIO may help illustrate the potential of ontological representation in supporting the scientific understanding of liver immunology. Such a potential, however, can be fully achieved only through the development of an application ontology integrating multiple reference ontologies. On the one hand, such integration allows for a detailed and tailored representation in a specific domain. On the other hand, because such a representation is based on more general reference ontologies, it preserves and benefits from their overall unifying character. By doing so, such a representation also preserves the capacity to be connected to other domains. This is because integration with other domains is facilitated from design at least, as long as the treatments of other application domains are also developed in a similar way.

The resulting benefits for connected research fields, such as the one concerned with liver immunology in the case of LIO, come from the explicit and specific character of the ontology and its capacity to integration from design. This is because bridging multiple ontologies creates a complex network of knowledge able i) to better reflect biological reality and also ii) to maintain a formal structure required for computational analysis, in general, and, in particular, the interoperability of high throughput technologies on which such analysis is based. Ways of evaluating this hypothesis are dual. First, the correctness of LIO and the adequacy of its coverage of the relevant domain can be evaluated in collaboration with domain experts. Second, the ability of LIO to facilitate data analysis can be evaluated in application driven tests and prototypes.

In the long term, our aim is to use LIO in order to support more comprehensive analysis of existing data than has heretofore been possible, resulting in novel insights, and the generation of new hypotheses. For example, in virtue of its structure, LIO supports an enhanced analysis of microarray data. This is because annotations using LIO would carry, if not no ambiguity at all, at least less ambiguities than context-insensitive annotations using the more general ontology.

Another application of LIO that we wish to explore is to support the biological annotation of relevant mathematical models and associated computer simulations. LIO provides, on the side of reality, the elements of an account of the liver's immune response: the objects involved (cells and molecules), the

environment in which they evolve (liver anatomy), and the relations between these objects and their environment. It provides therefore ontological resources readily conjoined to the mathematical modeling of the liver's immune response, both from the physiological and the pathological perspectives. At stake here is the increased rate of translation of modeling and simulation endeavors into clinical studies, for example, aimed at identifying genes involved in the progression and reversion of liver diseases.

The development of LIO is ongoing and much remains to do in the near future. Our next step will be concerned with the expansion of the representation of the immune response in relation to a variety of triggers such as, for example, infectious agents, alcohol abuse, obesity, autoimmunity and drugs. We believe that furthering the line of development sketched here and extending LIO to more connected domains and resources will provide a valuable resource, meeting the needs of the hepatic disease research community.

Acknowledgments

This work was supported by NIAID R01 AI077706 to LGC (AMM, PI subcontract to Duke University) and a Burroughs Wellcome Fund Career Award to LGC.

References

1. Tacke F, Luedde T, Trautwein C. Inflammatory pathways in liver homeostasis and liver injury. Clin Rev Allergy Immunol. 2009 Feb;36(1):4-12.
2. Crispe IN. The liver as a lymphoid organ. Annu Rev Immunol. 2009;27:147-63.
3. Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol. 2007 Nov;25(11):1251-5.
4. Rosse C, Mejino JL, Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform. 2003 Dec;36(6):478-500.
5. Bard J, Rhee SY, Ashburner M. An ontology for cell types. Genome Biol. 2005;6(2):R21.
6. Diehl AD, Augustine AD, Blake JA, Cowell LG, Gold ES, Gondre-Lewis TA, et al. Hematopoietic cell types: prototype for a revised cell ontology. J Biomed Inform. 2011 Feb;44(1):75-9.
7. Meehan TF, Masci AM, Abdulla A, Cowell LG, Blake JA, Mungall CJ, et al. Logical development of the cell ontology. BMC Bioinformatics. 2011;12:6.
8. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000 May;25(1):25-9.
9. Osborne JD, Flatow J, Holko M, Lin SM, Kibbe WA, Zhu LJ, et al. Annotating the human genome with Disease Ontology. BMC Genomics. 2009;10 Suppl 1:S6.
10. Masci AM, Arighi CN, Diehl AD, Lieberman AE, Mungall C, Scheuermann RH, et al. An improved ontological representation of dendritic cells as a paradigm for all cell types. BMC Bioinformatics. 2009;10:70.
11. Natale DA, Arighi CN, Barker WC, Blake JA, Bult CJ, Caudy M, et al. The Protein Ontology: a structured representation of protein forms and complexes. Nucleic Acids Res. 2011 Jan;39(Database issue):D539-45.
12. Smith B, Ceusters W, Klagges B, Kohler J, Kumar A, Lomax J, et al. Relations in biomedical ontologies. Genome Biol. 2005;6(5):R46.