OWL DL Ontology for classifications of Lipids

Hong-Sang Low, Christopher J.O. Baker, Alexander Garcia, Markus R. Wenk
International Conference on Biomedical Ontology
University at Buffalo, NY July 24-26, 2009
Motivation

- Semantic data integration is necessary for lipid research yet this is poorly achievable due to an absence of a single unified, consistent, and universally accepted lipid classification system.

- Lipid nomenclature is highly heterogeneous.
  - Not semantically explicit and very graphic dependent
  - Many conflicting nomenclatures and multiple synonyms
  - Lack of universal systematic nomenclature
Objective

- Formalize and represent lipid nomenclature & classification hierarchy in the Web Ontology Language (OWL-DL)

- Provide lipid definitions that are:
  - Semantically explicit
  - Independent of graphical descriptions
  - Amenable to inference and classification based reasoning

- To make available:
  - a systematic and formalized OWL-DL definitions of lipids for testing appropriateness of existing nomenclature to lipid structures.
  - Serve as a reusable standard for lipid researchers and the lipid bioinformatics community.
LIPIDMAPS definition

Hydrophobic or amphipathic small molecules that may originate entirely or in part by carbanion-based condensations of thioesters and / or by carbocation-based condensation of isoprene units
Lipids

- Lipids are defined by the presence of functional groups and other structural descriptors

**Mycolic Acid**

- Cyclopropane Group
- Cyclopropane Group
- Alpha-hydroxyl Acid Group
Analysis, reuse and extension of existing ontologies.

- Compliance to Basic Formal Ontology
Analysis, reuse and extension of existing ontologies.

CheBI

- Aims to promote the correct and consistent use of unambiguous biochemical terminology throughout the molecular database in EBI
- Not all relationship definitions from CheBI are relevant
- Lack of formal consistency
- Currently in a state of flux, final state uncertain
- Unsuitable for reuse in the context of lipid ontology
Analysis, reuse and extension of existing ontologies.

**Chemical Ontology**

- Originally applied to facilitate pharmacophore search and semantic comparison of small molecules
- OWL version available since 2007
- We re-use primarily the following axiom:

![Diagram showing relationships between Compound, PartOf, and Organic_Group]

- 32 functional group primitive classes found suitable to be re-used in Lipid Ontology
Analysis, reuse and extension of existing ontologies.

- Limit Compound to Lipid

Analysis, reuse and extension of existing ontologies.

Organic Group

Total no. of simple organic group = 95
Analysis, reuse and extension of existing ontologies.

**Organic Group**

Total no. of complex organic group = 23
Analysis, reuse and extension of existing ontologies.

**Ring System**

Total no. of ring system = 87
Analysis, reuse and extension of existing ontologies.

Incorporation of LIPIDMAPS hierarchy into the upper ontology

<table>
<thead>
<tr>
<th>Table</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of Classes</td>
<td>715</td>
</tr>
<tr>
<td>No. of Lipid Classes</td>
<td>492 UPDATED!</td>
</tr>
<tr>
<td>Primitive Lipid Classes</td>
<td>107 UPDATED!</td>
</tr>
<tr>
<td>Defined Lipid Classes</td>
<td>268 UPDATED!</td>
</tr>
<tr>
<td>Total No. of Restrictions</td>
<td>901</td>
</tr>
<tr>
<td>Total No. of Properties</td>
<td>41</td>
</tr>
<tr>
<td>DL Expressivity</td>
<td>ALCIQ(D)</td>
</tr>
</tbody>
</table>

Definition of lipid classes: DL Axioms

LC_Fatty_Acyl is a Lipid
((hasPart some Carboxylic_Acid_derivative_Group)
and (hasPart some Acyl.Chain)) or (hasPart some Alkyl.Chain)

LC_Fatty.acid is a LC_Fatty_Acyl
(hasPart some Carboxylic_Acid)
and (hasAcyl_Chain exactly 1)

LC_Mycolic.acid is a LC_Fatty.acid
(hasPart some Alpha-Hydroxy_Acid_Group) and
(hasMero(mycolic_Chain exactly 1)

LC_Docosanoid is a LC_Fatty_Acyl
(hasPart some Carboxylic_Acid) and (hasPart some Alcohol) and (hasPart some Alkenyl_Group)
and (hasPart some Cyclopentenone)
and (hasAcyl_Chain exactly 1);
hasPart only (Carboxylic_Acid or Alcohol or Alkenyl_Group or Cyclopentenone or Acyl.Chain);

LC_Fatty.acyl_derivative
is a LC_Fatty_Acyl
hasPart some alkyl.Chain

LC_Fatty.alcohol is a LC_Fatty.acyl_derivative
(hasPart some Alcohol) and (hasAlkyl.Chain exactly 1);
hasPart only (Alcohol or Alkyl.Chain);
Definition of lipid classes: DL Axioms

- As we travel down the hierarchy
  - Specify more specific organic group classes
  - Specify cardinality

- hasPart some Carboxylic_Acid_derivative_Group
- hasPart some Aldehyde
- hasAldehyde exactly 3

\[
\begin{align*}
&\text{R} & \text{R} \\
&\text{O} & \text{C} & \text{R} & \text{H}
\end{align*}
\]
Definition of lipid classes: DL

Axioms

Additional properties required prior to cardinality specification

Fatty Acyl

- Fatty acyls are a diverse lipid group synthesized by chain-elongation of an acetyl-CoA primer with malonyl-CoA/methylmalonyl-CoA groups.

((hasPart some Carboxylic_Acid_derivative_Group) and (hasPart some Acyl_Chain)) or (hasPart some Alkyl_Chain)
Docosanooid

LC_Fatty_Acyl
(hasPart some Carboxylic_Acid) and (hasPart some Alcohol) and (hasPart some Alkenyl_Group) and (hasPart some Cyclopentenone) and (hasAcyl_Chain exactly 1)

hasPart only (Carboxylic_Acid or Alcohol or Alkenyl_Group or Cyclopentenone or Acyl_Chain)
Fatty Alcohol

LC_Fatty_acyl_derivative
(hasPart some Alcohol) and (hasAlkyl_Chain exactly 1)
hasPart some Alkyl_Chain

hasPart only (Alcohol or Alkyl_Chain)
Mycolic acid

- Has parts some Carboxylic Acid and hasACYL_Chain exactly 1
- Has parts some Alpha-Hydroxy_Acid_Group
- Has part only (Alkenyl_Group or Alpha-hydroxy_Acid_Group or Cyclopropane or Carboxylic_Acid or Meromycolic_Chain)

New Mycolic Acid Classes

LC_Methoxy_mycolic_acid

LC_Omega-1_methoxy_mycolic_acid

LC_General_Methoxy_mycolic_acid

LC_General_Oxygenated_mycolic_acid

LC_Keto_mycolic_acid

LC_Epoxide_mycolic_acid

LC_Fatty_mycolic_acid

LC_General_methylated_mycolic_acid

LC_Mycolic_acid

LC_General_alpha_mycolic_acid

LC_Alpha_prime_mycolic_acids

LC_Alpha_1_mycolic_acid

LC_General_mycolic_acid

LC_Alpha_mycolic_acid

LC_Mycolic_acid is a LC_Fatty_mycolic_acid (hasPart some Alpha-Hydroxy_Acid_Group) and (hasMeromycolic_Chain exactly 1)

LC_General_alpha_mycolic_acid is a LC_Mycolic_acid hasPart some (Cyclopropane or Alkenyl_Group)

LC_Alpha_mycolic_acid is a LC_General_alpha_mycolic_acid hasPart some Cyclopropane; hasPart only (Alkenyl_Group or Alpha-Hydroxy_Acid_Group or Cyclopropane or Carboxylic_Acid or Meromycolic_Chain);
### New Classes of Mycolic Acids

<table>
<thead>
<tr>
<th>Structure</th>
<th>Class type of Mycolic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Structure" /></td>
<td>Alpha_mycolic_acid</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure" /></td>
<td>Alpha_prime_mycolic_acid</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure" /></td>
<td>Alpha_1_mycolic_acid</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure" /></td>
<td>Alpha_2_mycolic_acid</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure" /></td>
<td>Keto_mycolic_acid</td>
</tr>
<tr>
<td><img src="image6.png" alt="Structure" /></td>
<td>Epoxy_mycolic_acid</td>
</tr>
<tr>
<td><img src="image7.png" alt="Structure" /></td>
<td>Wax_ester_mycolic_acid</td>
</tr>
<tr>
<td><img src="image8.png" alt="Structure" /></td>
<td>Methoxy_mycolic_acid</td>
</tr>
<tr>
<td><img src="image9.png" alt="Structure" /></td>
<td>Omega-1_methoxy_mycolic_acid</td>
</tr>
</tbody>
</table>
Future Work

- Reflect changes for sphingolipid headgroup in lipid maps
- Formalize definitions of organic groups and rings
- Unresolved parenthood issues with quinone ring vs cyclic terpenoid
- Owl 2.0, Qualified Cardinality Restriction
- Re-adapting the ontology so that it becomes amenable to the application for integration of lipid databases
The role of Lipid Ontology

- **LIPID MAPS**
  - Nomenclature
    - non OBO compliant
  - ChEBI
    - towards OBO compliance

- **OBO Bio-Ontologies**
  - OBO compliant

- **Lipid Ontology**

- **Lipidomics**

Available for testing the appropriateness of existing nomenclature to classify lipid structures.

http://www.lipidprofiles.com/LipidOntology/LiPrO-02042009.owl

Lipid research is increasingly integrated within systems level biology such as lipidomics where lipid classification is required before appropriate annotation of chemical functions can be applied. The ontology describes the LIPIDMAPS nomenclature classification explicitly using description logics (OWL-DL). Lipid classes are organized hierarchically with the super-classes restricted by generic necessary conditions. More specific necessary conditions are used to define membership requirements for sub classes of lipid according to appropriate functional groups.
Summary

- The first reported OWL-DL Lipid Ontology covering 450+ classes of lipid nomenclature and suitable for classification of lipids, based on functional groups and other structural descriptors.

- At present, the Lipid Ontology has a total of 715 Classes, 901 Restrictions, 41 Properties with DL Expressivity of ALCHIQ(D).

- Successful representation of LIPIDMAPS hierarchy using DL Axioms.

- Reconcile LIPIDMAPS inconsistency in classification of lipids without Acyl group in Fatty Acyl without violating OWL-DL consistency.

- Added 9 novel classes of lipid, namely the mycolic acids extended from LIPIDMAPS original hierarchy.

- Ontology available from OBO, BioPortal and Lipid Profiles for reuse in Lipidomics – where collection/organization of lipid data via a “system biology” requires a comprehensive classification, nomenclature and chemical representation system capable of representing diverse classes of lipids that exist in nature.
Acknowledgements

University of New Brunswick, Canada
• Christopher J.O. Baker
University of Bremen, Germany
• Alexander Garcia
National University of Singapore
• Markus R. Wenk

Funding
• NUS Office of Life Science (R-183-000-607-712),
• ARF (R-183-000-160-112),
• BMRC A*STAR (R-183-000-134-305),
• Singapore NRF under CRP award No. 2007-04,
• NBIF, New Brunswick, Canada.