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
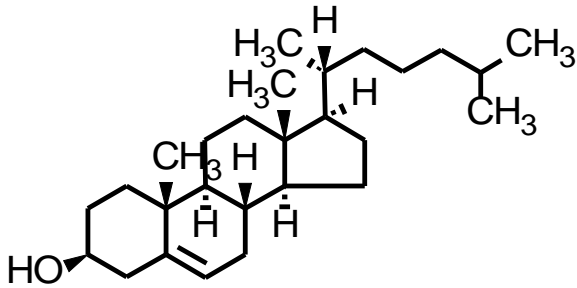
Lipids

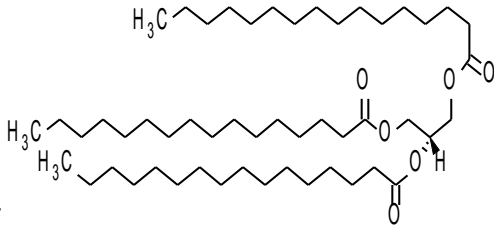Sterol lipid
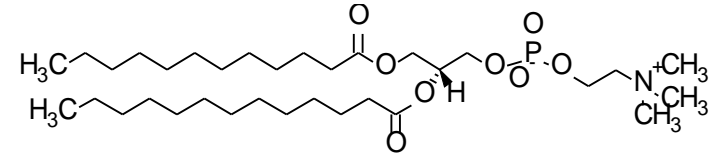
Glycerophospholipid

Glycerolipid

Fatty Acyl

Saccharolipid
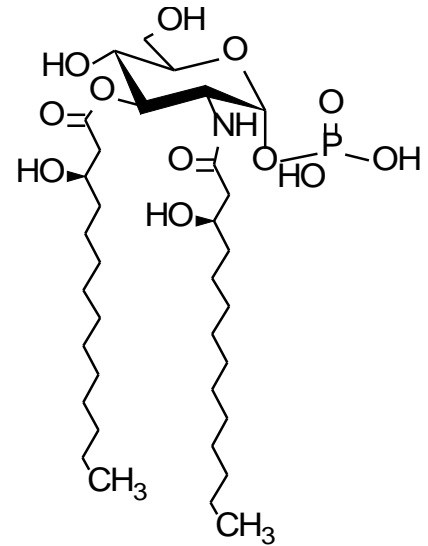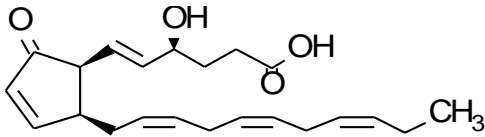
LIPIDMAPS definition

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
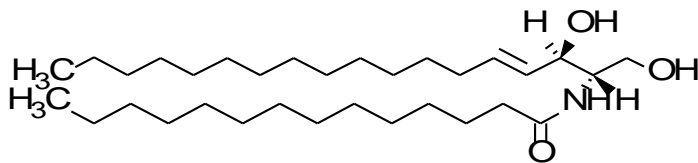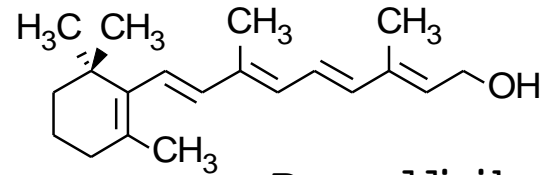
Sphingolipid
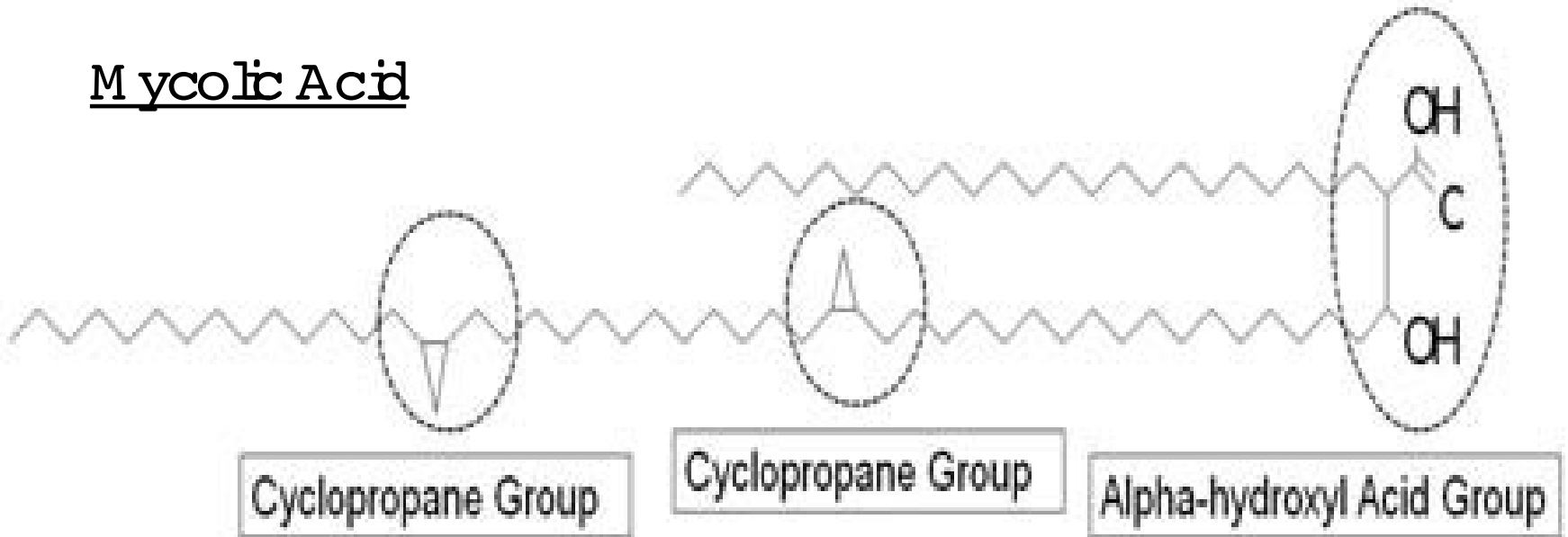
Prenol lipid

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
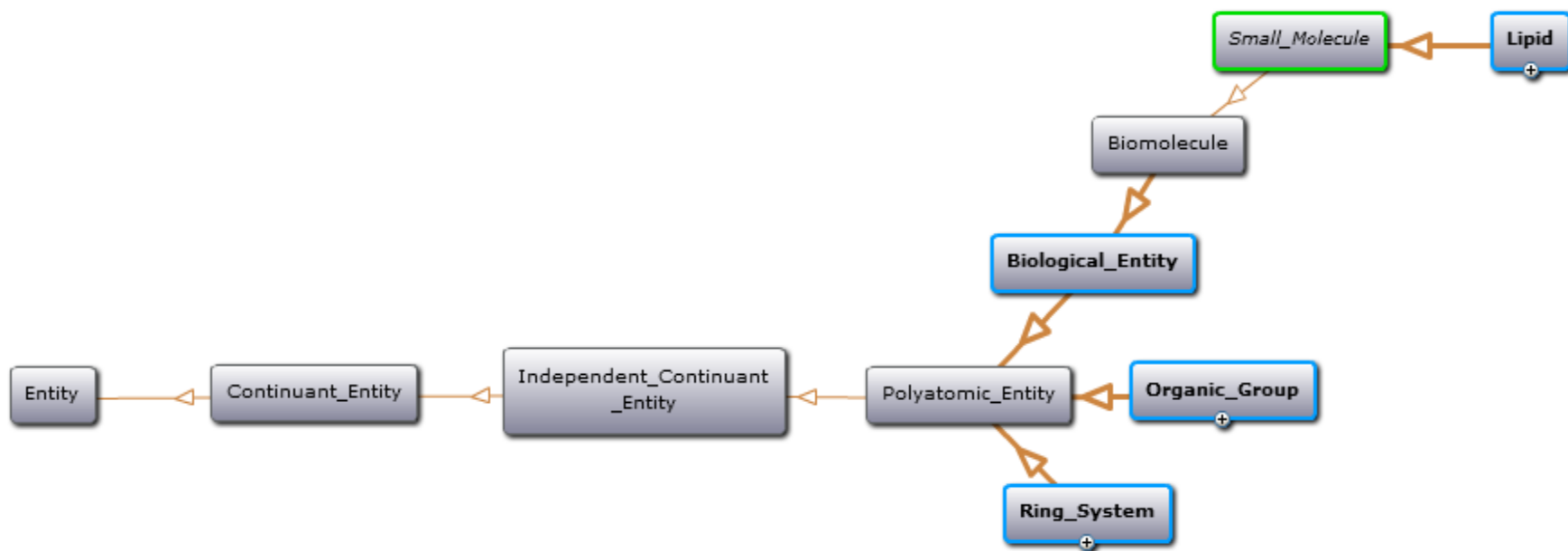
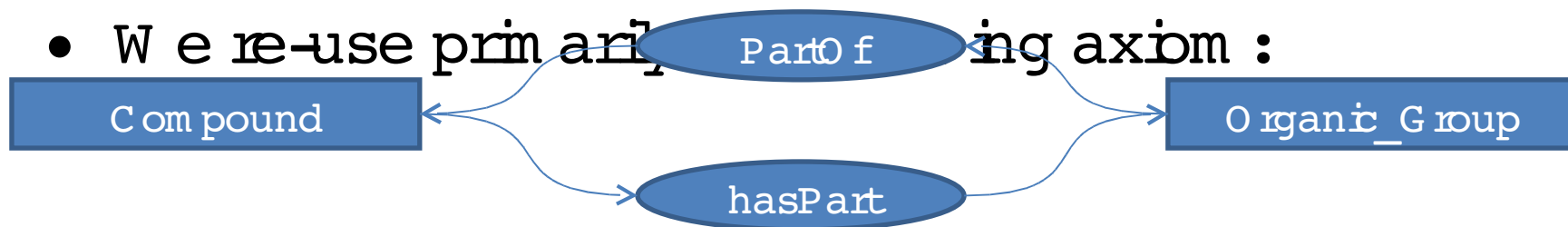# Analysis, reuse and extension of existing ontologies.

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 re-use in the context of lipid ontology

Analysis, reuse and extension of existing ontologies.

ChemicalOntology

- Originally applied to facilitate pharmacophore search and semantic comparison of small molecules

- OWL version available since 2007
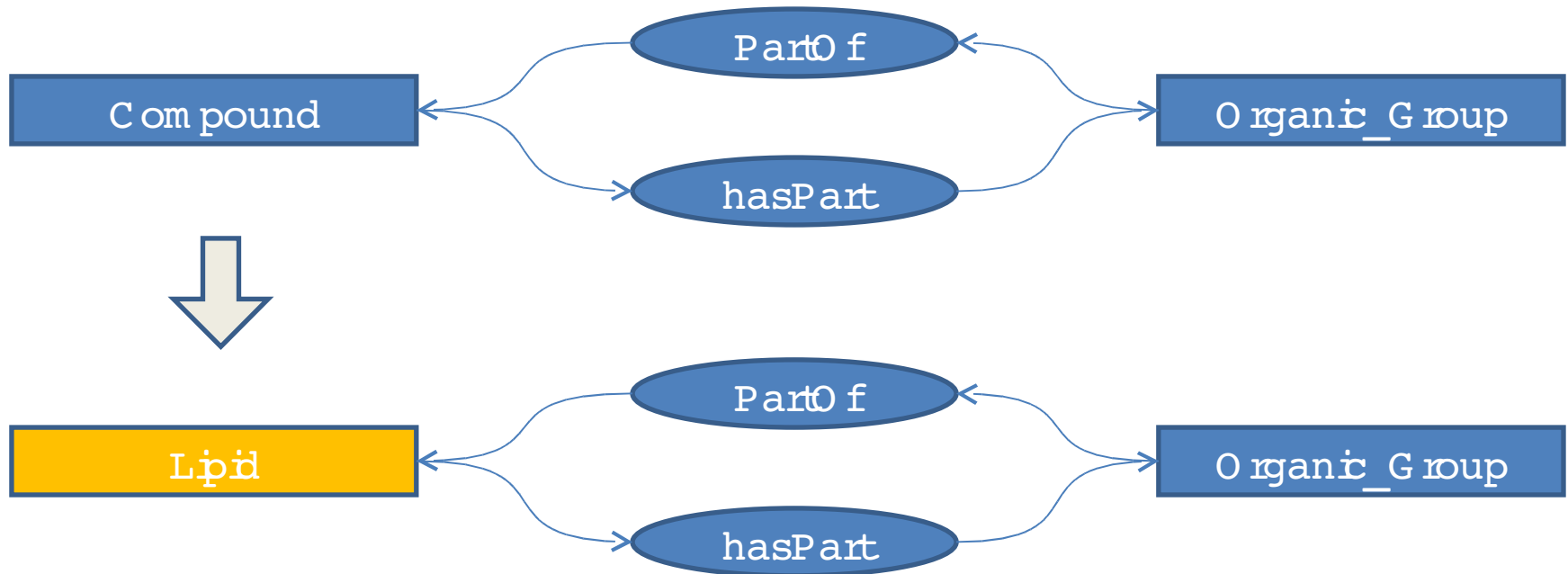
- We re-use primarily the following axiom:

- 32 functional group primitive classes found suitable to be re-used in LipidOntology

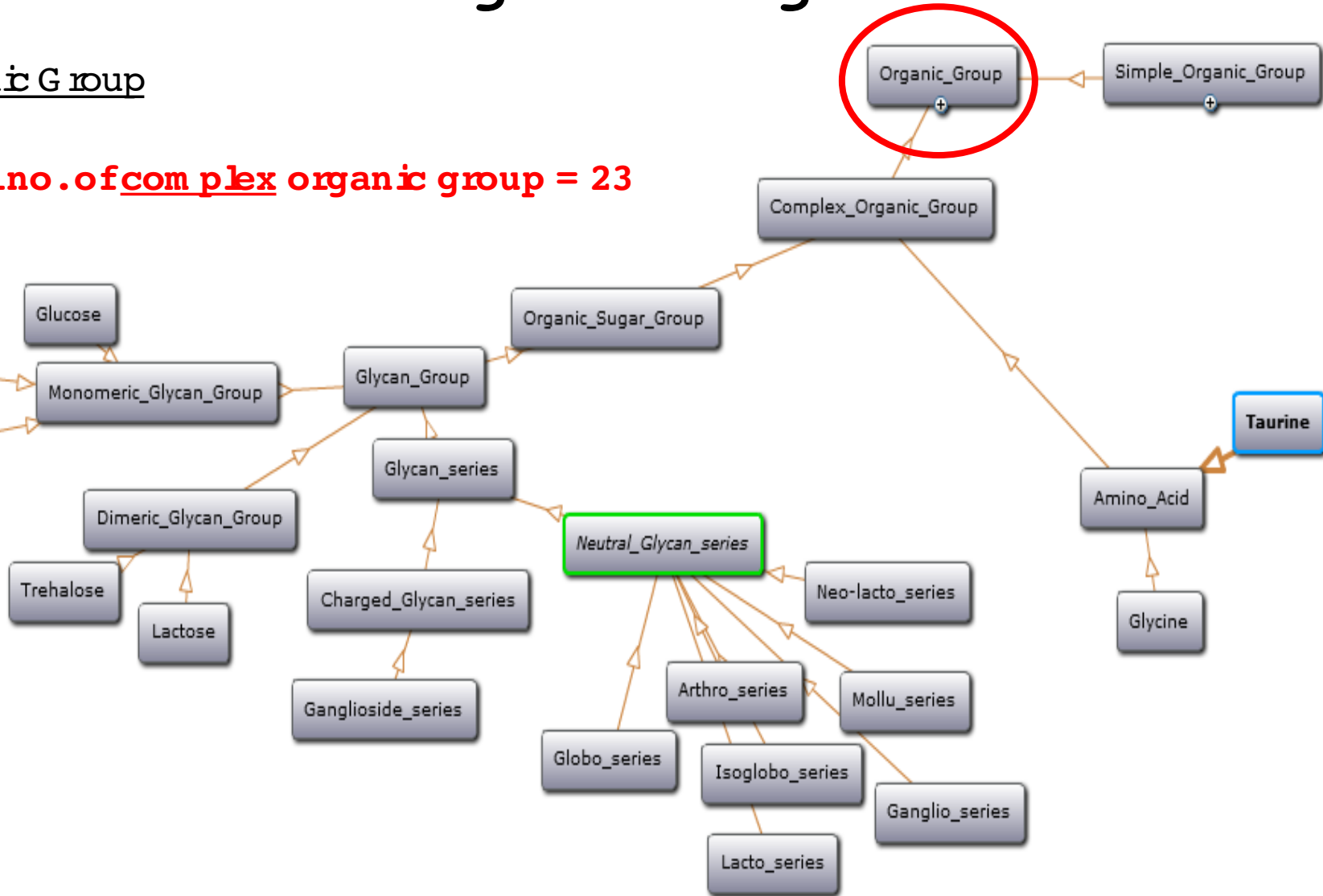# Analysis, reuse and extension of existing ontologies.

- Limit Compound to Lipid

Analysis, reuse and extension of existing ontologies.

Organic Group
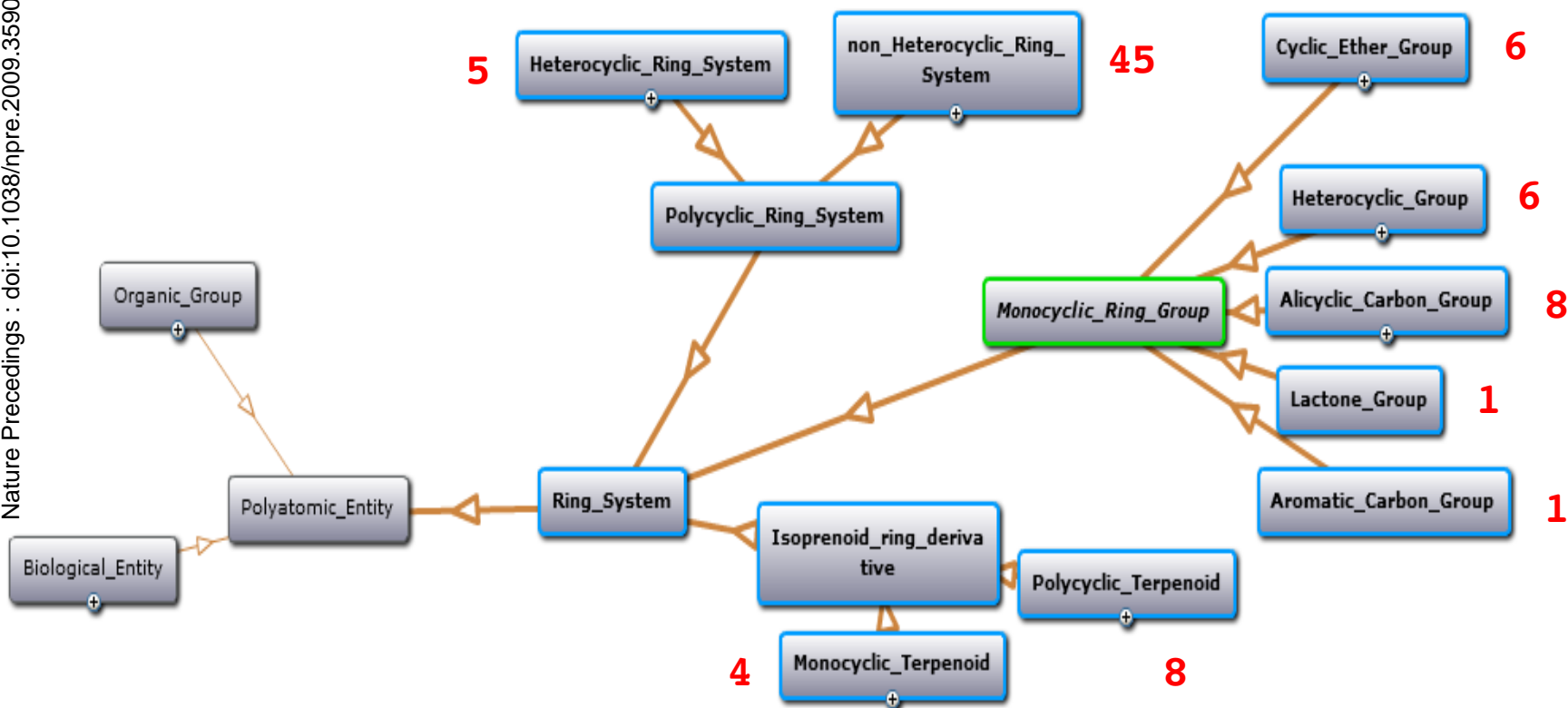
Total no. of simple organic group = 95

63

32

Extensions required to support lipid characterization

Organic group From Chemical Ontology

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 LIPID MAPS hierarchy into the upper ontology

Total No. of Classes	715
No. of Lipid Classes	**492 UPDATED!**
Primitive Lipid Classes	**107 UPDATED!**
Defined Lipid Classes	**268 UPDATED!**
Total No. of Restrictions	901
Total No. of Properties	41
DL Expressivity	ALCHIQ(D)
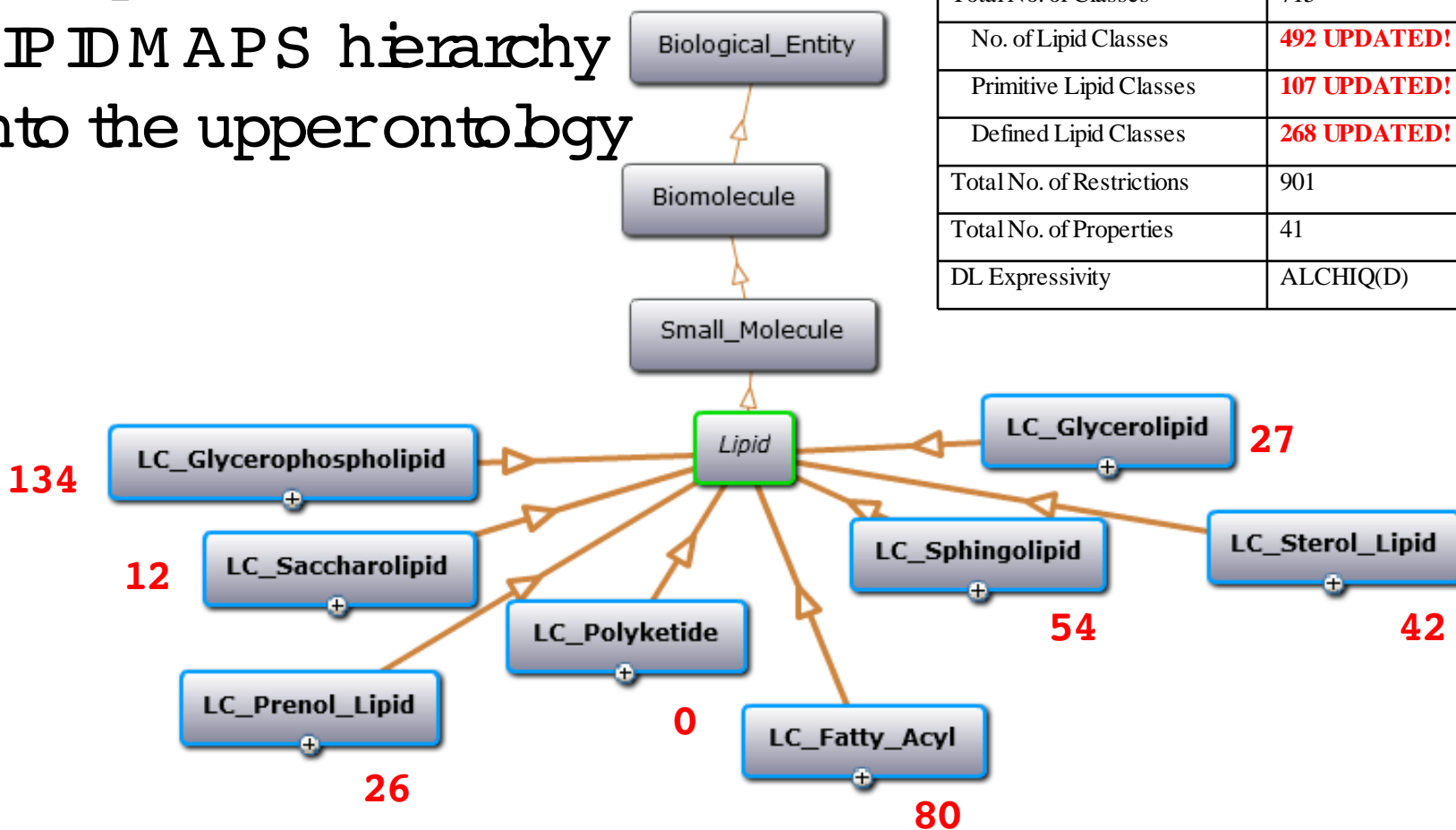
Definition of lipid classes: DL Axioms
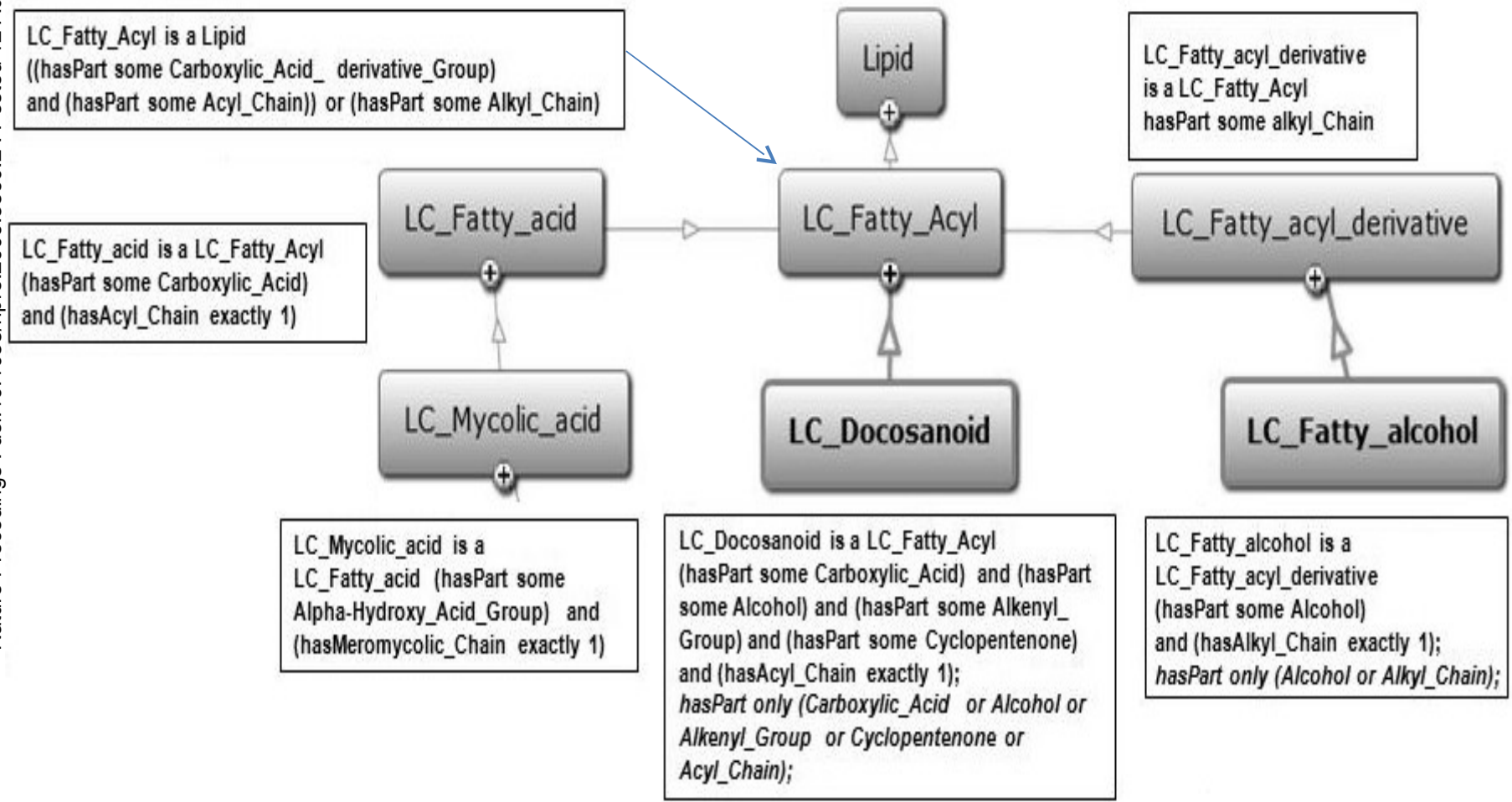
Definition of lipid classes: DL Axioms

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

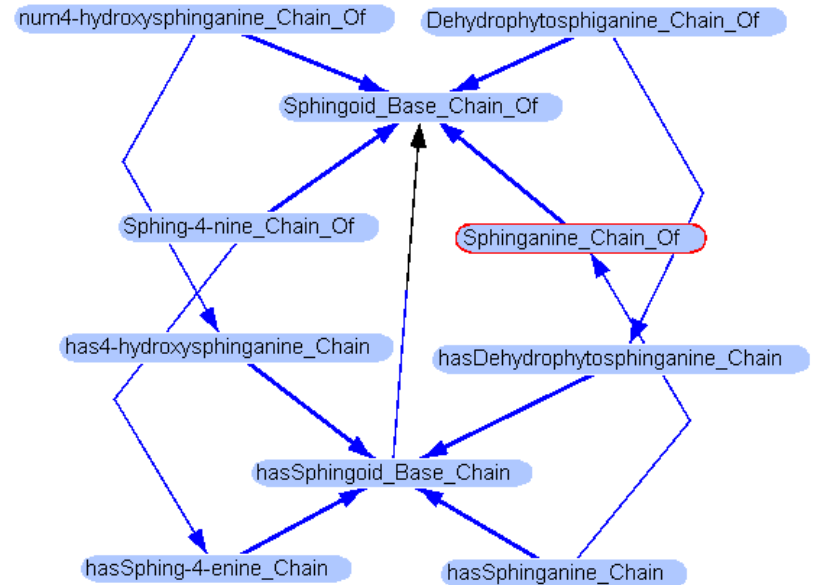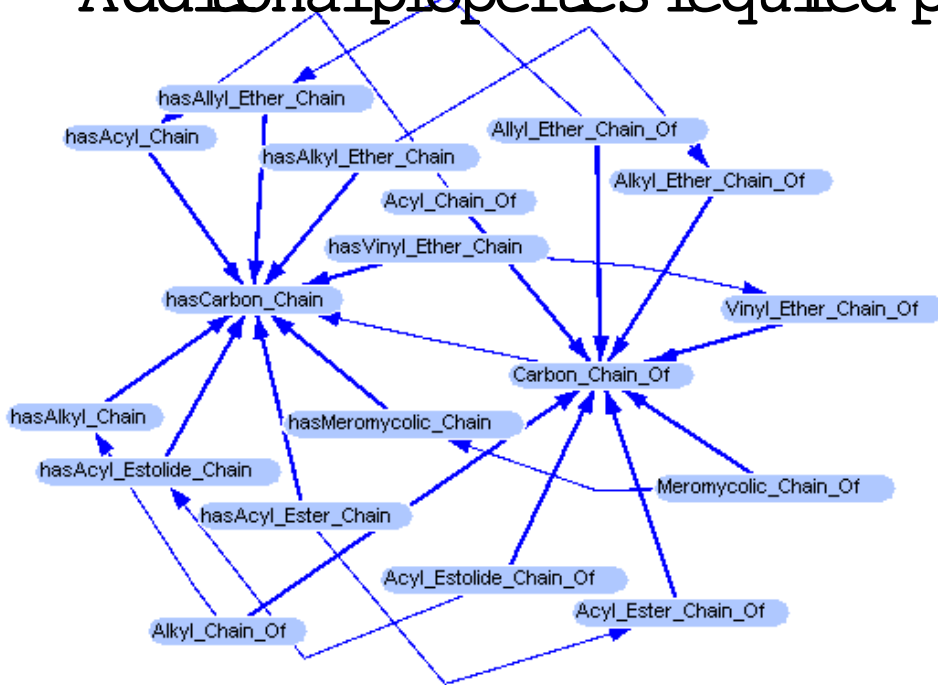- hasPart some Carboxylic_Acid_derivative_Group
- hasPart some Aldehyde
- hasAldehyde exactly 3
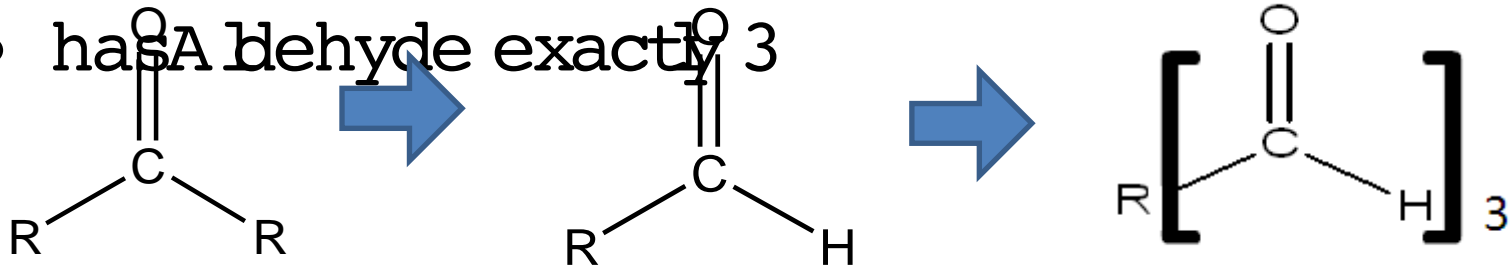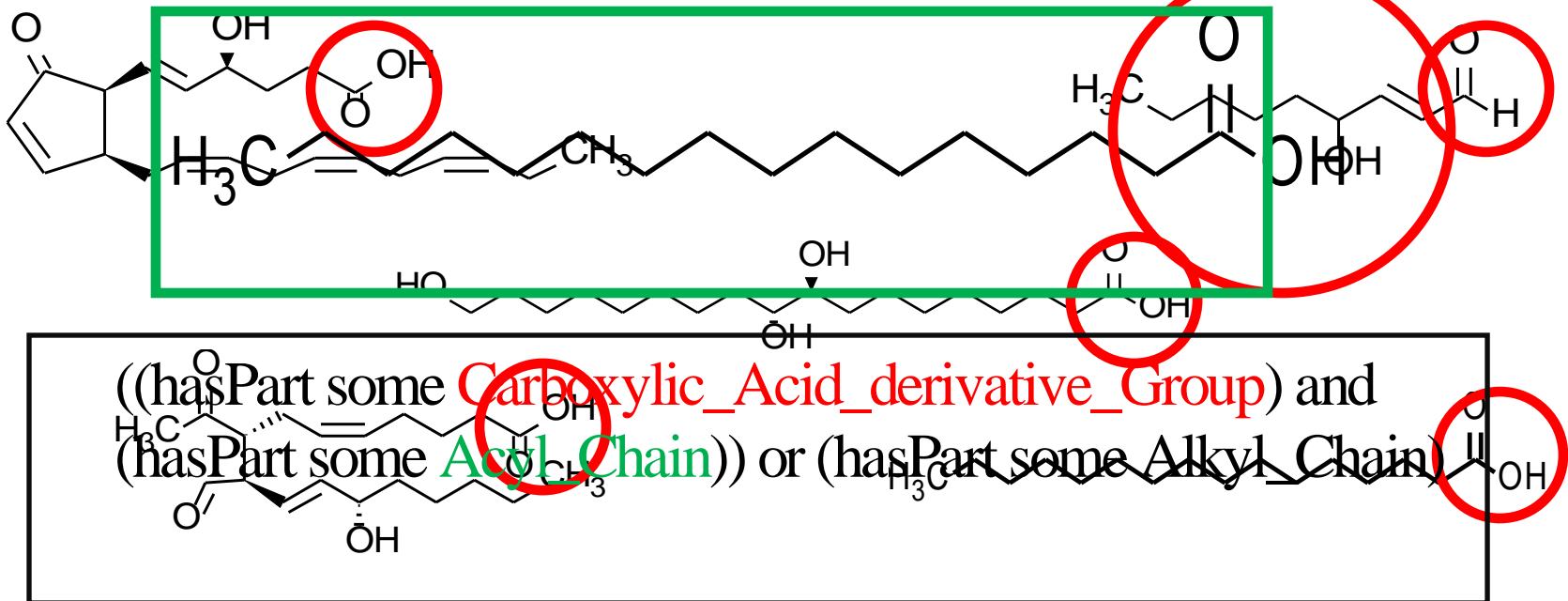
Definition of lipid classes: DL Axioms

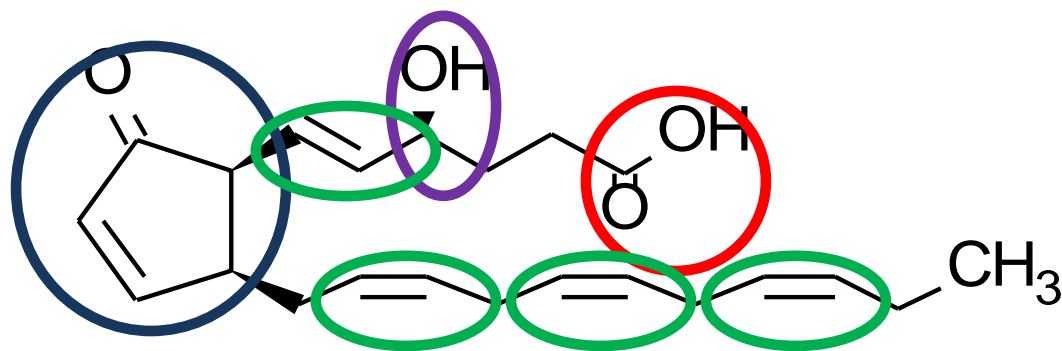## Additional properties required prior to cardinality specification
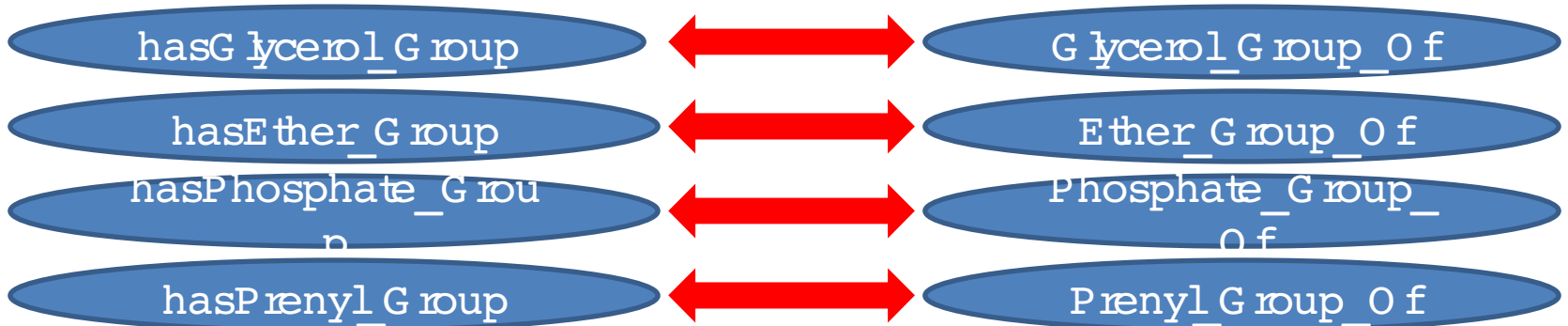
Fatty Acyl

- 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)

Docosanoid

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
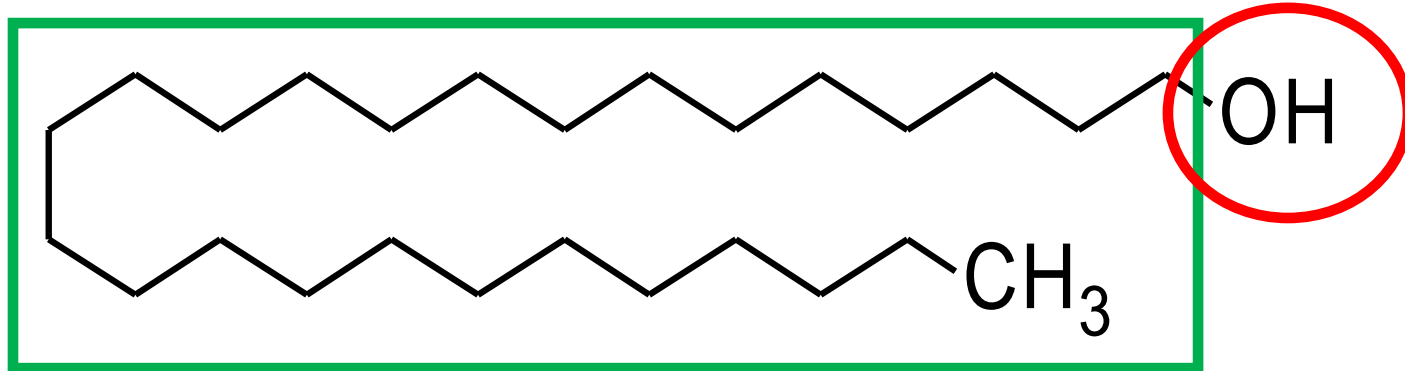
LC_Fatty_acyl_derivative

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

hasPart only (Alcohol or Alkyl_Chain)

Mycolic acid
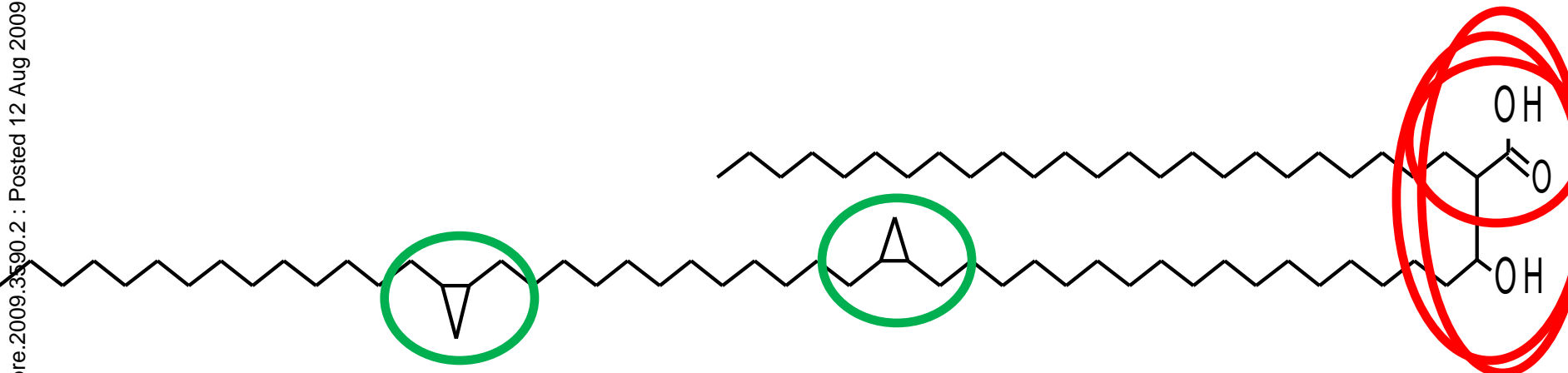
hasPart some (Cyclopropane or Alkenyl_Group)

hasPart some (Carboxylic_Acid) Acid(Group) Chain exactly 1)

(hasMeromycolic_Chain exactly 1)

LC_General_alpha_mycolic_acid

hasPart some Cyclopropane
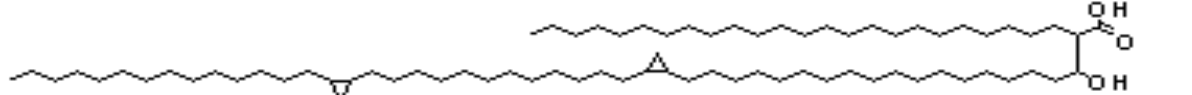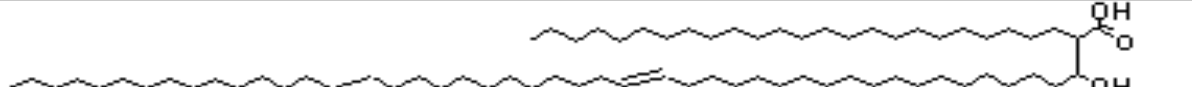
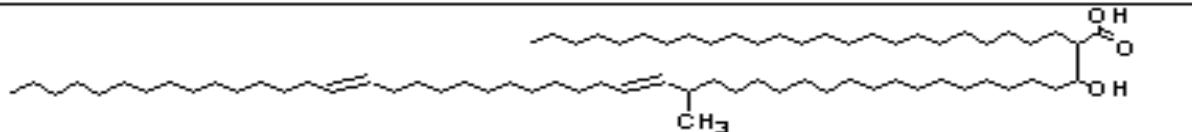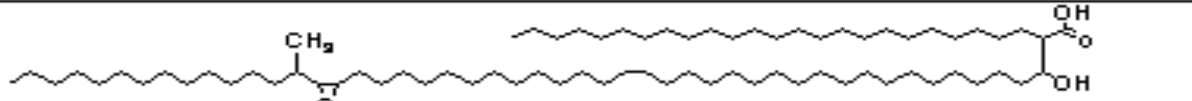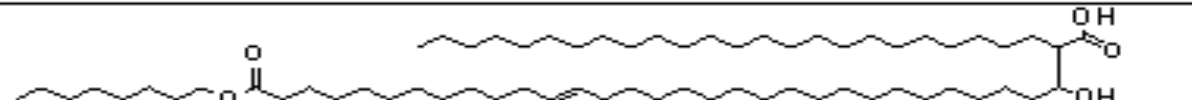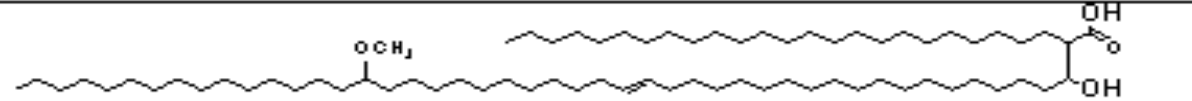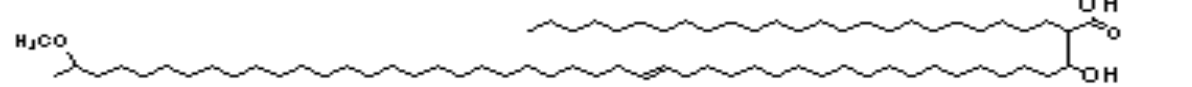hasPart only (Alkenyl_Group or Alpha-hydroxy_Acid_ Group or Cyclopropane or Carboxylic_Acid or Meromycolic_Chain)

New Mycolic Acid Classes
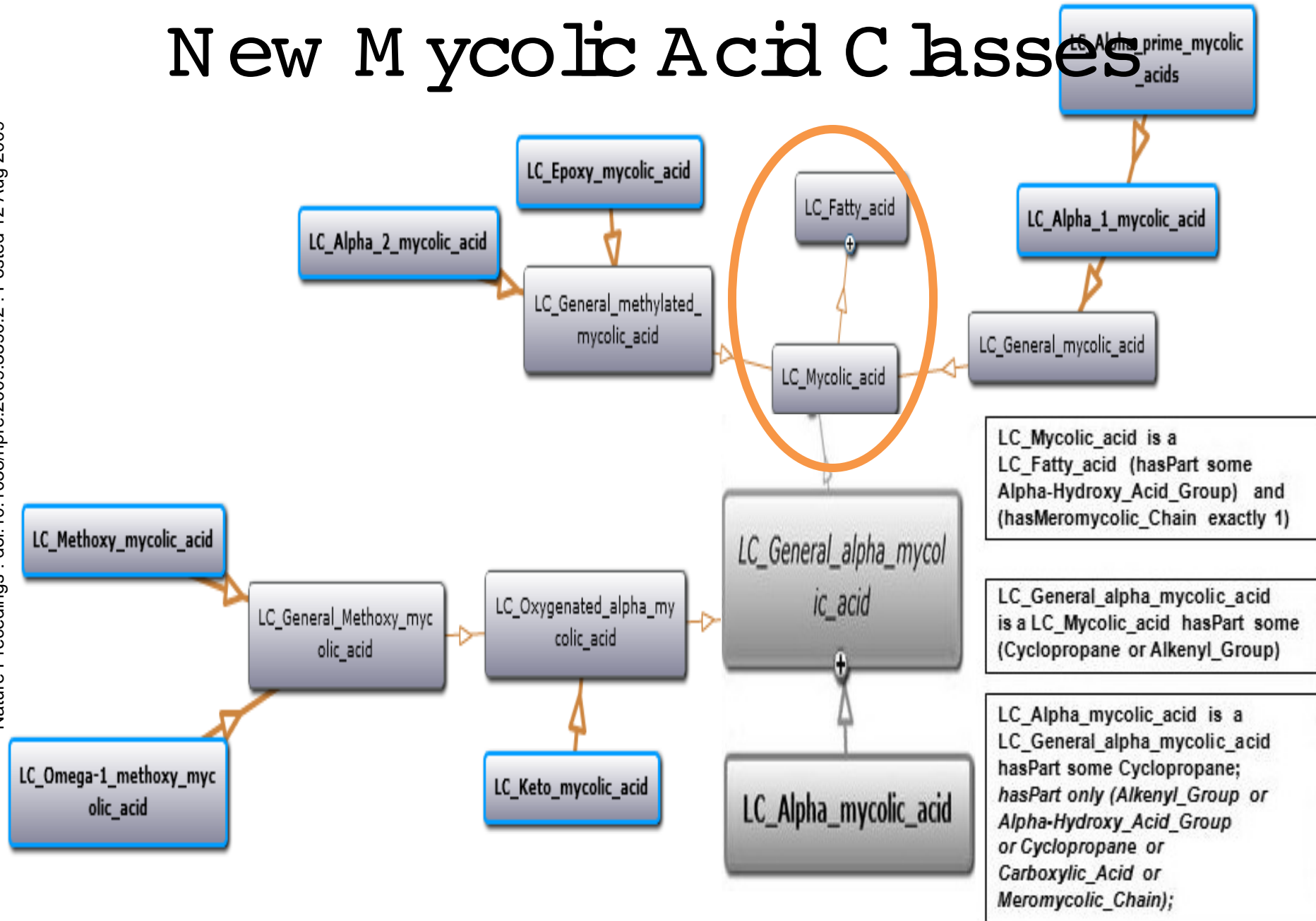
New Classes of Mycolic Acids

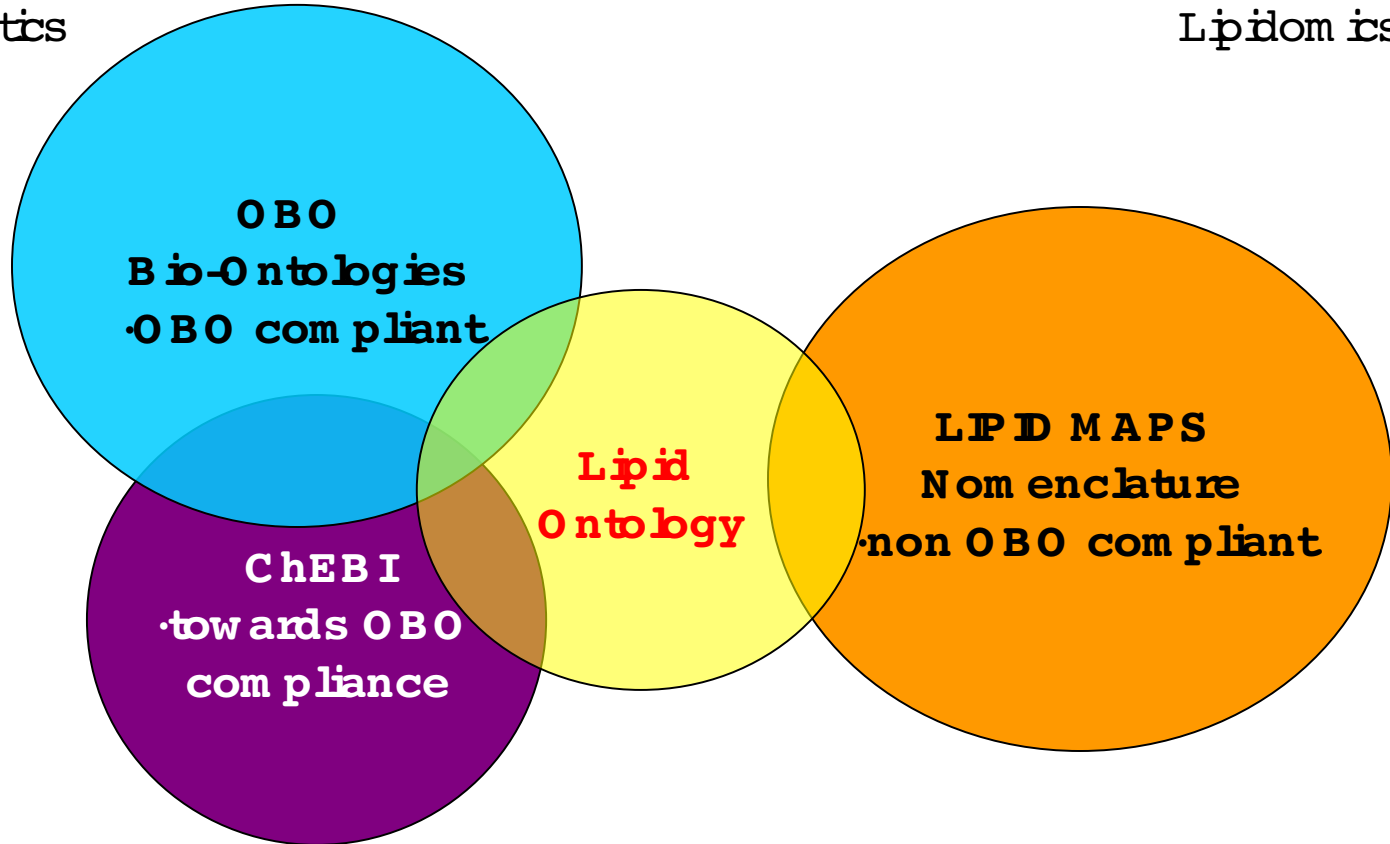| Structure	Class type of Mycolic acid
	Alpha_mycolic_acid
	Alpha_prime_mycolic_acid
	Alpha_1_mycolic_acid
	Alpha_2_mycolic_acid
	Keto_mycolic_acid
	Epoxy_mycolic_acid
	Wax_ester_mycolic_acid
	Methoxy_mycolic_acid
	Omega-1_methoxy_mycolic_acid

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
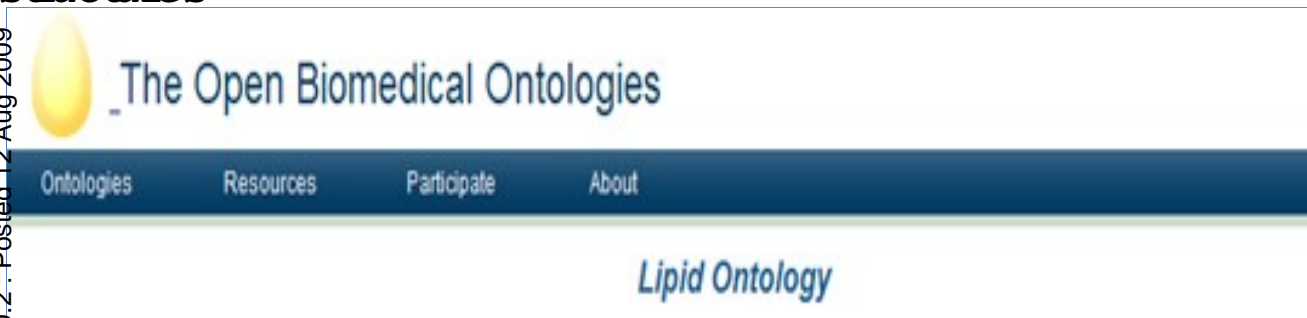
The role of Lipid Ontology

Bioinformatics

Lipidomics

OBO
Bio-Ontologies
·OBO compliant

Lipid
Ontology

LIPID MAPS
Nomenclature
·non OBO compliant

ChEBI
·towards OBO
compliance

Downbad

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

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

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.

Acknowledgments