

Sinorhizobium meliloti str. 1021
Gene Database

Team Name

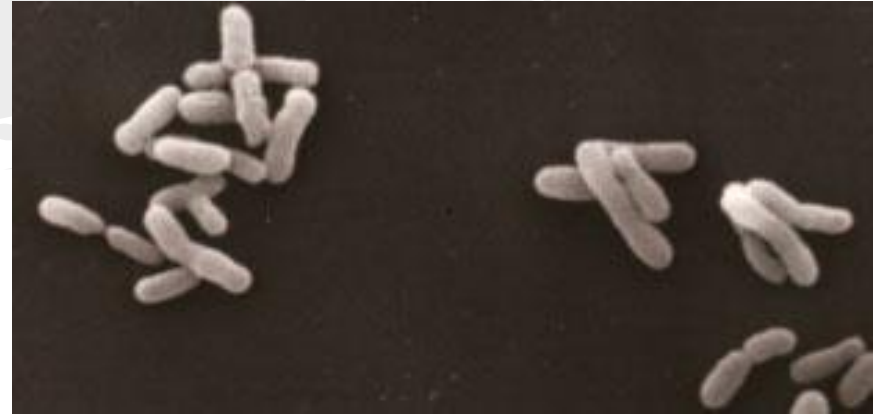
Miles Malefyt, Lauren Magee, Mitchell Petredis,
Stephen Louie

Outline

- *Sinorhizobium meliloti* is an important Plant Growth Promoting Rhizobacteria (PGPR)
- A topic of interest is the transcriptional responses of *S. meliloti*
- Raw Microarray Data was downloaded, processed, and tested in GenMAPP
- A Gene Database was created that compiled the gene expression changes of *S. meliloti*

Introduction to *Sinorhizobium meliloti*

- Rhizobial symbiont to legume alfalfa
- Creates nodules in the roots the convert Dinitrogen to ammonia
- Aerobic but lives in low Oxygen
- Affects stable agriculture and ecosystem function → nitrogen fixation



<http://cmgm.stanford.edu/~mbarnett/1021.jpg>

GenMAPP User Objectives

- Identify an article documenting the microarray data for *Sinorhizobium meliloti*
- Download this data in its rawest form
- Compare this download with the samples in the experiment documented in your article.
- Perform appropriate statistical analysis on the data
- Format file to be imported into GenMAPP.
- Run GenMAPP and create Color Sets
- Use MAPPFinder to determine relevant biological pathways for *Sinorhizobium meliloti*

Article Identification and Raw Data Downloads

- Experiment used 4 time points with 3 replicates each
- Data labelled Sm6kOligo#####_300N1-t15
 - includes concentration, time point, and replicate in its name
 - 300-mm concentration
 - N-NaCl or the salt version of experiment
 - N#-replicate
 - t##- the time point and which each replicate was made
- Data labelled Smc#####
 - 700-mm sucrose concentration
 - S- sucrose version of experiment
 - S#-replicate
 - t###- the time point (15, 30, 60, 240)

Statistical Analysis

NaCl-20160 microarray dots measured

- P value < .05
 - T15-5520
 - T30-7484
 - T60-6711
 - T240-5901

Sucrose- 20160 microarray dots measured

- P value < .05
 - T15- 3613
 - T30- 5225
 - T60- 5207
 - T240- 6790

GenMAPP Errors- 700mm Sucrose

- Gene IDs carrying extraneous information
- GenMAPP stopped responding entirely
 - possibly due to the overwhelming amount of errors produced, due to the IDs failing to match those in the Gene Database.
- Extraneous information was deleted
 - GenMAPP found ~5,000 errors due to Gene IDs not being in the Ordered Gene Locus.
 - Gene Ontology Examination yet to come.

GenMAPP Results for 300mm NaCl

Gene Ontology	Z score
Bacterial type flagellum basal body	2.928
tRNA aminoacylation for protein translation	4.562
ATP synthase activity coupled to transmembrane movement of ions	3.705
Ligase activity forming carbon-nitrogen bonds, rotational mechanism	4.562
structural constituents of Ribosome	8.655
plasma membrane ATP synthesis coupled proton transport	3.78
reductive pentose-phosphate cycle	2.928
lysine biosynthetic process via diaminopimelate	3.705
hydrogen transport	4.562
amino-acid activation	4.562

Comparison Between Article and Downloaded Data

- Data from microarray article

VOL. 188, 2006

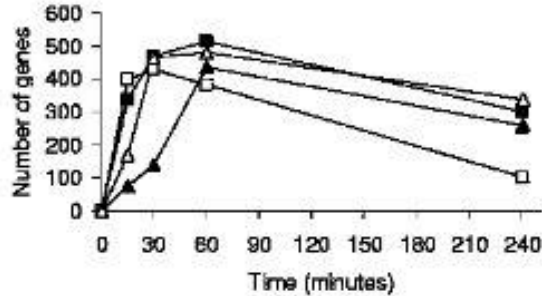
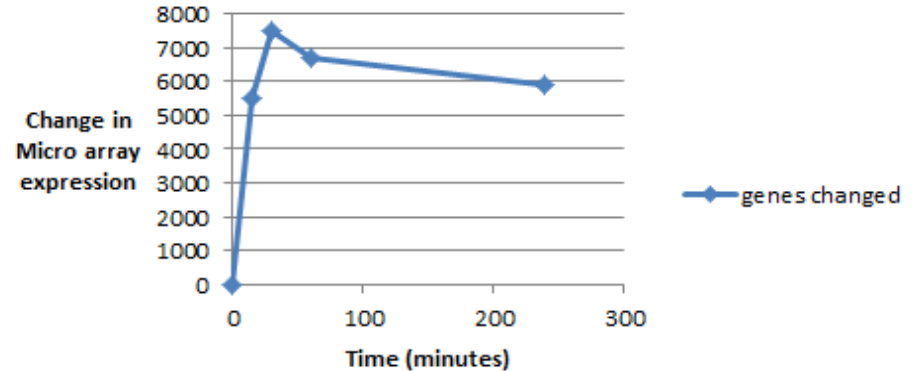


FIG. 1. Numbers of differentially expressed genes at different times after addition of 0.3 M NaCl (○), 0.4 M NaCl (▲), 0.5 M sucrose (□), and 0.7 M sucrose (△).

Microarray expression over time 300mm NaCl solution



Coding Objectives

- Create a custom version of gmbuilder specific to *Sinorhizobium meliloti*.
 - Tweak the code of gmbuilder to reduce errors.
- Create a new database in PostgreSQL
 - Import all important data to the database via gmbuilder.
- Export the data as a .gdb file
 - Done for every new version of gmbuilder

Coding: 1st Export

- gmbuilder-2.0b71
 - unedited program
- Initial testing to determine how many changes were needed.
- Result: > 20,000 errors!

Code Tweak #1

Summary of Milestone 3 in “Coder” wiki page.

1. Add a "species profile" to the GenMAPP Builder code base.
2. Customize the species profile: *Sinorhizobium meliloti*
 - a. Put in OrderedLocusNames record of the Systems Table
3. Customize the Link field in the OrderedLocusNames record of the Systems Table
 - a. <http://cmr.jcvi.org/tigr-scripts/CMR/shared/GenePage.cgi?locus=~>
4. Record any feedback from QA and GenMAPP Users about errors and determine the cause
5. Adjust code to correct errors
6. Commit the species profile to SourceForge
7. Release a new version of GenMAPP Builder.
8. Repeat steps 4-7.

Coding: 2nd Export

- GenMAPP_Builder_2.0b72
 - 1st custom version
- Result: ~20,000 errors
 - Progress!
- What went wrong?

The screenshot shows a software dialog box titled "Export Wizard Dialog". The main section is "Export to GenMAPP: Basic Information". It contains several input fields and a notes field:

- Profile:** A dropdown menu set to "org.uniprot.uniprot.Uniprot". To its right is a text block: "This profile defines the requirements for any UniProt-centric Gene Database."
- Owner:** A text box containing "S. meliloti".
- Version (MM/dd/yyyy):** A text box containing "11/21/2013".
- MODSystem:** A text box containing "UniProt".
- Species:** A text box containing "Taxon ID 266834, Sinorhizobium meliloti". To its right is a text block: "Info for selected species: Taxon ID 266834, Sinorhizobium meliloti, custom profile."
- Modify (MM/dd/yyyy):** A text box containing "11/21/2013".
- Notes:** A text box containing "Exported by GenMAPP Builder 2.0b72".

At the bottom of the dialog are three buttons: "Back", "Next", and "Cancel".

Quality Assurance Objectives

- Assist Coder with Import/Export Cycle
- Inspect and Identify Gene IDs
- Troubleshoot Discrepancies within Gene IDs

R. meliloti before *S. meliloti*

Originally named as *Rhizobium meliloti*

- R#####
- RA#####
- RB#####

Change to *S. meliloti* created complications

Renaming caused IDs to switch

- SM_b#####
- SMa#####
- SMb#####
- SMc#####

Switch in name made things confusing

Issues with Identification:

- Both types of Gene IDs were present in dataset.
- Some outlier IDs (RB10006.1 and SMc02725.1)

Both Types were accounted for

- Created code that Identified all types of Gene IDs
 - Tags “ordered locus” and “ORF” were added to ID search
- No synonymous IDs were present, so overlap was not an issue (ex. RA1234 and SMa1234)

Code Tweak #2

- New code for exceptions
 - Searches for Gene IDs within ordered locus and ORF

```
@Override
public TableManager getSystemTableManagerCustomizations(TableManager tableManager, TableManager primarySystemTableManager, Date version) throws SQLException,
List<String> comparisonList = new ArrayList<String>();
comparisonList.add("ordered locus");
comparisonList.add("ORF");

return systemTableManagerCustomizationsHelper(tableManager, primarySystemTableManager, version, "OrderedLocusNames", comparisonList);
}
```

```
package edu.lmu.xmlpipedb.gmbuilder.databasetoolkit.profiles;
```

```
import java.sql.SQLException;
```

```
public class SinorhizobiumMelilotiUniProtSpeciesProfile extends UniProtSpeciesProfile {
```

```
    public SinorhizobiumMelilotiUniProtSpeciesProfile() {  
        super("Sinorhizobium meliloti",  
              266834,  
              "This profile customizes the GenMAPP Builder export for " +  
                "Sinorhizobium meliloti" +  
                " data loaded from a UniProt XML file.");  
    }
```

```
    @Override
```

```
    public TableManager getSystemsTableManagerCustomizations(TableManager tableManager, DatabaseProfile dbProfile) {  
        super.getSystemsTableManagerCustomizations(tableManager, dbProfile);  
        tableManager.submit("Systems", QueryType.update, new String[][] {  
            { "SystemCode", "N" },  
            { "Species", "|" + getSpeciesName() + "|" }  
        });  
  
        tableManager.submit("Systems", QueryType.update, new String[][] {  
            { "SystemCode", "N" },  
            { "Link", "http://cmr.jcvi.org/tigr-scripts/CMR/shared/GenePage.cgi?locus=~" }  
        });  
  
        return tableManager;  
    }
```

```
    @Override
```

```
    public TableManager getSystemTableManagerCustomizations(TableManager tableManager, TableManager primarySystemTableManager, Date version) throws SQLException,  
        List<String> comparisonList = new ArrayList<String>();  
        comparisonList.add("ordered locus");  
        comparisonList.add("ORF");  
  
        return systemTableManagerCustomizationsHelper(tableManager, primarySystemTableManager, version, "OrderedLocusNames", comparisonList);  
    }
```

```
}
```

Coding: 3rd Export

- SmelilotiGenMAPP_Builder_2.0b73
- Results: ~1,500 errors
 - Significant Progress!
- Why are there still errors?

Match test was conducted to ensure no errors were made

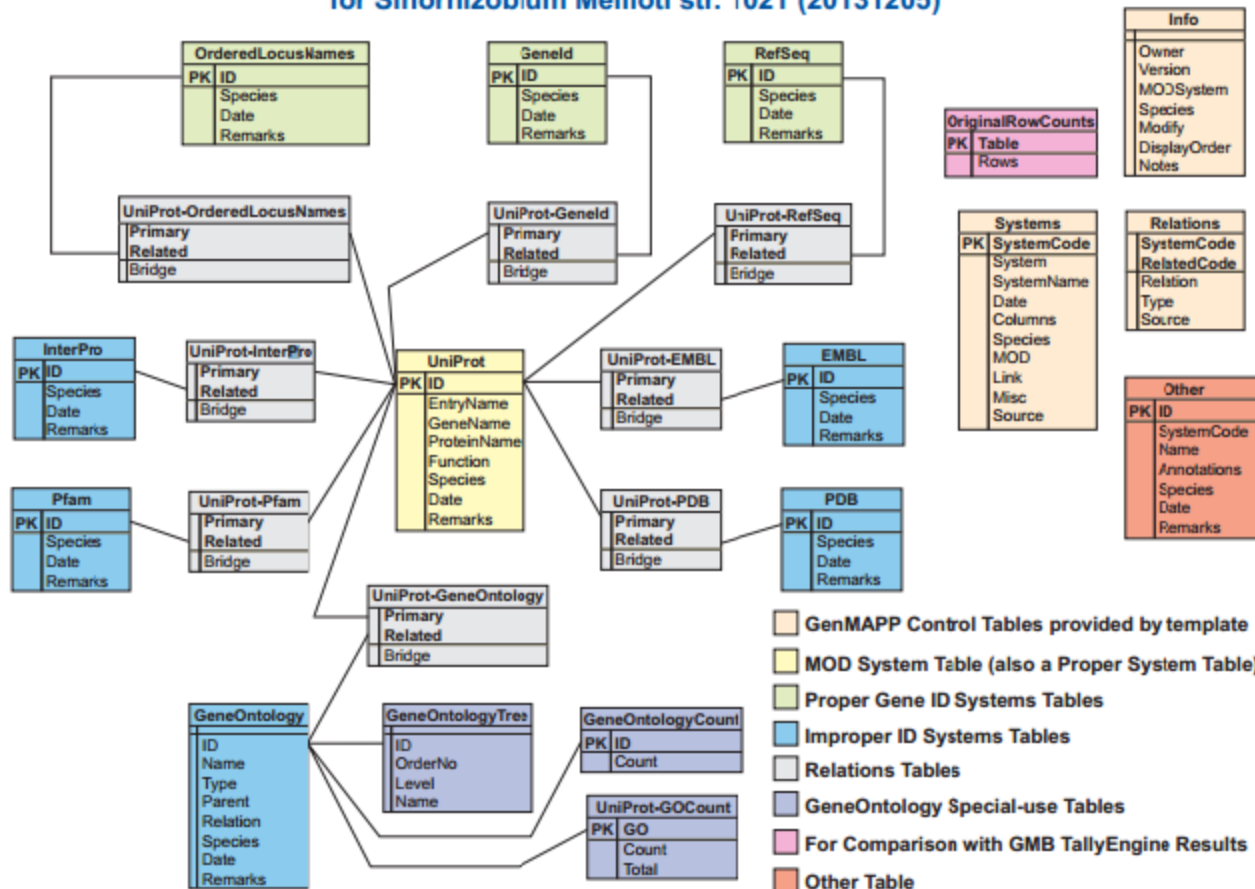
- The 1527 exceptions were compared with the original Gene ID list from the xml file.
- All comparisons showed up as #N/A which meant that the IDs were never present to begin with.

Final Export was more successful

- 13% error
- IDs found in microarray dataset were not present in xml file.
 - Ex. SMc02725 stringency control

Gene ID 1 (errors)	Gene ID 2 (xml)	like 1 to 2	like 2 to 1
SMa0046	R00001	#N/A	#N/A
SMa0046	R00001	#N/A	#N/A
SMa0046	R00002	#N/A	#N/A
SMa0278	R00002	#N/A	#N/A
SMa0278	R00003	#N/A	#N/A
SMa0278	R00003	#N/A	#N/A
SMa0292	R00004	#N/A	#N/A
SMa0292	R00004	#N/A	#N/A
SMa0292	R00005	#N/A	#N/A
SMa0356	R00005	#N/A	#N/A
SMa0356	R00006	#N/A	#N/A
SMa0356	R00006	#N/A	#N/A
SMa0370	R00007	#N/A	#N/A
SMa0370	R00007	#N/A	#N/A
SMa0370	R00008	#N/A	#N/A
SMa0409	R00008	#N/A	#N/A
SMa0409	R00009	#N/A	#N/A
SMa0409	R00009	#N/A	#N/A

GenMAPP Gene Database Schema for *Sinorhizobium Meliloti* str. 1021 (20131205)



Conclusions

- Gene Database was created for the *S. meliloti*
- GenMAPP analysis showed that gene expression increases under salt and sucrose stress.

Acknowledgements

Thanks to Dr. Lum for her expertise in
Sinorhizobium meliloti

Thanks to Dr. Dionisio for his help with coding

Thanks to Dr. Dahlquist for help with
GenMAPP and formatting issues

References

Domínguez-Ferreras, A., Pérez-Arnedo, R., Becker, A., Olivares, J., Soto, M.J., Sanjuán, J. (2006) Transcriptome Profiling Reveals the Importance of Plasmid pSymB for Osmoadaptation of *Sinorhizobium meliloti* *Journal of Bacteriology* 188:7617-7625

Galibert, F., Finan, T.M., Long, S., Puhler, A., et al. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti* *Science* 293:668-672