Searching the LIPID MAPS databases

Dawn Cotter[a], Eoin Fahy[a], Manish Sud[a], Shankar Subramaniam[a,b]

www.lipidmaps.org

National Institute of General Medical Sciences
LIPID MAPS funded by Glue Grant from :www.nigms.nih.gov

Introduction

Given the importance of these molecules in cellular function and pathology, it is essential to have a well-organized database of lipids and lipid-related genes and proteins that is extensible, flexible and scalable. An object-relational database of lipids, the LIPID MAPS Structure Database (LMSD) is based on an internationally accepted classification system and is the largest public lipid-specific repository in the world with over 22,000 individual structures. The LIPID MAPS Proteome Database (LMPD) was developed to provide a catalog of genes and proteins involved in lipid metabolism and signaling. We present extensive browsing and searching tools available on the LIPID MAPS website for these and other lipid-specific resources.

LIPID MAPS Lipid classification system[1,2]

Category[2,3]	Abbrev	Example
Fatty Acyls	FA	Arachidonic acid
Glycerolipids	GL	1-hexadecanoyl-sn-glycerol
Glycerophospholipids	GP	1-hexadecanoyl-2-(9Z-octadecenoyl)-sn-glycero-3-phosphocholine
Sphingolipids	SP	Sphingosine
Sterol Lipids	ST	Cholesterol
Prenol Lipids	PR	Retinol
Saccharolipids	SL	Kdo_2-lipid A
Polyketides	PK	epothilone D

[1] A comprehensive classification system for lipids. Journal of Lipid Research, Vol. 46, 839-862, 2005

[2] Update of the LIPID MAPS comprehensive classification system for lipids. Journal of Lipid Research, Vo l50: S9-S14, 2009

Lipid categories in LIPID MAPS classification system

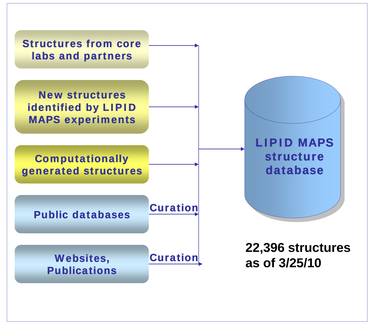
Name: PGE_2
LM_ID: LMFA03010003

LM_ID description:

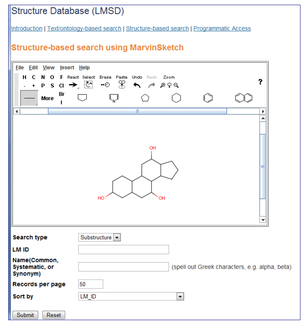
Database: LM (LIPID MAPS)
Category: FA (Fatty Acyls)
Main Class: 03 (Eicosanoids)
Sub Class: 01 (Prostaglandins)
Unique identifier within a sub class: 0003

LIPID MAPS LM_ID description

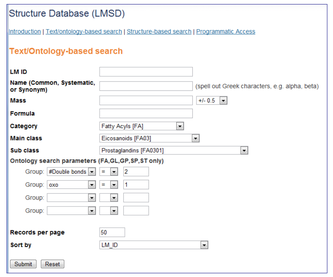
Lipid structure database (LMSD)[3]



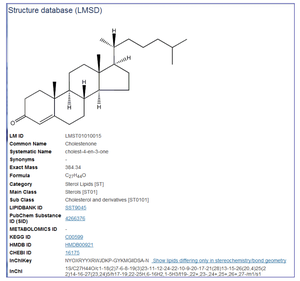
Populating LIPID MAPS structure database



Substructure search interface



Text and ontology-based search interface

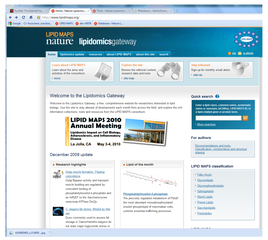


Example of a detail-view page for a lipid structure

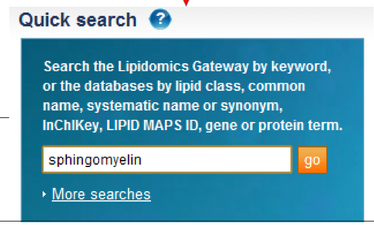
[3] LMSD: LIPID MAPS structure database. Nucleic Acid Research, 2007, Vol. 35, Database issue D527-D32

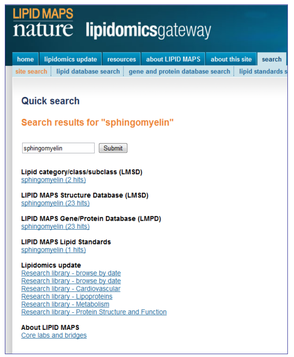
Querying Lipidomics Gateway website and LIPID MAPS databases via "Quick search"

Lipidomics Gateway website: http://www.lipidmaps.org



QUICK SEARCH
- ❖*Small "footprint"*
- ❖*High visibility (on home page)*
- ❖*Multi-purpose*



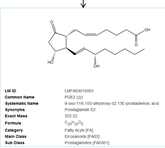


Database hits
(lipid,gene,standards db's Classification system)

Web page hits
(LM text information, updates,references,etc.)

LIPID MAPS LM_ID
LMFA03010003

InChI Key
XEYBRNLFEZDVAW-ARSRFYASSA-N



Lipid LM_ID record

Lipid classification term: choline

Lipid Classification Search

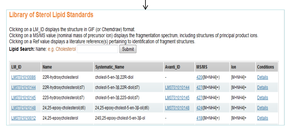
Lipid class/subclass

Lipid molecular formula: $C_{12}H_{24}O_2$

Lipid common/systematic name or synonym: "Linoleic", "HETE", "5Z,8Z,14Z-eicosatrienoic", "PC(16:0/18:1(9Z))" "MGDG", "PAF", "PGE", "docosa", "phytosphingosine"

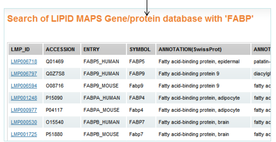
Lipid Structures

Lipid standard (name or LMID): sterol



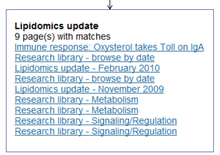
Lipid Standards

Gene / protein name/synonym: FABP



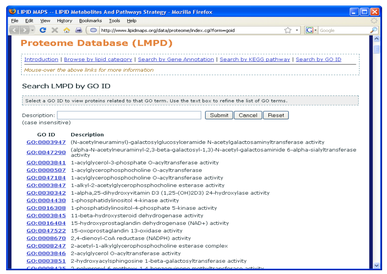
Lipid-related gene / proteins

Keywords on Lipidomics Gateway website pages (personnel, publications, news, updates, etc.): "Atherosclerosis", "Dennis", "homeostasis"

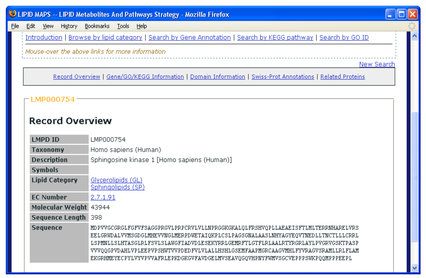


Lipidomics Gateway html pages

Lipid proteome database (LMPD)[4]

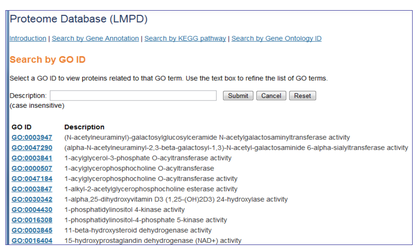


LMPD database search



LMPD search results: Overview

LMPD search by KEGG pathway

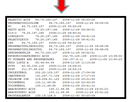


LMPD search by Gene ontology (GO) annotation

[4] LMPD: LIPID MAPS proteome database. Nucleic Acid Research, 2006, Vol. 34, Database issue D507-D510

★ *Obtaining end-user feedback to improve and update the databases*

Databases



New entries (lipid structures, lipid / gene names, synonyms,etc)

Log file of quick search queries

★ *Improving the effectiveness of text-based searching via "Quick Search"*

Problem:
No matches due to issues of whitespace, parentheses, hypens, etc. in user input which are different from those in database fields.

Solution:
Use of concatenated versions of lipid / gene / classification common names, systematic names and synonyms (stored in database):

Removal of whitespace characters and ,-_[]{}()' from concatenated entries

These characters are also removed from user's query prior to searching

LIPID MAPS Bioinformatics Core[a], SDSC, UCSD, 9500 Gilman Dr, La Jolla, CA 92093; Department of Bioengineering[b], UCSD, 9500, Gilman Dr, La Jolla, CA, 92093;