

# GenMAPP Gene Database for *Sinorhizobium meliloti* str. 1021

[Sinorhizobium\\_meliloti\\_1021\\_mpetredi\\_2013123-2.gdb](#)

## ReadMe

Last revised: 12/06/2013

This document contains the following:

1. Overview of GenMAPP application and accessory programs
2. System Requirements and Compatibility
3. Installation Instructions
4. Gene Database Specifications
  1. Gene ID Systems
  2. Species
  3. Data Sources and Versions
  4. Database Report
5. Contact Information for support, bug reports, feature requests
6. Release notes
  1. Current version:[Sinorhizobium\\_meliloti\\_1021\\_mpetredi\\_2013123-2.gdb](#)
7. Database Schema Diagram

### 1. Overview of the GenMAPP application and accessory programs

GenMAPP (Gene Map Annotator and Pathway Profiler) is a free computer application for viewing and analyzing DNA microarray and other genomic and proteomic data on biological pathways. MAPPFinder is an accessory program that works with GenMAPP and Gene Ontology to identify global biological trends in gene expression data. The GenMAPP Gene Database (file with the extension *.gdb*) is used to relate gene IDs on MAPPs (*.mapp*, representations of pathways and other functional groupings of genes) to data in Expression Datasets (*.gex*, DNA microarray or other high-throughput data). GenMAPP is a stand-alone application that requires the Gene Database, MAPPs, and Expression Dataset files to be stored on the user's computer. GenMAPP and its accessory programs and files may be downloaded from <http://www.GenMAPP.org>. GenMAPP requires a separate Gene Database for each species. This ReadMe describes a Gene Database for *Sinorhizobium Meliloti 1021* that was built by the Loyola Marymount University (LMU) Bioinformatics Group using the program GenMAPP Builder 2.0, part of the open source XMLPipeDB project <http://xmlpipedb.cs.lmu.edu/>.

### 2. System Requirements and Compatibility:

- This Gene Database is compatible with GenMAPP 2.0 and 2.1 and MAPPFinder 2.0. These programs can be downloaded from <http://www.genmapp.org>.
- System Requirements for GenMAPP 2.0/2.1 and MAPPFinder 2.0:

Operating System: Windows 98 or higher, Windows NT 4.0 or higher (2000, XP, etc)

Monitor Resolution: 800 X 600 screen or greater (SVGA)

Internet Browser: Microsoft Internet Explorer 5.0 or later

Minimum hardware configuration:

Memory: 128 MB (512 MB or more recommended)

Processor: Pentium III

Disk Space: 300 MB disk (more recommended if multiple databases will be used)

### 3. Installation Instructions

- Extract the zipped archive and place the file “[Sinorhizobium\\_meliloti\\_1021\\_mpetredi\\_2013123-2.gdb](#)” in the folder you use to store Gene Databases for GenMAPP. If you accept the default folder during the GenMAPP installation process, this folder will be C:\GenMAPP 2 Data\Gene Databases.
- To use the Gene Database, launch GenMAPP and go to the menu item *Data > Choose Gene Database*. Alternatively, you can launch MAPPFinder and go to the menu item *File > Choose Gene Database*.

### 4. Gene Database Specifications

#### 1. Gene ID Systems

This *Sinorhizobium meliloti* Gene Database is UniProt-centric in that the main data source (primary ID System) for gene IDs and annotation is the UniProt complete proteome set for *Sinorhizobium meliloti*, made available as an XML download. In addition to UniProt IDs, this database provides the following proper gene ID systems that were cross-referenced by the UniProt data: OrderedLocusNames, GeneID (NCBI), and RefSeq (protein IDs of the form NP\_#####). It also supplies UniProt-derived annotation links from the following systems: EMBL, InterPro, PDB, and Pfam. The Gene Ontology data has been acquired directly from the Gene Ontology Project. The GOA project was used to link Gene Ontology terms to UniProt IDs. Links to data sources are listed in the section below.

Proper ID System	SystemCode
UniProt	S
OrderedLocusNames	N
GeneID (NCBI)	L
RefSeq	Q

#### 1. Species

This Gene Database is based on the UniProt proteome set for *Sinorhizobium meliloti* str. 1021

#### 1. Data Sources and Versions

- This *Sinorhizobium meliloti* Gene Database was built on December 3, 2013; this build date is reflected in the filename [Sinorhizobium\\_meliloti\\_1021\\_mpetredi\\_2013123-2.gdb](#). All date fields internal to the Gene Database (and not usually seen by regular GenMAPP users) have been filled with this build date.
- UniProt complete proteome set for *Sinorhizobium meliloti* str. 1021, downloaded from this page:

<[http://www.uniprot.org/uniprot/?query=organism%3a266834+keyword%3a181&format=\\*](http://www.uniprot.org/uniprot/?query=organism%3a266834+keyword%3a181&format=*)>

Filename:

“<http://www.uniprot.org/uniprot/?query=organism%3a266834+keyword%3a181&format=xml>”

(downloaded as a compressed .gz file and extracted)

Version information for the proteome sets can be found at <<http://www.uniprot.org/news/>>

The proteome set used for this version of the *Sinorhizobium meliloti* Gene Database was based on UniProt release 2013\_10 released on October 16, 2013.

· Gene Ontology gene associations are provided by the GOA project:

<<http://www.ebi.ac.uk/GOA/>> as a tab-delimited text file. The *Sinorhizobium meliloti* GOA file was accessed from the GOA proteomes FTP site: <

<http://ftp.ebi.ac.uk/pub/databases/GO/goa/proteomes/>>

Filename: “58.R\_meliloti.goa”. Version 12/10/2013 7:19:00 PM.

· Gene Ontology data is downloaded from

<<http://beta.geneontology.org/page/download-ontology>>

Data is released daily. For this version of the *Sinorhizobium meliloti* Gene Database we used the ontology version 2013-10-21 12:54:57 PM.

Filename: “go\_daily-termdb.obo-xml.gz”.

### 1. Database Report

· UniProt is the primary ID system for the *Sinorhizobium meliloti* Gene Database. The UniProt table contains all 6169 UniProt IDs contained in the UniProt proteome set for this species.

· The OrderedLocusNames ID system was derived from the cross-references in the UniProt proteome set. Because *Sinorhizobium meliloti* has the alternate name *Rhizobium meliloti* there were two types of IDs. For *Rhizobium meliloti* the IDs used were R#####, RA#####, and RB#####. For *Sinorhizobium meliloti* the IDs used were SM\_b#####, SMa#####, SMb#####, and SMc#####. Both of these IDs were used in the literature. We compared this table with the list of gene IDs in the JCVI Comprehensive Microbial Resource (CMR) at <<http://cmr.jcvi.org/tigr-scripts/CMR/GenomePage.cgi?org=ntsm01>>. There are 6205 protein coding genes listed there. Of the protein coding genes, 36 gene IDs do not appear in our Gene Database because they are not cross-listed in the UniProt XML file.

· The following table lists the numbers of gene IDs found in each gene ID system:

ID System	ID Count Current version
EMBL	11908
GeneID (NCBI)	6210
GeneOntology	5218
InterPro	4393

OrderedLocusNames	10939
PDB	108
Pfam	2188
RefSeq	6210
UniProt	6169

## 5. Contact Information for support, bug reports, feature requests

- The Gene Database for *Sinorhizobium meliloti* was built by the Loyola Marymount University (LMU) Bioinformatics Group using the program GenMAPP Builder, part of the open source XMLPipeDB project <<http://xmlpipedb.cs.lmu.edu/>>.
- For support, bug reports, or feature requests relating to XMLPipeDB or GenMAPP Builder, please consult the XMLPipeDB Manual found at <<http://xmlpipedb.cs.lmu.edu/documentation.shtml>> or go to our SourceForge site <<http://sourceforge.net/projects/xmlpipedb/>>.
- For issues related to the *Sinorhizobium meliloti* Gene Database, please contact:

Kam D. Dahlquist, PhD.

Department of Biology

Loyola Marymount University

1 LMU Drive, MS 8220

Los Angeles, CA 90045-2659

[kdahlquist@lmu.edu](mailto:kdahlquist@lmu.edu)

- For issues related to GenMAPP 2.0/2.1 or MAPPFinder 2.0 please contact GenMAPP support directly by e-mailing [genmapp@gladstone.ucsf.edu](mailto:genmapp@gladstone.ucsf.edu) or [GenMAPP@googlegroups.com](mailto:GenMAPP@googlegroups.com).

## 6. Release Notes

### a. Current version: [Sinorhizobium\\_meliloti\\_1021\\_mpetredi\\_2013123-2.gdb](#)

- Miles Malefyt, Lauren Magee, Mitchel Petredis, and Stephen Louie contributed to this first release.