Expanding upon the microarray results of Yang et al. (2009) strengthened the gene relationships for *Shewanella oneidensis* in iron depleted and repleted states

Mary C Alverson, Joshua K Kuroda, Ronald A Legaspi, Emily R Simso

BIOL/CMSI 367-01 Biological Databases

December 18, 2015

**Abstract**

*Shewanella oneidensis* is a pathogen to humans and fish, but has not been extensively studied in the past. Yang et al. (2009) performed a microarray experiment to look how iron depletion and repletion affected biological processes for the organism. Three major modules were found (iron acquisition, anaerobic energy metabolism, and protein degradation), but this data did not completely explain or visualize how *S. oneidensis* deals with environmental stress. To get more information about specific genes, this study used GenMAPP to look at two genes for each of the above modules. The results from this analysis were similar to those from Yang et al., strengthening the argument that these are influential pathways in *S. oneidensis*. Additionally, this study used MAPPFinder to create gene pathways to look how changing iron levels affected genes, specifically the ribosome pathway and fatty acid degradation. This study found that while the three previously mentioned major modules are significantly impacted by iron depletion and repletion, there are many more genes involved as well. *S. oneidensis* used anaerobic energy metabolism and rapid iron uptake in the ribosome and fatty acid pathways, which was clearly confirmed by creating comprehensive gene pathways that used all of the time points from the original study. Creating a gene database for *S. oneidensis* was very beneficial for detailed analysis and can be used in future studies to learn more about the pathogen’s metabolism and other processes.

**Introduction**

*Shewanella oneidensis* is a relatively unknown organism, particularly in the realm of biological genomics. According to a microarray study done by Yang et al. in 2009, *S. oneidensis* is a pathogen to fish and humans, making it critical to study its pathways to better understand its interaction with these organisms. The genome of this organism was sequenced by Heidelberg et al. in 2002. In anaerobic conditions, S. oneidensis was shown to be able to reduce the levels of oxidized metals, which shows its current bioremediation capabilities (Heidelberg et al, 2002). The sequencing revealed a 51,857 base pair phage that is similar to the lambda phage present in Escherichia coli (Heidelberg et al, 2002). The Lambda-like phage was discovered to be both integrated in the S. oneidensis genome as well as present in a non-integrated form, suggesting that it is a functional phage (Heidelberg et al, 2002). This phage could potentially play an important role in genomically engineering S. oneidensis, which could increase S. oneidensis’ capability for bioremediation (Heidelberg et al, 2002).

The study by Yang et al. used a microarray of *S. oneidensis* to study how the pathogen maintains iron homeostasis, which is involved with metabolism and cell function. To do this, the researchers collected time points during an iron-deprived state using an iron-chelator and from an iron-repleted state using ferrous sulfate (Figure 1). They first determined that using a 160uM concentration of 2,2’-dipyridyl would allow for the most significant difference from the control. They used anaerobic culturing because *S. oneidensis* primarily uses anaerobic energy metabolism pathways, which they found from an initial test. To do this, the research team grew the cells to mid-log phase in 10ml LB medium with 10mM fumarate and 10mM lactate. They were spun down, washed with the LB medium, and divided into 5x10^7 aliquots. These were then transferred to a 5mL LB medium that contained 10mM lactate as an electron donor and 10mM Fe(III) dioxide as an electron acceptor. They could then use the 2,2’-dipyridyl to deplete the iron for a total of 60 minutes. They then repleted the iron using a ferrous sulfate solution for 60 minutes. Data was collected at 0, 5, 10, 20, 40, and 60 minutes for the depleted and repleted conditions (Yang et al, 2009).

Unfortunately, due to the lack of knowledge on *S. oneidensis*, its genome has not previously been compiled into a gene database, and so could not be analyzed using GenMAPP or MAPPFinder software. Therefore, the goal of this project was to use XMLPipeDB and GenMAPP Builder to create a gene database for *S. oneidensis* to better understand its biological processes. More information can be drawn from the microarray data using GenMAPP, allowing future researchers to discover more about *S. oneidensis*, as it is a relevant pathogen to human life.

**Materials and Methods**

We first downloaded the UniProt XML proteome set file (UniProt release 2015\_10), GO association file (GOA Proteome Sets 124), and the GO OBO-XML file (version 2015-11-01) on November 20, 2015. Next, we created a new database in PostgreSQL by executing the sql code taken from the sql folder of the latest GenMAPP Builder build. This code was run in PostgreSQL to create 167 empty tables. Now that we set the foundation for our database, we configured GenMAPP Builder to connect to our PostgreSQL database and imported the UniProt XML file, GOA file, and GO OBO-XML file using GenMAPP Builder. We were now able to export a GenMAPP gene database, making sure that it also exported all molecular function, cellular component, and biological process gene ontology terms. This process took one hour and 18 minutes.

Inspecting and validating our gene database was a long but significant process. Although we had successfully exported the database, it would mean nothing unless we verified that the data within the database was valid and accurate. The first check we made used the TallyEngine in GenMAPP Builder to record the number of records for UniProt and GO in the XML data and in the Postgres databases. The table (Image 1) we got from running TallyEngine confirmed that the XML and PostGres Ordered Locus counts were both 4196. The next check we made used the XMLPipeDB Match function to validate the results from the TallyEngine table. Initially, the regEx pattern we used only caught 4079 IDs because there were over 100 Ordered Locus names that contained the extra character ‘A’ in their ID. So we accounted for those IDs in the regEx pattern and got a count of 4207, which was 11 more than we were expecting. A quick look at the raw XML file told us that these 11 IDs were not picked up by the TallyEngine because they were missing gene tags, so the XMLPipeDB Match function recognized the pattern in another section that TallyEngine did not check. Our next check used an SQL query to validate the PostgreSQL database results from the TallyEngine. Using a regEx pattern similar to the one used for XMLPipeDB Match, we were able to get a confirmed count of 4196 IDs. Finally, we made a visual inspection of the gene database itself using Microsoft Access. We checked the UniProt, RefSeq, and OrderedLocusNames tables to make sure all of the IDs were in the correct form and found that there were no discrepancies.

To come to a logical conclusion in regards to the 11 IDs that were missing a gene tag in the XML file, we simply searched for each ID on the UniProt website and found that they were part of the "STRING" protein-protein interaction database. This meant that we could safely ignore these IDs in our database. A more detailed explanation of our verification process is provided in our Results section.

A comprehensive Gene Database Report can be found in the Appendix (Image 2).

The first step in organizing the data was to download each of the replicates for the given time points in both the Cy3 and Cy5 conditions from Array Express (<https://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-15334/>)  and compile them into an Excel spreadsheet. First, the unnecessary information at the top of each sheet was deleted. The columns used for the analysis were G, K, L, and H from the originally downloaded files, so these were copied into new sheets, one for each replicate at each time point. The data for the Cy3 and Cy5 for each replicate were then copied into new, separate spreadsheets so the data could be analyzed. The Log2 had to be calculated for each of the replicates to account for the signal and background medians and combine the data. First, the Cy5 background median was subtracted from the Cy5 signal median and the Cy3 background median was subtracted from the Cy3 signal median. The Cy5 value was then divided by the Cy3 value, which was then used to calculate the final value using Log2.

All of the Log2 values for the various time points were then collected into a single Excel spreadsheet. However, we took the replicates only from the C0, C5, C20, C60, F5, F20, and F60 conditions due to the vast amount of data. These time points show a range of the experiment and were believed to show significance. The data had to be split because there were two data points on each of the original microarray chips. Once we had the split data, we could calculate the biological average for each of the timepoints from the four replicates, giving us a total of seven time point averages. From the averages, we found the Average Log Ratio by comparing C5, C20, and C60 to C0 and F5, F20, and F60 to C60. We also calculated the p-value using a t-test of these same comparisons. From the p-value, we performed Bonferroni and Benjamini & Hochberg p-value corrections to ensure our data was valid. We used the Benjamini & Hochberg because it was a more stringent test of the relevancy of the results.

The files then had to be prepared for import into GenMAPP. We did this by creating a new sheet in our Excel file that contained the averages for the replicates of each time point, the biological averages for each time point, the Average Log Ratios, the t-tests, and the Benjamini & Hochberg p-value corrections. We saved this as a .txt file to import it into GenMAPP.

We imported our .txt file using Expression Dataset Manager into GenMAPP to look at the relevant results and see how the genes interacted within various pathways. We used MAPPFinder to interpret the results from GenMAPP. To determine which pathways were the most relevant, we looked at the top Gene Ontology terms based upon which factors increased or decreased the most during the microarray experiment. Based upon this analysis, we determined that the ribosome pathway and the fatty acid degradation pathway were the most significant. Using KEGG for *S. oneidensis*, we determined which genes were part of both of these pathways (Images 3 and 4). We placed the genes into MAPPFinder and created an Expression Dataset that included all of the comparison data. From this, we drew a ribosome pathway and a fatty acid degradation pathway, one that shows just the F60-C60 results and one that shows the results from all of the time comparisons.

**Results**

The schema of our gene database is represented visually in Figure 2. Each box represents a table in the gene database, with the lines connecting each box representing foreign keys. The schema was created with our GenMAPP Builder code as well as with the UniProt proteome set, GOA, and GO files, which have been previously discussed.

Table 1 shows the number of OrderedLocusNames ID’s found by XMLPipeDB match and TallyEngine in the UniProt XML file, the PostgreSQL database, our gene database and in the EnsemblBacteria database.  In order to generate these results using XMLPipeDB match to search for ID’s in the UniProt XML file, we used the command java -jar xmlpipedb-match-1.1.1.jar “SO\_A?[0-9][0-9][0-9][0-9]” < SOneidensisUNIPROT which was effective in detecting 4207 ID’s. When counting the ID’s in our gene database, we used the SQL query: select count(\*) from genenametype where type = 'ordered locus' and value ~ 'SO\_A?[0-9][0-9][0-9][0-9]'; in PGAdmin III which resulted in reporting 4196 ID’s. Figure (table) shows the number of ID’s in our gene database as being 8392 because a modification to the code of GenMAPP builder resulted in half of the ID’s containing underscores and the other half without. Half of the number of ID’s that exist is the number equal to the number of ID’s found by TallyEngine, which can be seen in Image 1, as well as the number of coding genes found in the external EnsemblBacteria database (Table 1).

In order to determine the cause for the inconsistency of the ID count found by XMLPipeDB match, we located the eleven OrderedLocusNames ID’s present in the XMLPipeDB search that were missing from our gene database.  The ID’s are : SO\_3699, SO\_1312,   SO\_4269, SO\_2875, SO\_4532, SO\_4580, SO\_2662, SO\_4423, SO\_3156, SO\_2967, and SO\_2024. Next we located each of these ID’s in the XML document itself and identified the reason that they were not found by TallyEngine as failing to be marked with a <gene> tag. In order to assess the importance of these ID’s, we searched for each of them in UniProt. We found that since each of them were only found in "protein-protein interaction” databases, they could be safely left out of our database.

Even though the OrderedLocusNames ID’s were successfully imported into our database without considerable code manipulation, we did in fact rewrite a section of GenMAPP Builder code so that our database would be able to accommodate more than one ID format. The changes, which can be viewed online at  <https://github.com/lmu-bioinformatics/xmlpipedb/blob/s-oneidensis/gmbuilder/src/edu/lmu/xmlpipedb/gmbuilder/databasetoolkit/profiles/ShewanellaOneidensisUniProtSpeciesProfile.java> , create a copy of each ID in the database and remove the underscore from each copied ID. This results in the number of ID’s in our database being double the amount of genes that are represented in our database.

The DNA microanalysis yielded a variety of results, beginning with the sanity check to see which data point comparison was most significant (Table 2). The F60-C60 time point was the most significant, as 788 genes had an average log fold change >0.25 and a pvalue <0.05 (a significant increase), and 963 genes with an average log fold change <-0.25 and a pvalue <0.05 (a significant decrease), as shown in Table 3. Tables 4 and 5 show the filtered list of MAPPFinder results of the non-redundant GO terms for the increased and decreased genes, respectively.

The GO terms that significantly increased includes ribosome, ribonucleoprotein complex, and structural component of ribosome (Table 4). The GO terms that significantly decreased includes fatty acid catabolic process and short-chain fatty acid metabolic process (Table 5). Based upon these results, ribosome and fatty acid degradation pathways were analyzed in more detail. The Yang et al. paper did not specifically look at either of these pathways, however, they did look at energy metabolism, protein synthesis, and central intermediary metabolism, all of which can be associated with the two pathways studied in this trial. Therefore, this study is more specific than the Yang et al. study simply because the filtered GO terms provided more in-depth information.

To look at this further, a GenMAPP MAPP was produced for fatty acid degradation and ribosome activity (Figures 3 and 4, respectively). Additionally MAPPs were made to compare all of the gene expression changes across the time points for these two pathways (Figures 5 and 6).

**Discussion**

According to Yang et al., there were three major modules for *S. oneidensis*: iron acquisition, anaerobic energy metabolism, and protein degradation. Therefore, we wanted to look at specific genes using GenMAPP that Yang et al. found to be significant to see if our results agreed with their findings (Table 6). According to our sanity check, the F60-C60 comparison had the most significant results, so we based our analysis upon this repletion time point. Yang et al. did not mention the criteria used to determine statistical significance, but values from 0-1 indicate gene expression decreases and values above 1 indicate increases. Therefore, our values will vary somewhat due to the different methods of analysis.

For iron acquisition, we looked at SO1111 and SO1784. According to GenMAPP, SO1111 increases iron ion homestasis, iron ion transport, the oxidation-reduction process, cellular components within the cell, metal ion binding, and ferric iron binding. The Yang et al. paper found that SO1111 increased by 5.61, while our results found the biological average to be 2.21 and the average log ratio as 1.69. This is understandable, as in repletion, the cell would need to use the increased iron levels in cell processes. In GenMAPP, SO1784 decreased in iron acquisition and is involved with iron ion homeostasis, transport, ferrous iron transport, ferrous iron transmembrane transport, and ferrous iron transmembrane transporter activity. Yang et al. found that SO1784 changed by 0.14, while the experimental results from the database found the biological average to be -1.15 and the average log ratio to be -1.7. Decreases in this gene are also logical in repletion, as the cell would not need to use as much energy to move iron around the cell, as it is more prevalent.

The two genes critically observed for anaerobic energy metabolism were SO0261 and SO0262. SO0261 increased heme transport, heme binding, and heme transporter activity according to GenMAPP. Yang et al. found its change to be 2.57, while the experimental results from this study found the biological average to be 0.74 and the average log ratio as 0.58. SO0261 should increase in repletion because, in anaerobic energy metabolism, the cell takens iron from alternative sources, such as hemoglobin. The second gene was SO0262 was not a part of the pathway during the repleted state at F60 according to Yang et al., however, it was present in the depleted state. GenMAPP found that it was relevant to cytochrome complex assembly, establishment of localization, specifically transport, the membrane, integral component of the membrane, plasma membrane, and heme transport activity. Additionally, our analysis found that SO0262 had a biological average at F60 of 0.44 and an average log ratio of 0.95. It is understandable that SO0262 would not be present in repletion, as the cell does not need to expend as much energy to maintain the membranes due to the increased iron levels.

Finally, protein degradation was analyzed using SO2016 and SO0052. SO2016 decreased by a factor of 0.06 in Yang et al.’s study and is involved in protein folding, response to stress, nucleotide binding, ATP binding, and unfolded protein binding in GenMAPP. Comparatively, it had a biological average of -0.74 and an average log ratio of -2.3. SO2016 would understandable decrease during protein degradation because there is less stress to the cell and less need to break down proteins for energy. SO0052 was not involved in repletion but was present in depletion and is involved in protein tetramerization, protein folding, transport, protein transport, and unfolded protein binding. Interestingly, the analysis performed in this test found a biological average of 1.05 and an average log ratio of -1.36. Finally, SO0052 is more logically part of depletion because the cell would need to use proteins as energy in its stressed condition.

The GenMAPP Builder process worked well for *S. oneidensis* because the data could be imported into the program and analyzed. GenMAPP allowed us to compare multiple replicates, giving a more complete picture of how iron depletion and repletion affect ribosome and fatty acid degradation in *S. oneidensis*. Specifically, *S. oneidensi*s increases genes in the ribosome pathway during the latter stages of repletion (see Figure 6). It also decreases fatty acid degradation in the late time points of repletion (see Figure 5). This connects to the conclusions by Yang et al. because ribosomes are an active part of protein production. It make sense that ribosome activity would increase at F20 and F60 because the cell would be able to expand energy for protein development. Similarly, the fatty acid degradation pathway shows decreased activity at F20 and F60, which relates to Yang et al.’s findings on metabolism. *S. oneidensis* does not need to break down fatty acids in the late repletion stage because it can get energy from iron. This pathway also strengthens the idea that *S. oneidensis* is able to use a variety of pathways for energy.

Yang et al. mentioned six genes involved in ribosome processes (10). The three genes with the most significant changes according to their analysis were SO1205, SO3927, and SO4120. However, only two of these genes (SO3927 and SO412) were marked as increased in the GenMAPP (see Table 7). The other gene that increased from these six was SO0227, which had a lower change in expression according to Yang et al. The experimental GenMAPP also had many more genes than Yang et al. discussed in their paper (Figure 4). More analysis would need to be done on these genes to see how they interact within *S. oneidensis*.

**Conclusion**

The findings for iron acquisition, anaerobic energy metabolism, and protein degradation correspond fairly well to the data found by Yang et al. The two differences were that Yang et al. did not have data for SO0262 and SO0052 for the F60 time point, while our analysis found that these genes did change. In terms of pvalues, there were some minor differences, but the general increasing and decreasing trends were the same between the two studies. The ribosome and fatty acid degradation MAPPs also strengthened these previously known gene relationships, however our study provided more information. For future analysis, it would be necessary to learn more about the criteria Yang et al. used for significance to better understand how our findings correlate to the original data.

A significant difference between the two data sets was that many more genes are involved in the ribosome pathway according to our analysis than Yang et al. reported. This may be because they were more interested in the other three major module pathways previously discussed or because there is so much information on ribosomes for organisms.

Yang et al. did not mention fatty acid degradation specifically in their paper, but incorporated this into other major categories. Therefore, it was interesting to discover that genes in *S. oneidensis* decrease fatty acid degradation only late in iron repletion. This may be because the cells need to break down the fatty acids for energy in the depletion and early repletion phase. Due to the rapid uptake of iron into *S. oneidensis*, the organism would be able to stop breaking down fatty acids once enough iron was re-introduced into the cells for energy.

Much more analysis could be done on *S. oneidensis* in terms of mapping the gene pathways. Since our report found many more genes involved in the ribosome, this would be an interesting place to start. Additionally, Yang et al. mentioned five genes specifically involved in anaerobic energy metabolism (10). A study could be done to see if these genes are involved in the fatty acid degradation pathway to better understand how that process connects to iron stress in *S. oneidensis*. Finally, other stressors could be introduced into *S. oneidensis*, such as other heavy metals, to compare the gene expression responses.

**Acknowledgments**

    The authors would like to thank Dr. Dahlquist (LMU Biology) and Dr. Dionisio (LMU Computer Science) for their help and guidance throughout the project, along with the Biology and Computer Science Departments at Loyola Marymount University for their resources.

**References**

"Fatty Acid Degradation - Shewanella Oneidensis." *KEGG*. Kanehisa Laboratories, n.d. Web. 13 Dec. 2015. <<http://www.genome.jp/kegg-bin/show_pathway?org_name=son&mapno=00071&mapscale=&show_description=show>>.

Heidelberg, J. F., Paulsen, I. T., Nelson, K. E., Gaidos, E. J., Nelson, W. C., Read, T. D., ... & Fraser, C. M. (2002). Genome sequence of the dissimilatory metal ion–reducing bacterium Shewanella oneidensis. *Nature biotechnology, 20*(11), 1118-1123. doi:10.1038/nbt749

"Ribosome - Shewanella Oneidensis." *KEGG*. Kanehisa Laboratories, n.d. Web. 13 Dec. 2015. <<http://www.genome.jp/kegg-bin/show_pathway?org_name=son&mapno=03010&mapscale=&show_description=show>>.

Yang, Yunfeng, et al. "Snapshot of iron response in Shewanella oneidensis by gene network reconstruction." *BMC genomics* 10.1 (2009): 131.

**Appendix**

Figure 1: experimental procedure for Yang et al. microarray study using the 160 µM concentration

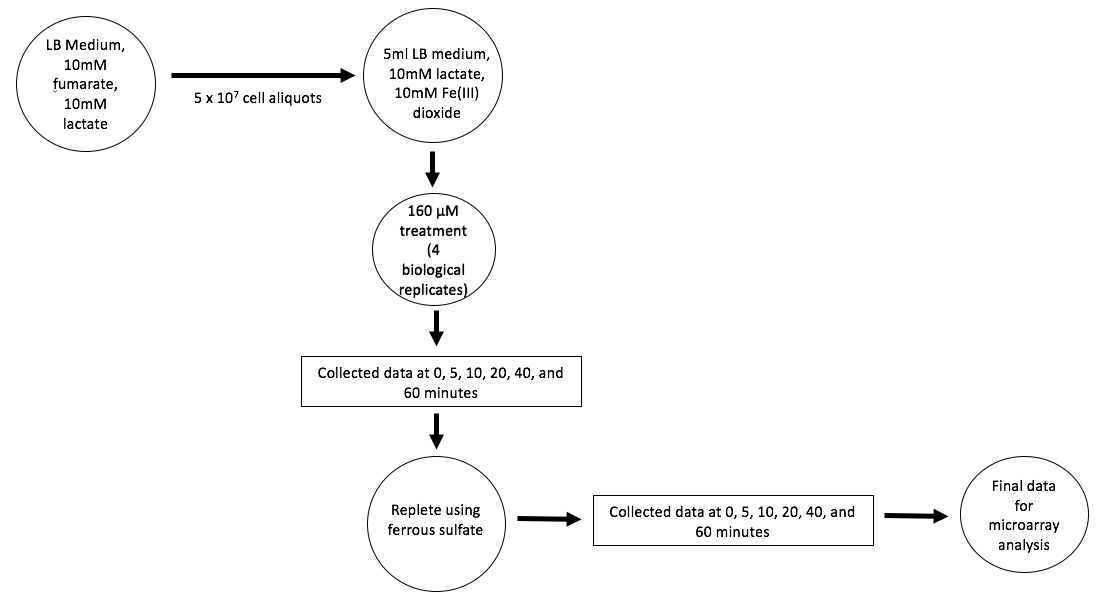


Image 1: TallyEngine results

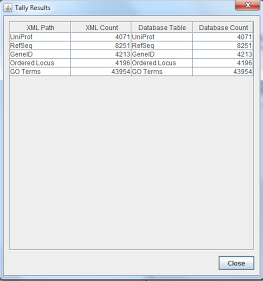
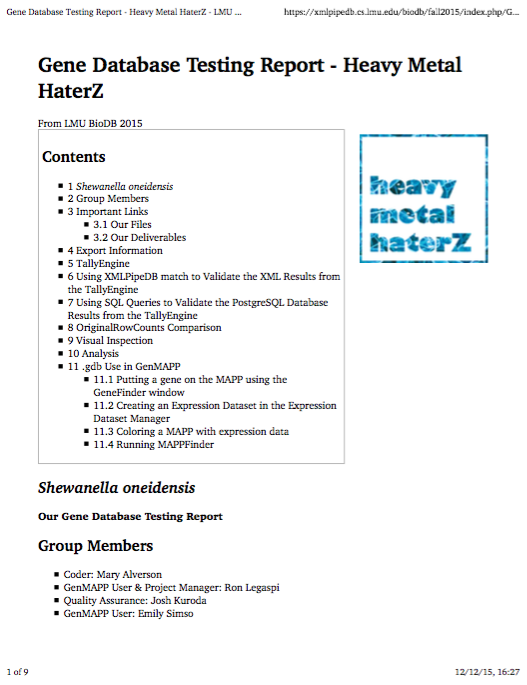
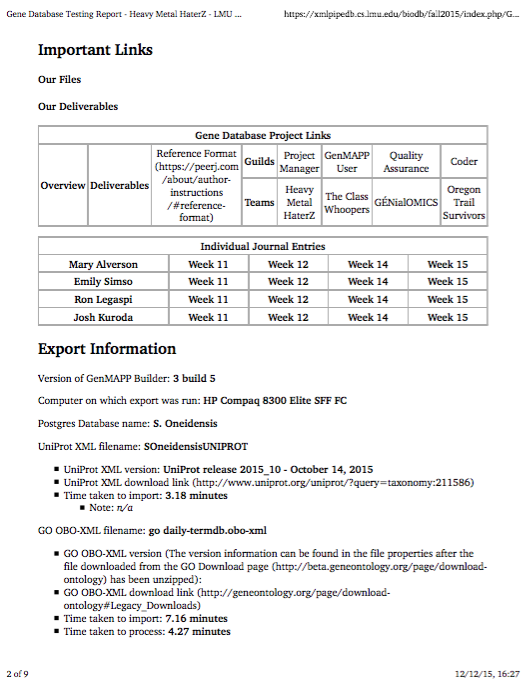
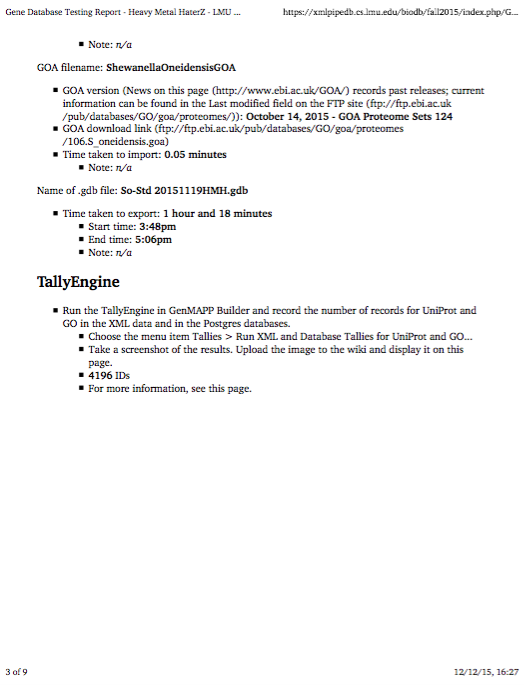
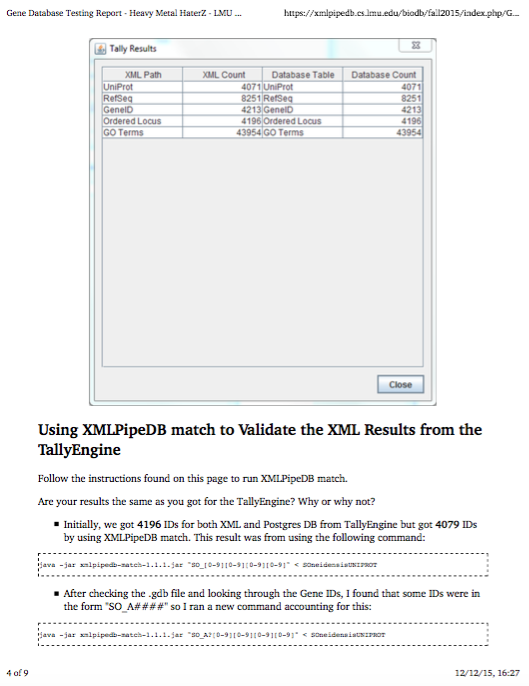


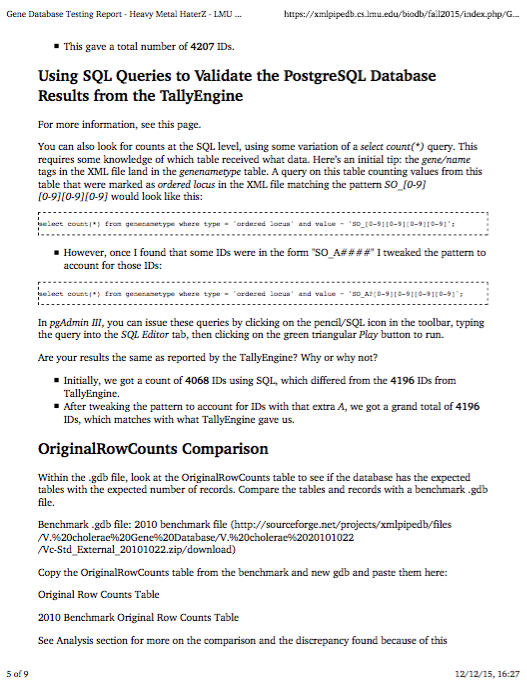
Image 2: Gene Database Report

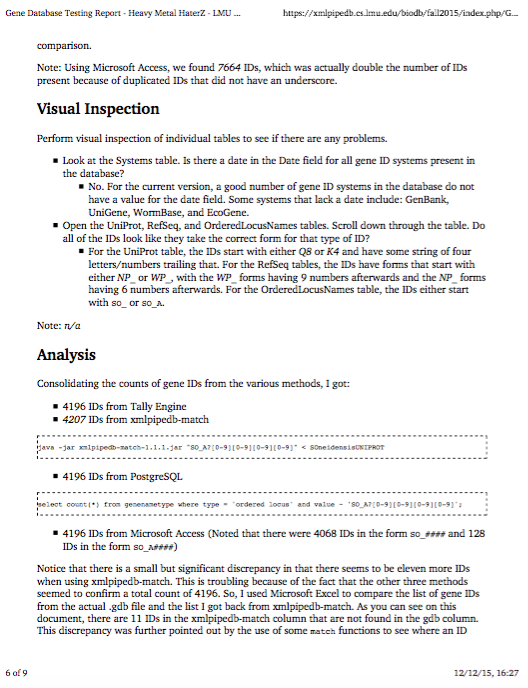


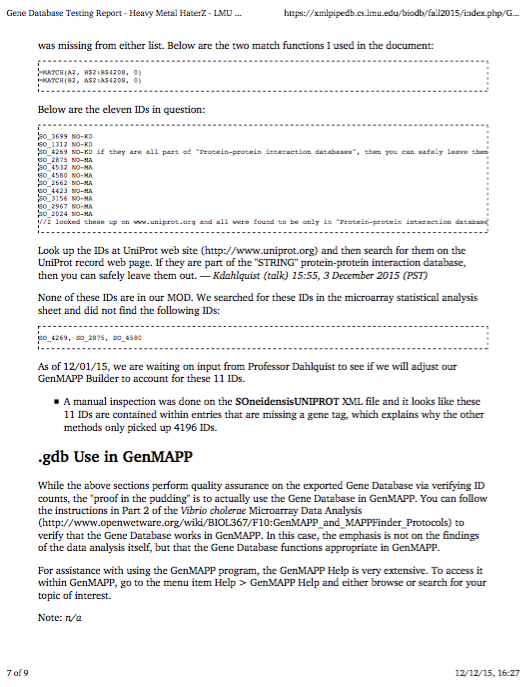


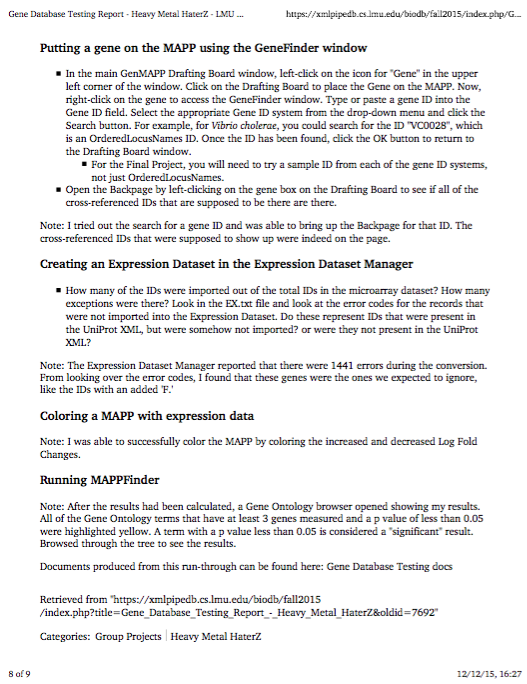












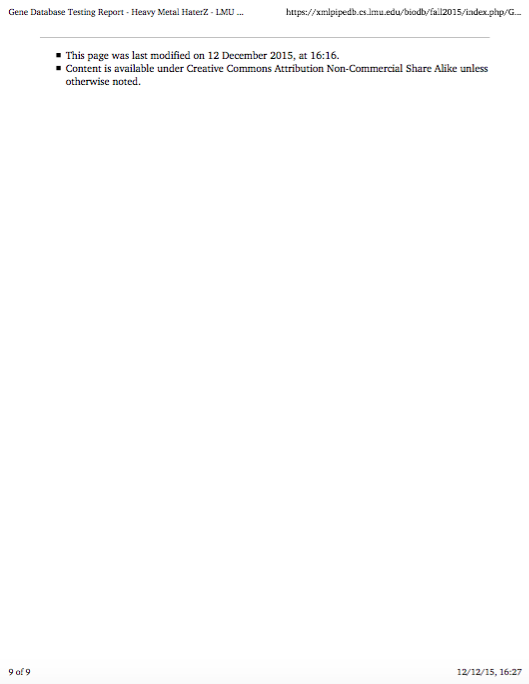


Image 3: ribosomal pathway for *S. oneidensis* from KEGG

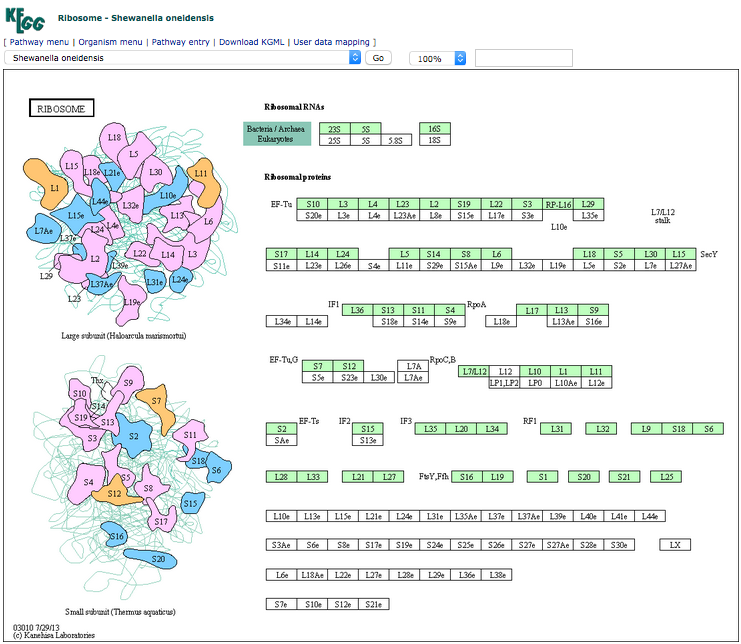


Image 4: fatty acid degradation pathway for *S. oneidensis* from KEGG

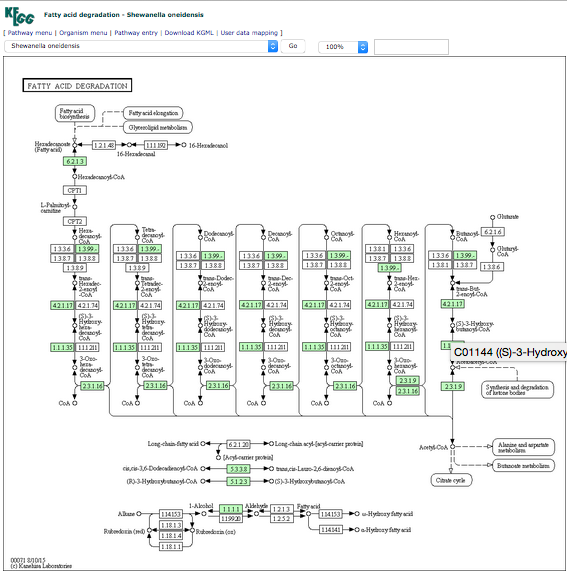


Figure 2: the gene database schema figure showing the relationships between the programs used

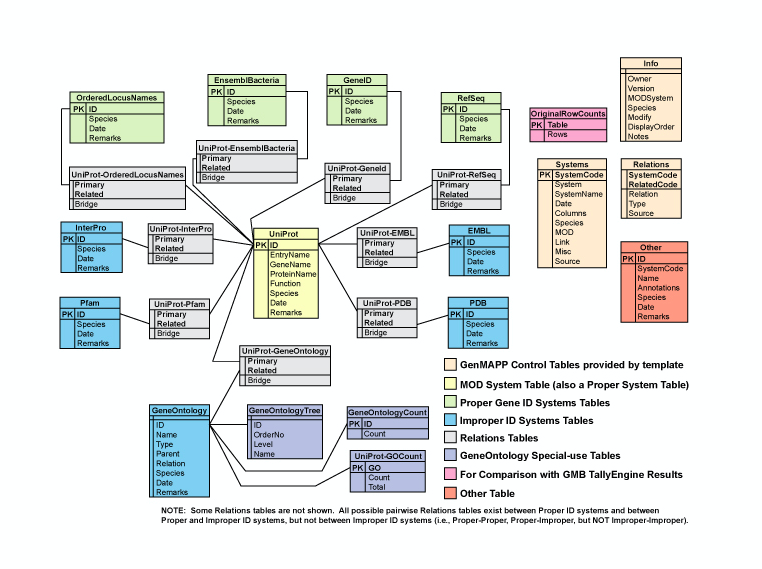


Table 1: OrderedLocusNames IDs were found by checking the values against other sources including the following EnsemblBacteria link

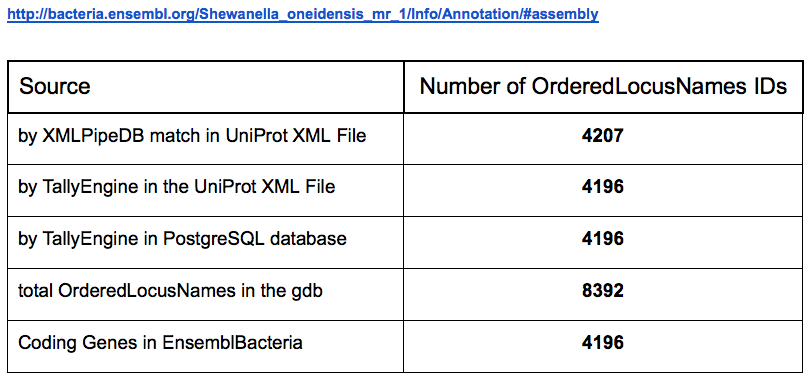


Table 2: the results of the sanity check, comparing the C5, C20, and C60 time points to C0 and the F5, F20, and F60 time points to C60. Based upon these results, F60-C60 will be used for future analysis because the genes had the most significant changes.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | How many genes have a pvalue<0.05 (with percentage) | How many genes have a pvalue<0.01 (with percentage) | How many genes have a pvalue<0.001 (with percentage) | How many genes have a pvalue <0.0001 (with percentage) | How many genes are p<0.05 for the Bonferroni-corrected p value (with percentage) | How many genes are p<0.05 for the Benjamini and Hochberg corrected p value (with percentage) | Genes with an average log fold change greater than zero | Genes with average log fold change less than zero | Genes with average log fold change of >0.25 and p<0.05 (with percentage) | Genes with average log fold change <-0.25 and p<0.05 (with percentage) |
| C5 and C0 | 344 genes, 6.36% | 94 genes, 1.74% | 18 genes, 0.33% | 5 genes, 0.09% | 2 genes, 0.04% | 2 genes, 0.037% | 180 genes, 3.33% | 164 genes, 3.03% | 161 genes, 2.98% | 149 genes, 2.76% |
| C20 and C0 | 868 genes, 16.05% | 342 genes, 6.32% | 79 genes, 1.46% | 14 genes, 0.26% | 1 gene, 0.01% | 34 genes, 0.63% | 452 genes, 8.36% | 416 genes | 437 genes, 7.69% | 405 genes, 7.49% |
| C60 and C0 | 1017 genes, 18.81% | 471 genes, 8.71% | 163 genes, 3.01% | 53 genes, 0.98% | 13 genes, 0.24% | 229 genes, 4.23% | 487 genes, 9.01% | 530 genes, 9.80% | 475 genes, 8.78% | 513 genes, 9.49% |
| F5 and C60 | 969 genes, 17.95% | 315 genes, 5.82% | 40 genes, 0.74% | 7 genes, 0.13% | 1 gene, 0.01% | 4 genes, 0.07% | 479 genes, 8.86% | 490 genes, 9.06% | 441 genes, 8.15% | 431 genes, 7.97% |
| F20 and C60 | 1838 genes, 33.99% | 892 genes, 16.49% | 239 genes, 4.42% | 54 genes, 1.00% | 10 genes, 0.18% | 707 genes, 13.07% | 826 genes, 15.27% | 1012 genes, 18.71% | 788 genes, 14.57% | 963 genes, 17.81% |
| F60 and C60 | 2070 genes, 38.28% | 1140 genes, 21.08% | 387 genes, 7.16% | 120 genes, 2.22% | 33 genes, 0.61% | 1193 genes, 22.06% | 870 genes, 16.09% | 1200 genes, 22.19% | 828 genes, 15.31% | 1146 genes, 21.19% |

Table 3: the criterea used to determine significant increases or decreases from the GenMAPP analysis

|  |  |  |  |
| --- | --- | --- | --- |
|  | Z score | PermuteP | Number Change |
| Increase | >2 | <0.05 | > or = 4 |
| Decrease | >2 | <0.05 | > or = 4 |

Table 4: the GO terms that significantly increased based upon the previously determined criteria

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| GOID | GO Name | Number Changed | Number Measured | Number in GO | %Changed | Percent Present | Z Score | Permute P | Adjusted P |
| 5840 | ribosome | 21 | 58 | 58 | 36.2069 | 100 | 4.8 | 0 | 0.08 |
| 30529 | ribonucleoprotein complex | 21 | 59 | 59 | 35.59322 | 100 | 4.706 | 0 | 0.081 |
| 3735 | structural constituent of ribosome | 19 | 54 | 54 | 35.18518 | 100 | 4.413 | 0 | 0.106 |
| 5198 | structural molecule activity | 20 | 65 | 65 | 30.76923 | 100 | 3.823 | 0 | 0.566 |
| 44444 | cytoplasmic part | 24 | 92 | 58 | 26.08696 | 158.6207 | 3.266 | 0 | 1 |
| 9156 | ribonucleoside monophosphate biosynthetic process | 13 | 39 | 10 | 33.33333 | 390 | 3.411 | 0.001 | 1 |
| 1902600 | hydrogen ion transmembrane transport | 9 | 24 | 24 | 37.5 | 100 | 3.255 | 0.002 | 1 |
| 30151 | molybdenum ion binding | 7 | 15 | 15 | 46.66667 | 100 | 3.586 | 0.003 | 0.684 |
| 9124 | nucleoside monophosphate biosynthetic process | 14 | 43 | 9 | 32.55814 | 477.7778 | 3.438 | 0.003 | 1 |
| 16836 | hydro-lyase activity | 11 | 34 | 6 | 32.35294 | 566.6667 | 3.019 | 0.003 | 1 |
| 1901137 | carbohydrate derivative biosynthetic process | 27 | 112 | 26 | 24.10714 | 430.7692 | 3.006 | 0.003 | 1 |
| 44205 | ’de novo’ UMP biosynthetic process | 4 | 6 | 6 | 66.66666 | 100 | 3.666 | 0.004 | 0.663 |
| 9311 | oligosaccharide metabolic process | 6 | 13 | 9 | 46.15385 | 144.4444 | 3.285 | 0.004 | 1 |
| 15078 | hydrogen ion transmembrane transporter activity | 9 | 25 | 9 | 36 | 277.7778 | 3.108 | 0.004 | 1 |
| 6818 | hydrogen transport | 9 | 26 | 26 | 34.61538 | 100 | 2.968 | 0.004 | 1 |
| 15992 | proton transport | 9 | 26 | 26 | 34.61538 | 100 | 2.968 | 0.004 | 1 |
| 19843 | rRNA binding | 12 | 39 | 39 | 30.76923 | 100 | 2.951 | 0.004 | 1 |
| 43232 | intracellular non-membrane-bounded organelle | 26 | 110 | 58 | 23.63636 | 189.6552 | 2.836 | 0.005 | 1 |
| 1901566 | organonitrogen compound biosynthetic process | 72 | 377 | 101 | 19.09814 | 373.2673 | 2.797 | 0.005 | 1 |
| 32991 | macromolecular complex | 32 | 141 | 2 | 22.69504 | 7050 | 2.899 | 0.006 | 1 |
| 43603 | cellular amide metabolic process | 33 | 150 | 101 | 22 | 148.5148 | 2.745 | 0.007 | 1 |
| 1901564 | organonitrogen compound metabolic process | 89 | 487 | 101 | 18.27515 | 482.1782 | 2.676 | 0.007 | 1 |
| 15412 | molybdate transmembrane-transporting ATPase activity | 4 | 7 | 7 | 57.14286 | 100 | 3.24 | 0.008 | 1 |
| 6412 | translation | 26 | 111 | 101 | 23.42342 | 109.901 | 2.784 | 0.008 | 1 |
| 43229 | intracellular organelle | 26 | 115 | 58 | 22.6087 | 198.2759 | 2.582 | 0.008 | 1 |
| 1901271 | lipooligosaccharide biosynthetic process | 5 | 10 | 9 | 50 | 111.1111 | 3.228 | 0.009 | 1 |
| 9247 | glycolipid biosynthetic process | 5 | 10 | 9 | 50 | 111.1111 | 3.228 | 0.009 | 1 |
| 46467 | membrane lipid biosynthetic process | 5 | 10 | 9 | 50 | 111.1111 | 3.228 | 0.009 | 1 |
| 6664 | glycolipid metabolic process | 5 | 10 | 9 | 50 | 111.1111 | 3.228 | 0.009 | 1 |
| 1901269 | lipooligosaccharide metabolic process | 5 | 10 | 9 | 50 | 111.1111 | 3.228 | 0.009 | 1 |
| 90407 | organophosphate biosynthetic process | 27 | 120 | 13 | 22.5 | 923.0769 | 2.605 | 0.009 | 1 |
| 6206 | pyrimidine nucleobase metabolic process | 4 | 8 | 4 | 50 | 200 | 2.886 | 0.01 | 1 |
| 43043 | peptide biosynthetic process | 26 | 115 | 101 | 22.6087 | 113.8614 | 2.582 | 0.01 | 1 |
| 9312 | oligosaccharide biosynthetic process | 5 | 11 | 9 | 45.45454 | 122.2222 | 2.955 | 0.011 | 1 |
| 6643 | membrane lipid metabolic process | 5 | 11 | 10 | 45.45454 | 110 | 2.955 | 0.012 | 1 |
| 16835 | carbon-oxygen lyase activity | 12 | 43 | 2 | 27.90698 | 2150 | 2.562 | 0.012 | 1 |
| 44769 | ATPase activity, coupled to transmembrane movement of ions, rotational mechanism | 4 | 7 | 7 | 57.14286 | 100 | 3.24 | 0.013 | 1 |
| 46933 | proton-transporting ATP synthase activity, rotational mechanism | 4 | 7 | 7 | 57.14286 | 100 | 3.24 | 0.013 | 1 |
| 42777 | plasma membrane ATP synthesis coupled proton transport | 4 | 7 | 7 | 57.14286 | 100 | 3.24 | 0.013 | 1 |
| 42625 | ATPase activity, coupled to transmembrane movement of ions | 9 | 27 | 9 | 33.33333 | 300 | 2.834 | 0.014 | 1 |
| 15238 | drug transmembrane transporter activity | 5 | 11 | 11 | 45.45454 | 100 | 2.955 | 0.015 | 1 |
| 15098 | molybdate ion transmembrane transporter activity | 4 | 8 | 8 | 50 | 100 | 2.886 | 0.015 | 1 |
| 8654 | phospholipid biosynthetic process | 9 | 28 | 10 | 32.14286 | 280 | 2.706 | 0.017 | 1 |
| 45259 | proton-transporting ATP synthase complex | 4 | 8 | 3 | 50 | 266.6667 | 2.886 | 0.018 | 1 |
| 15986 | ATP synthesis coupled proton transport | 4 | 8 | 8 | 50 | 100 | 2.886 | 0.018 | 1 |
| 16469 | proton-transporting two-sector ATPase complex | 4 | 8 | 3 | 50 | 266.6667 | 2.886 | 0.018 | 1 |
| 6754 | ATP biosynthetic process | 4 | 8 | 8 | 50 | 100 | 2.886 | 0.018 | 1 |
| 15985 | energy coupled proton transport, down electrochemical gradient | 4 | 8 | 8 | 50 | 100 | 2.886 | 0.018 | 1 |
| 42455 | ribonucleoside biosynthetic process | 11 | 39 | 8 | 28.20513 | 487.5 | 2.492 | 0.019 | 1 |
| 9129 | pyrimidine nucleoside monophosphate metabolic process | 5 | 12 | 6 | 41.66667 | 200 | 2.711 | 0.02 | 1 |
| 9130 | pyrimidine nucleoside monophosphate biosynthetic process | 5 | 12 | 6 | 41.66667 | 200 | 2.711 | 0.02 | 1 |
| 1903509 | liposaccharide metabolic process | 7 | 21 | 14 | 33.33333 | 150 | 2.497 | 0.02 | 1 |
| 9173 | pyrimidine ribonucleoside monophosphate metabolic process | 4 | 9 | 6 | 44.44444 | 150 | 2.585 | 0.021 | 1 |
| 46049 | UMP metabolic process | 4 | 9 | 6 | 44.44444 | 150 | 2.585 | 0.021 | 1 |
| 9174 | pyrimidine ribonucleoside monophosphate biosynthetic process | 4 | 9 | 6 | 44.44444 | 150 | 2.585 | 0.021 | 1 |
| 6222 | UMP biosynthetic process | 4 | 9 | 6 | 44.44444 | 150 | 2.585 | 0.021 | 1 |
| 44711 | single-organism biosynthetic process | 65 | 348 | 13 | 18.67816 | 2676.923 | 2.442 | 0.022 | 1 |
| 9163 | nucleoside biosynthetic process | 11 | 40 | 9 | 27.5 | 444.4445 | 2.396 | 0.022 | 1 |
| 1901659 | glycosyl compound biosynthetic process | 11 | 40 | 9 | 27.5 | 444.4445 | 2.396 | 0.022 | 1 |
| 9161 | ribonucleoside monophosphate metabolic process | 16 | 66 | 9 | 24.24242 | 733.3333 | 2.326 | 0.022 | 1 |
| 90484 | drug transporter activity | 5 | 12 | 11 | 41.66667 | 109.0909 | 2.711 | 0.024 | 1 |
| 6855 | drug transmembrane transport | 6 | 16 | 16 | 37.5 | 100 | 2.655 | 0.024 | 1 |
| 15689 | molybdate ion transport | 4 | 9 | 9 | 44.44444 | 100 | 2.585 | 0.025 | 1 |
| 9123 | nucleoside monophosphate metabolic process | 17 | 70 | 9 | 24.28572 | 777.7778 | 2.407 | 0.025 | 1 |
| 43604 | amide biosynthetic process | 29 | 138 | 101 | 21.01449 | 136.6337 | 2.292 | 0.025 | 1 |
| 6644 | phospholipid metabolic process | 9 | 31 | 4 | 29.03226 | 775 | 2.351 | 0.027 | 1 |
| 6518 | peptide metabolic process | 26 | 120 | 101 | 21.66667 | 118.8119 | 2.34 | 0.027 | 1 |
| 42493 | response to drug | 6 | 17 | 16 | 35.29412 | 106.25 | 2.477 | 0.028 | 1 |
| 15893 | drug transport | 6 | 17 | 16 | 35.29412 | 106.25 | 2.477 | 0.028 | 1 |
| 8610 | lipid biosynthetic process | 15 | 62 | 16 | 24.19355 | 387.5 | 2.242 | 0.028 | 1 |
| 46390 | ribose phosphate biosynthetic process | 13 | 51 | 10 | 25.4902 | 510 | 2.296 | 0.029 | 1 |
| 9206 | purine ribonucleoside triphosphate biosynthetic process | 4 | 9 | 8 | 44.44444 | 112.5 | 2.585 | 0.032 | 1 |
| 9145 | purine nucleoside triphosphate biosynthetic process | 4 | 9 | 8 | 44.44444 | 112.5 | 2.585 | 0.032 | 1 |
| 46493 | lipid A metabolic process | 4 | 9 | 9 | 44.44444 | 100 | 2.585 | 0.032 | 1 |
| 9245 | lipid A biosynthetic process | 4 | 9 | 9 | 44.44444 | 100 | 2.585 | 0.032 | 1 |
| 6730 | one-carbon metabolic process | 4 | 9 | 9 | 44.44444 | 100 | 2.585 | 0.037 | 1 |
| 9069 | serine family amino acid metabolic process | 7 | 23 | 5 | 30.43478 | 460 | 2.216 | 0.037 | 1 |
| 44283 | small molecule biosynthetic process | 36 | 183 | 71 | 19.67213 | 257.7465 | 2.124 | 0.04 | 1 |
| 6213 | pyrimidine nucleoside metabolic process | 5 | 15 | 7 | 33.33333 | 214.2857 | 2.109 | 0.04 | 1 |
| 96 | sulfur amino acid metabolic process | 6 | 18 | 8 | 33.33333 | 225 | 2.311 | 0.041 | 1 |
| 44391 | ribosomal subunit | 5 | 14 | 7 | 35.71429 | 200 | 2.292 | 0.041 | 1 |
| 44281 | small molecule metabolic process | 79 | 450 | 22 | 17.55556 | 2045.455 | 2.095 | 0.043 | 1 |
| 6526 | arginine biosynthetic process | 4 | 10 | 10 | 40 | 100 | 2.324 | 0.044 | 1 |
| 9165 | nucleotide biosynthetic process | 15 | 65 | 13 | 23.07692 | 500 | 2.037 | 0.047 | 1 |
| 43565 | sequence-specific DNA binding | 11 | 42 | 42 | 26.19048 | 100 | 2.212 | 0.048 | 1 |
| 98796 | membrane protein complex | 9 | 33 | 3 | 27.27273 | 1100 | 2.137 | 0.048 | 1 |
| 6525 | arginine metabolic process | 5 | 15 | 14 | 33.33333 | 107.1429 | 2.109 | 0.048 | 1 |
| 34248 | regulation of cellular amide metabolic process | 5 | 14 | 6 | 35.71429 | 233.3333 | 2.292 | 0.049 | 1 |
| 6417 | regulation of translation | 5 | 14 | 6 | 35.71429 | 233.3333 | 2.292 | 0.049 | 1 |

Table 5: the GO terms that significantly decreased based upon the previously determined criteria

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| GOID | GO Name | Number Changed | Number Measured | Number in GO | Percent Changed | Percent Present | Z Score | Permute P | Adjusted P |
| 17076 | purine nucleotide binding | 87 | 347 | 309 | 25.07205 | 112.2977 | 4.761 | 0 | 0.029 |
| 9062 | fatty acid catabolic process | 7 | 10 | 7 | 70 | 142.8571 | 4.642 | 0 | 0.038 |
| 19184 | nonribosomal peptide biosynthetic process | 4 | 4 | 2 | 100 | 200 | 4.567 | 0 | 0.103 |
| 72329 | monocarboxylic acid catabolic process | 7 | 12 | 1 | 58.33333 | 1200 | 3.985 | 0 | 0.226 |
| 160 | phosphorelay signal transduction system | 32 | 116 | 117 | 27.58621 | 99.1453 | 3.415 | 0 | 0.767 |
| 35556 | intracellular signal transduction | 32 | 118 | 119 | 27.11864 | 99.15966 | 3.305 | 0 | 0.834 |
| 5057 | receptor signaling protein activity | 18 | 57 | 50 | 31.57895 | 114 | 3.201 | 0 | 1 |
| 7154 | cell communication | 40 | 165 | 153 | 24.24242 | 107.8431 | 2.906 | 0 | 1 |
| 32555 | purine ribonucleotide binding | 86 | 346 | 309 | 24.85549 | 111.9741 | 4.639 | 0.001 | 0.039 |
| 46459 | short-chain fatty acid metabolic process | 4 | 4 | 3 | 100 | 133.3333 | 4.567 | 0.001 | 0.103 |
| 19541 | propionate metabolic process | 4 | 4 | 3 | 100 | 133.3333 | 4.567 | 0.001 | 0.103 |
| 30554 | adenyl nucleotide binding | 78 | 310 | 309 | 25.16129 | 100.3236 | 4.521 | 0.001 | 0.115 |
| 32550 | purine ribonucleoside binding | 84 | 342 | 309 | 24.5614 | 110.6796 | 4.454 | 0.001 | 0.143 |
| 35639 | purine ribonucleoside triphosphate binding | 84 | 342 | 309 | 24.5614 | 110.6796 | 4.454 | 0.001 | 0.143 |
| 1883 | purine nucleoside binding | 84 | 342 | 309 | 24.5614 | 110.6796 | 4.454 | 0.001 | 0.143 |
| 1882 | nucleoside binding | 84 | 343 | 309 | 24.4898 | 111.0032 | 4.424 | 0.001 | 0.143 |
| 32549 | ribonucleoside binding | 84 | 343 | 310 | 24.4898 | 110.6452 | 4.424 | 0.001 | 0.143 |
| 32559 | adenyl ribonucleotide binding | 77 | 309 | 309 | 24.91909 | 100 | 4.393 | 0.001 | 0.144 |
| 5524 | ATP binding | 77 | 309 | 309 | 24.91909 | 100 | 4.393 | 0.001 | 0.144 |
| 32553 | ribonucleotide binding | 89 | 372 | 309 | 23.92473 | 120.3884 | 4.314 | 0.001 | 0.152 |
| 44242 | cellular lipid catabolic process | 7 | 11 | 1 | 63.63636 | 1100 | 4.294 | 0.001 | 0.159 |
| 16042 | lipid catabolic process | 7 | 11 | 5 | 63.63636 | 220 | 4.294 | 0.001 | 0.159 |
| 97367 | carbohydrate derivative binding | 89 | 375 | 309 | 23.73333 | 121.3592 | 4.227 | 0.001 | 0.161 |
| 30163 | protein catabolic process | 6 | 9 | 8 | 66.66666 | 112.5 | 4.131 | 0.001 | 0.209 |
| 15891 | siderophore transport | 6 | 9 | 9 | 66.66666 | 100 | 4.131 | 0.001 | 0.209 |
| 15688 | iron chelate transport | 6 | 9 | 9 | 66.66666 | 100 | 4.131 | 0.001 | 0.209 |
| 5515 | protein binding | 15 | 42 | 4 | 35.71429 | 1050 | 3.476 | 0.001 | 0.747 |
| 16772 | transferase activity, transferring phosphorus-containing groups | 50 | 208 | 122 | 24.03846 | 170.4918 | 3.199 | 0.001 | 1 |
| 8170 | N-methyltransferase activity | 8 | 19 | 6 | 42.10526 | 316.6667 | 3.09 | 0.002 | 1 |
| 51716 | cellular response to stimulus | 49 | 208 | 153 | 23.55769 | 135.9477 | 3.005 | 0.002 | 1 |
| 1901678 | iron coordination entity transport | 8 | 17 | 9 | 47.05882 | 188.8889 | 3.479 | 0.003 | 0.745 |
| 4872 | receptor activity | 28 | 100 | 99 | 28 | 101.0101 | 3.278 | 0.003 | 0.837 |
| 50896 | response to stimulus | 62 | 278 | 153 | 22.30216 | 181.6993 | 2.917 | 0.004 | 1 |
| 60089 | molecular transducer activity | 33 | 126 | 93 | 26.19048 | 135.4839 | 3.131 | 0.005 | 1 |
| 23052 | signaling | 37 | 153 | 153 | 24.18301 | 100 | 2.773 | 0.006 | 1 |
| 7165 | signal transduction | 37 | 153 | 153 | 24.18301 | 100 | 2.773 | 0.006 | 1 |
| 44700 | single organism signaling | 37 | 153 | 153 | 24.18301 | 100 | 2.773 | 0.006 | 1 |
| 16775 | phosphotransferase activity, nitrogenous group as acceptor | 16 | 53 | 51 | 30.18868 | 103.9216 | 2.808 | 0.007 | 1 |
| 4176 | ATP-dependent peptidase activity | 4 | 6 | 6 | 66.66666 | 100 | 3.372 | 0.008 | 0.828 |
| 5315 | inorganic phosphate transmembrane transporter activity | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.008 | 1 |
| 1901677 | phosphate transmembrane transporter activity | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.008 | 1 |
| 30246 | carbohydrate binding | 8 | 21 | 22 | 38.09524 | 95.45454 | 2.749 | 0.008 | 1 |
| 97588 | archaeal or bacterial-type flagellum-dependent cell motility | 9 | 24 | 24 | 37.5 | 100 | 2.86 | 0.009 | 1 |
| 48870 | cell motility | 9 | 24 | 24 | 37.5 | 100 | 2.86 | 0.009 | 1 |
| 1539 | cilium or flagellum-dependent cell motility | 9 | 24 | 24 | 37.5 | 100 | 2.86 | 0.009 | 1 |
| 51674 | localization of cell | 9 | 24 | 24 | 37.5 | 100 | 2.86 | 0.009 | 1 |
| 6928 | movement of cell or subcellular component | 9 | 24 | 24 | 37.5 | 100 | 2.86 | 0.009 | 1 |
| 51082 | unfolded protein binding | 6 | 12 | 13 | 50 | 92.30769 | 3.199 | 0.01 | 1 |
| 16072 | rRNA metabolic process | 11 | 32 | 32 | 34.375 | 100 | 2.823 | 0.01 | 1 |
| 6364 | rRNA processing | 11 | 32 | 32 | 34.375 | 100 | 2.823 | 0.01 | 1 |
| 70011 | peptidase activity, acting on L-amino acid peptides | 26 | 104 | 42 | 25 | 247.619 | 2.501 | 0.01 | 1 |
| 34440 | lipid oxidation | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.011 | 1 |
| 6635 | fatty acid beta-oxidation | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.011 | 1 |
| 19395 | fatty acid oxidation | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.011 | 1 |
| 16783 | sulfurtransferase activity | 6 | 13 | 9 | 46.15385 | 144.4444 | 2.952 | 0.011 | 1 |
| 4672 | protein kinase activity | 18 | 62 | 57 | 29.03226 | 108.7719 | 2.791 | 0.011 | 1 |
| 6468 | protein phosphorylation | 18 | 62 | 61 | 29.03226 | 101.6393 | 2.791 | 0.011 | 1 |
| 18202 | peptidyl-histidine modification | 13 | 42 | 42 | 30.95238 | 100 | 2.632 | 0.011 | 1 |
| 18106 | peptidyl-histidine phosphorylation | 13 | 42 | 42 | 30.95238 | 100 | 2.632 | 0.011 | 1 |
| 155 | phosphorelay sensor kinase activity | 15 | 50 | 50 | 30 | 100 | 2.69 | 0.012 | 1 |
| 23014 | signal transduction by protein phosphorylation | 15 | 50 | 50 | 30 | 100 | 2.69 | 0.012 | 1 |
| 9401 | phosphoenolpyruvate-dependent sugar phosphotransferase system | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.014 | 1 |
| 6457 | protein folding | 10 | 29 | 28 | 34.48276 | 103.5714 | 2.702 | 0.014 | 1 |
| 38023 | signaling receptor activity | 15 | 51 | 50 | 29.41176 | 102 | 2.602 | 0.014 | 1 |
| 31167 | rRNA methylation | 8 | 22 | 20 | 36.36364 | 110 | 2.592 | 0.014 | 1 |
| 8649 | rRNA methyltransferase activity | 8 | 22 | 8 | 36.36364 | 275 | 2.592 | 0.014 | 1 |
| 154 | rRNA modification | 8 | 22 | 20 | 36.36364 | 110 | 2.592 | 0.014 | 1 |
| 4673 | protein histidine kinase activity | 15 | 51 | 51 | 29.41176 | 100 | 2.602 | 0.015 | 1 |
| 8134 | transcription factor binding | 5 | 11 | 11 | 45.45454 | 100 | 2.652 | 0.018 | 1 |
| 71973 | bacterial-type flagellum-dependent cell motility | 8 | 23 | 23 | 34.78261 | 100 | 2.444 | 0.019 | 1 |
| 3684 | damaged DNA binding | 4 | 7 | 7 | 57.14286 | 100 | 2.956 | 0.02 | 1 |
| 16773 | phosphotransferase activity, alcohol group as acceptor | 23 | 90 | 65 | 25.55556 | 138.4615 | 2.468 | 0.02 | 1 |
| 3887 | DNA-directed DNA polymerase activity | 6 | 14 | 14 | 42.85714 | 100 | 2.728 | 0.022 | 1 |
| 34061 | DNA polymerase activity | 6 | 14 | 14 | 42.85714 | 100 | 2.728 | 0.022 | 1 |
| 46915 | transition metal ion transmembrane transporter activity | 4 | 8 | 3 | 50 | 266.6667 | 2.611 | 0.022 | 1 |
| 8168 | methyltransferase activity | 22 | 89 | 88 | 24.7191 | 101.1364 | 2.236 | 0.024 | 1 |
| 32259 | methylation | 22 | 89 | 89 | 24.7191 | 100 | 2.236 | 0.024 | 1 |
| 16741 | transferase activity, transferring one-carbon groups | 24 | 99 | 88 | 24.24242 | 112.5 | 2.231 | 0.025 | 1 |
| 16226 | iron-sulfur cluster assembly | 4 | 8 | 8 | 50 | 100 | 2.611 | 0.026 | 1 |
| 3774 | motor activity | 4 | 8 | 8 | 50 | 100 | 2.611 | 0.026 | 1 |
| 31163 | metallo-sulfur cluster assembly | 4 | 8 | 8 | 50 | 100 | 2.611 | 0.026 | 1 |
| 6508 | proteolysis | 29 | 126 | 122 | 23.01587 | 103.2787 | 2.145 | 0.026 | 1 |
| 50789 | regulation of biological process | 83 | 424 | 225 | 19.57547 | 188.4444 | 2.059 | 0.027 | 1 |
| 65007 | biological regulation | 85 | 435 | 225 | 19.54023 | 193.3333 | 2.068 | 0.028 | 1 |
| 71897 | DNA biosynthetic process | 6 | 15 | 15 | 40 | 100 | 2.522 | 0.029 | 1 |
| 42254 | ribosome biogenesis | 12 | 43 | 40 | 27.90698 | 107.5 | 2.117 | 0.032 | 1 |
| 22613 | ribonucleoprotein complex biogenesis | 12 | 43 | 40 | 27.90698 | 107.5 | 2.117 | 0.032 | 1 |
| 8643 | carbohydrate transport | 5 | 12 | 8 | 41.66667 | 150 | 2.412 | 0.034 | 1 |
| 6817 | phosphate ion transport | 4 | 9 | 9 | 44.44444 | 100 | 2.315 | 0.034 | 1 |
| 6631 | fatty acid metabolic process | 10 | 34 | 22 | 29.41176 | 154.5455 | 2.12 | 0.038 | 1 |
| 4871 | signal transducer activity | 23 | 98 | 93 | 23.46939 | 105.3763 | 2.009 | 0.038 | 1 |
| 16435 | rRNA (guanine) methyltransferase activity | 4 | 9 | 2 | 44.44444 | 450 | 2.315 | 0.048 | 1 |
| 17111 | nucleoside-triphosphatase activity | 36 | 164 | 76 | 21.95122 | 215.7895 | 2.081 | 0.048 | 1 |

Figure 3: the MAPPFinder pathway showing gene changes for F60-C60 in fatty acid degradation



Figure 4: MAPPFinder pathway showing gene changes for F60-C60 for a ribosome

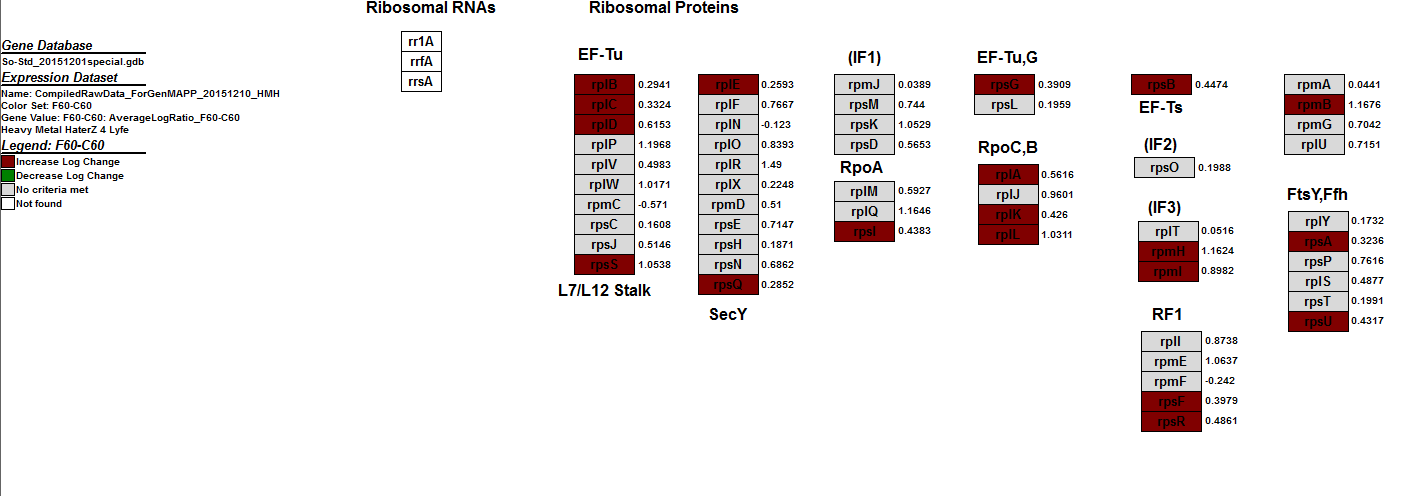


Figure 5: compiled MAPPFinder pathway showing the gene changes for all time point comparisons for fatty acid degradation

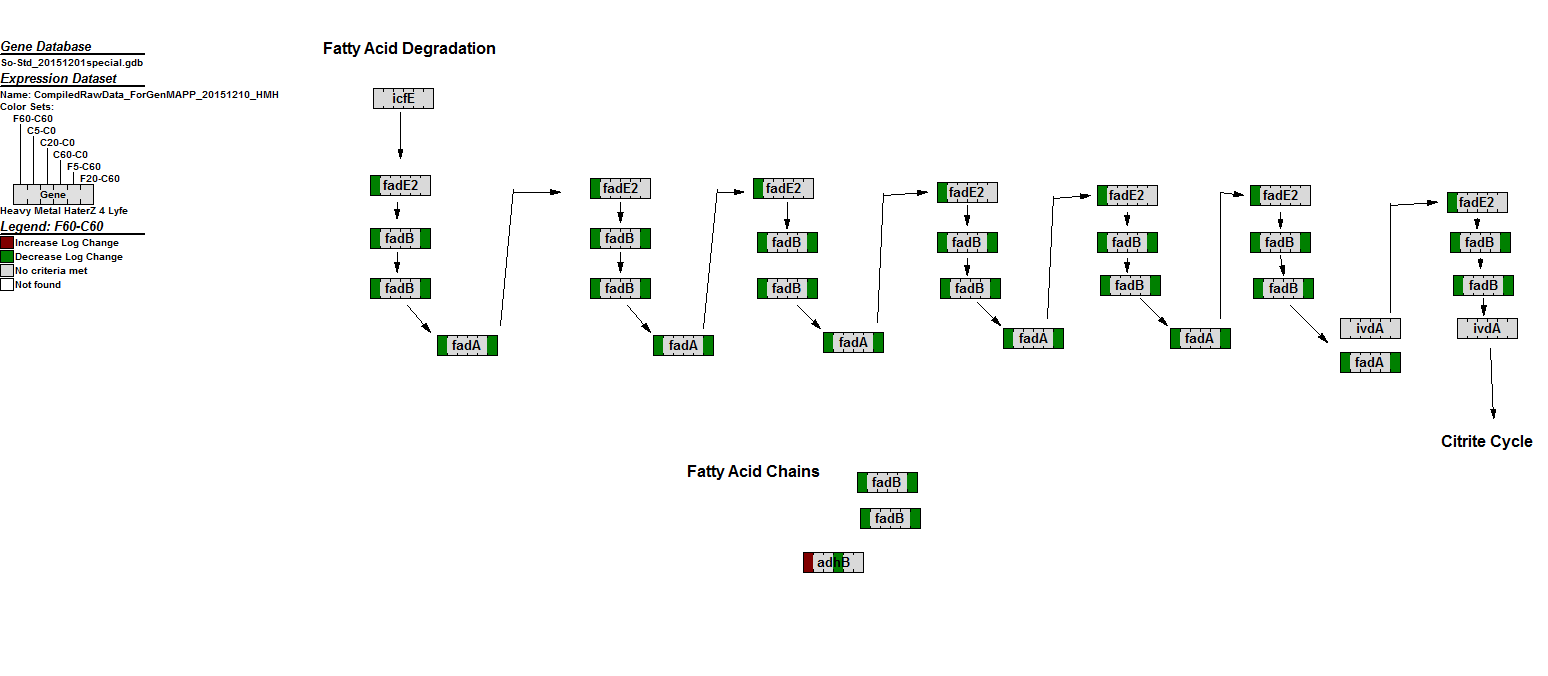


Figure 6: compiled MAPPFinder pathway showing the gene changes for all time point comparisons for a ribosome

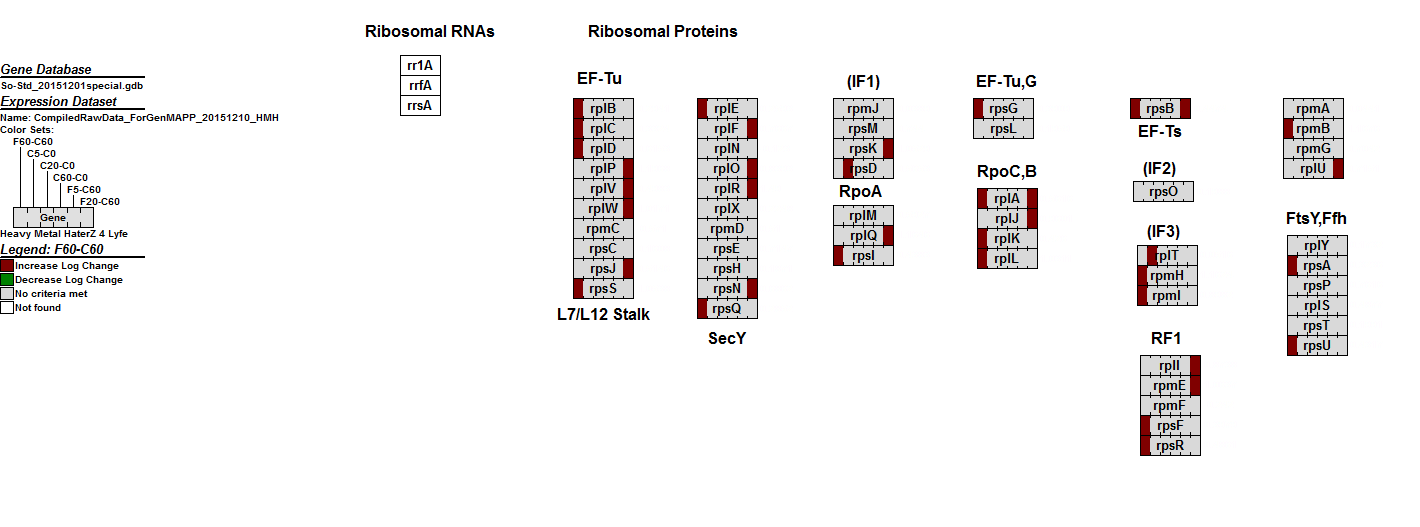


Table 6: changes in gene expression between Yang et al. and the experimental study

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Gene | Change in expression according to Yang et al. at F60 | Calculated F60 biological average | Calculated Average Log Ratio F60-C60 | Calculated Ttest F60-C60 |
| SO111 | 5.61 | 2.21 | 1.69 | 0 |
| SO1784 | 0.14 | -1.15 | -1.7 | 0 |
| SO0261 | 2.57 | 0.74 | 0.58 | 0.002 |
| SO0262 | NA | 0.44 | 0.95 | 0.0006 |
| SO0052 | NA | 1.05 | -1.36 | 0.0081 |
| SO2016 | 0.06 | -0.74 | -2.3 | 0 |

Table 7: changes in gene expression for genes relevant to the ribosome pathway

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Gene | Change in expression according to Yang et al. for F60 | Presence in ribosome GenMAPP | Calculated F60 biological average | Calculated Average Log Ratio F60-C60 | Calculated Ttest F60-C60 |
| SO0227 | 2.37 | Increased | 2.32 | 0.39 | 0.048 |
| SO0233 | NA | No criterea met | 3.03 | 1.02 | 0 |
| SO0246 | 2.87 | No criterea met | 2.34 | 0.77 | 0.0035 |
| SO1205 | 3.56 | Not in MAPP | 1.05 | 0.61 | 0.0088 |
| SO3927 | 3.83 | Increased | 1.82 | 0.87 | 0.0002 |
| SO4120 | 6 | Increased | 1.04 | 1.06 | 0.0031 |