Virtual Institute of Microbial Stress and Survival
DOE Genomes To Life Project
Progress Report: May, 2003

I. Overview

The objective of this monthly progress report is to provide an update of the technical and administrative actions from the previous month as well as forecast upcoming progress for the VIMSS Genomes To Life Project. I want to remind everyone how important to make sure everyone is communicating. The discussion boards (http://genomics.lbl.gov/~aparkin/discus) provide a forum for people to ask questions about direction of the project, priorities, and technical issues that can be read and answered by the entire group. I know email is often the most efficient means but it does privatize some of the important communications. Also, posting project data and information to BioFiles (https://tayma.lbl.gov/perl/biofiles) is EXTREMELY important. We are in the process of adding user help files to BioFiles – if you have user questions, please contact Keith Keller (tel: 510.495.2766 or email: kkeller@lbl.gov). This is the best metric I can give to the DOE leadership that we are making progress aside from the VIMSS website. Please make us and yourselves visible by donating data and information to the website.

II. Applied Environmental Microbiology Core

LBNL

SR-FTIR. We continue to modify the existing SR-FTIR spectromicroscopy apparatus to study *Desulfovibrio vulgaris* under anaerobic conditions. We are testing the modified system using reference samples. Succeeded in confirming that the source of the problem is on the thickness of the infrared-transparent crystal window and the distance between the window and the sample surface. This problem is corrected for, and the system is ready.

The group at LBNL helped draft a proposed experimental plan to members of the group for oxygen stress related experiments. The final edition can be seen on VIMMS discussion board. This initial experiment will be a batch experiment in 500 ml glass jars. The control will be anaerobic growth under N_2 purge and the stressed condition will be continuous aeration. The experiment will be conducted for 24hr, at 30°C, and in tripicate.

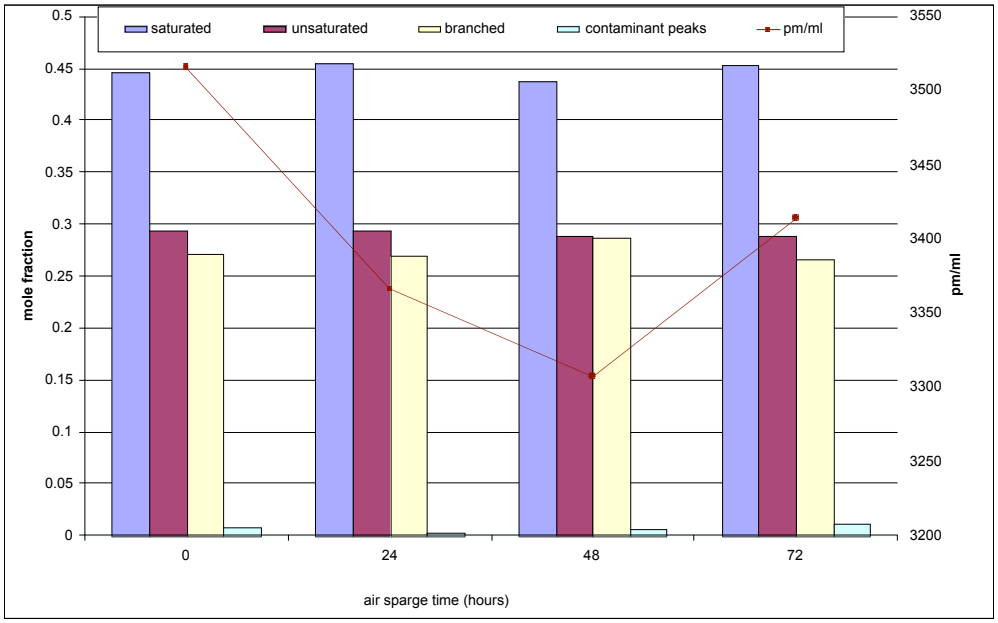
Much of the effort for the month has involved ordering necessary equipment for the large scale production of *D vulgaris* and determining the best way to perform accurate oxygen measurements of gas and aqueous phase of cultures, controlled delivery of gasses to experimental cultures for oxygen stress related experiments under sulfate reducing conditions.

Effort was also spent in working with Judy Wall's group to work out a standard protocol for the Lactate sulfate growth medium.

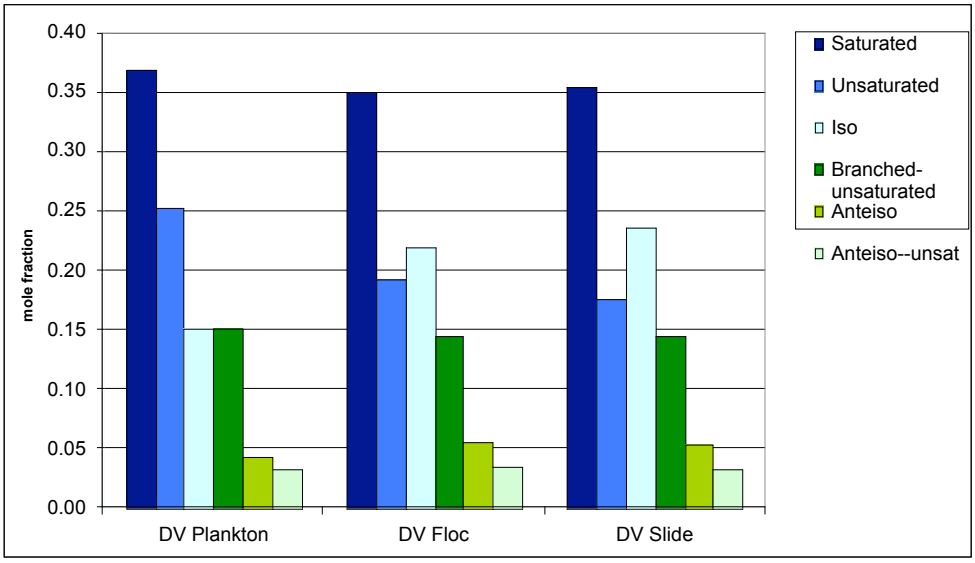
Work was continued to investigate the application of PLFA techniques to monitor responses of *D vulgaris* to oxygen stress and to various growth conditions. The oxygen stress experiment was extended to 72 hr. For this test, duplicate cultures were grown for three days in modified Baars medium. The aeration was staged so that all samples were harvested simultaneously to avoid growth time bias. PLFA results indicate no significant changes in lipid composition or viable biomass during aeration.

Due to the consistent appearance of signature lipids for pure cultures of *D. vulgaris*, spurious peaks were combined to indicate contamination. This may be useful as a complement to traditional methods to routinely monitor the chemostat culture.

Summary of PLFA signatures of D vulgaris under air sparging.



As part of the development of the design of the chemostats, lipid signatures of D. vulgaris were also compared during different growth conditions. Cells were grown in a container on a stir plate and stirred at a moderate speed. Glass slides were suspended to provide a vertical surface for cell growth. After 7 days, the culture was divided into three fractions: planktonic, attached, and flocculated. Shifts were seen in the type of lipids depending on the growth condition, especially in the fraction of unsaturated lipids and branched (iso) lipids.



University of Washington

MPN enrichments which were set up in March and April for sediment cores from FB609-10-00 (background area, core depth 180-239in), FB060-01-00 (area 1, depth 180-228 in), FB052-01-00 (area 2, depth 245-300 in) and FB056-01-00 (area 3, depth 194-198 in) on B2 and supplemented with one of the following substrates: 1) lactate, 2) acetate, 3) propionate 4) pyruvate and 5) hydrogen with carbon dioxide still show no SRB activity. To evaluate hydrogen sulfide production, $CuSO_4$ assays were preformed on all enrichments supplemented with lactate. Although growth has been observed in some of the enrichments, all test negative for sulfide.

The enrichments from ground water TPB-25Y of FRC area 2 and sediments FW-109 and 107 of FRC area 3 demonstrating sulfate reducing activity (provided by Dr. M. Fields) have been transferred to B3 medium amended with lactate and ethanol. Roll tubes were inoculated using an enrichment dilution series for colony isolation.

To explore resistance of *Desulfovibrio* species to high concentration of several salts, potential stressors in environment, the effect of high concentrations of sodium chloride (NaCl), magnesium sulfate ($MgSO_4$), and sodium nitrate ($NaNO_3$) on the growth of *D. vulgaris* Hildenborough and *Desulfovibrio* sp. PT2 was characterized. Growth of both strains was not affected at 40g/L magnesium sulfate. However, growth of both was slowed at 10 and 15g/L of sodium chloride or sodium nitrate, suggesting that the concentration of sodium ion was the major stressing factor. 20g/L of sodium chloride completely inhibited the growth of both strains. 20g/L and 30g/L of sodium nitrate completely inhibited growth of *D. vulgaris* and PT-2 respectively.

We initiated studies of two different *Desulfovibrio* species (*D. vulgaris* Hildenborough and *Desulfovibrio* sp. PT2) growing syntrophically on lactate with a hydrogenotrophic methanogen (*Methanococcus maripaludis*) in the absence of sulfate. Cultures were grown on three sets of media; 'full' medium (50 mM lactate, 50 mM acetate, 25 mM sulfate, 80%/20% H_2/CO_2 head space), no-sulfate medium (50 mM lactate, N_2 head space), and limiting-sulfate medium (50 mM lactate, 4 mM sulfate, N_2 head space). The limiting sulfate medium was evaluated as a mechanism to stimulate initial growth of SRB to promote establishment of a syntrophic co-culture. Cultures were monitored over time for optical density (OD_{600}), CO_2, H_2, methane, lactate, acetate, and sulfate. The production of methane on the sulfate-minus medium, and following depletion of sulfate on the sulfate-limiting medium, confirmed the establishment of syntrophic growth. Gas production was accompanied by modest cell growth. In contrast, very little growth (0.1 OD_{600}) of either *Desulfovibrio* sp. was observed on 50 mM lactate monocultures. Samples were taken from duplicate cultures for planned Northern blot analysis to assess differential expression of genes in the sulfate reduction pathway under the different growth conditions.

One fixed bed bioreactor has been assembled and hardware and software have been tested.

The FairMenTec Bioreactor installation and training has been completed.

Immediate future work
Continue development of co-culture growth conditions

16S rDNA and dsrAB genes of approximately 13 SRB isolated from Lake Depue metal-contaminated sediments on a lactate medium will be PCR amplified and sequenced.

The fixed bed reactor will be tested for its ability to maintain prolonged anaerobic conditions. Fresh sediment cores from the FRC have been received and will be used to inoculate the reactor.

The FairMenTec Bioreactor will be tested by running batch cultures of *D. vulgaris Hildenborough* and *Desulfovibrio* strain PT-2.

New SRB enrichments from fresh FRC cores will be set up on sulfate medium amended with1) lactate, 2) acetate, 3) propionate 4) pyruvate and 5) hydrogen with carbon dioxide

Oak Ridge National Laboratory

No update.

Diversa

A summary of the DNA extractions from the LBNL soil cores follows:

Sample ID	Core Segment	Core Depth	Date Received	Shipped By	Comments	Small Insert DNA	Small Insert Library	Large Insert DNA	Large Insert Library
FWB302-03 Top	? (~10" long)	?	?	LBNL	Background	4ng/□l		2nd-in progress	
Bottom						3ng/□l		2nd-in progress	
FB609-01-00 Top	24-43" long	180-234"	3/5/2003	LBNL	Background,	5&4ng/□l	in progress	Complete	in progress
Bottom					Anaerobic	4ng/□l x2	in progress	Complete	in progress
FB060-01-27 Top	27-46" long	180-228 "	4/10/2003	LBNL	<25ppm U238	5ng/□l	in progress	2nd-in progress	
Bottom					Anaerobic	5 ng/□l	in progress	2nd-in progress	
FB052-01-12 Top	12-36" long	245-300"	4/10/2003	LBNL	<25ppm U238	5ng/□l	in progress	2nd-in progress	
Bottom					Anaerobic	5 ng/□l	in progress	2nd-in progress	
FB056-01-34 Top	34-58" long	144-198"	4/10/2003	LBNL	<25ppm U238	2ng/□l	in progress	Complete	in progress
Bottom						2ng/□l	in progress	Complete	in progress

- *S.diversa* genomic DNA, which is GC rich (~70%), has been amplified and hybridized to an Affymetric GeneChip to look at genome coverage. In this experiment, a portion of the *S.diversa* genome was not amplified, so high GC random hexamers are being tried to increase the genome coverage of high GC DNA.

- Diversity indexing of the amplified large fragment gDNA from sample FB609 showed that we successfully amplified undetectable amounts of gDNA of multiple organisms from an environmental sample.

Issues

Inefficient cloning of blunt-ended amplified DNA for library construction may require several rounds of DNA amplification or library preparation to obtain efficient library titers.

Actions

- Large and small insert DNA libraries are being constructed for several of the samples from amplified DNA.

- Diversity indexing will be used to evaluate amplification coverage and bias by comparing DNA from direct small fragment extractions, amplified small fragments, and amplified large fragments.

- Work is ongoing to optimize the large insert FACS biopanning protocol in gel microdroplets. Recovery of DNA directly from hybridized GMDs has been shown to be feasible, but inefficient, so work is ongoing to optimize RCA protocols.

III. Functional Genomics Core

OVERVIEW OF THE CORE

Significant Milestones reached
Significant milestones were reached by the Functional Genomics core over the first six months of research. These include:

- The construction of a *Shewanella oneidensis* MR-1 microarray and extensive characterization of gene expression profiles resulting from to various environmental stresses such as pH, salt (osmotic) and H_2O_2-induced oxidative stress (under the support of the Microbial Cell Project).

- The design and ordering of oligonucleotide probes for the construction of the *Desulfovibrio vulgaris* whole-genome microarrays.

- The completion of the initial survey of the total proteome (soluble, urea-solubilized and SDS-solubilized) of *D. vulgaris* by 3D-LC-MS/MS.

- The initiation of experiments aimed at establishing and validating quantitative proteomics methodologies to quantify the change of protein expression upon cell stress.

- The development of separation and detection methods for nucleotides and coenzyme-A metabolites.

- The development of a simple and rapid metabolite extraction procedure.

Research integration of the core
As the facilities of the Functional Genomics core group are well underway in establishing their respective protocols and methodologies, the first steps at integrating the efforts of the core was initiated. A first experiment was design to investigate stress changes that exposure of *D. vulgaris* to air (O_2–stress). The rationale and objectives of the experiment are described below and details can be found in the protocol posted on the VMISS discussion board

Rationale
Since most contaminated subsurface environments are episodically exposed to oxygenated water via infiltration of snow melt and rain from the surface, and from recharge areas upgradient where O_2 has not been depleted, O_2 is a common stressor for anaerobes like *D. vulgaris* in the environment. In addition, metal and radionuclide reduction by bacteria is dependent upon fairly stringent reducing (anaerobic) conditions which can be upset by any exposure to oxygen. *D. vulgaris* is tolerant of oxygen via a number of metabolic mechanisms, negative aerotaxis, and via production of EPS. The exact response sequence that *D. vulgaris* and how oxygen affects that response are unknown.

Objectives
The overall objective of this first experiment is to document the array of changes that occur in cells exposed to air for 24h using the array of techniques being developed by the project. The objectives of this experiment are:
Define basic O_2 stress changes that occur as measured by the techniques being developed, establish *D. vulgaris* baseline parameters for the techniques being tested, determine logistical constraints for sampling, shipping etc., establish QA/QC verification of assay protocols, develop functional design criteria for future experiments.

PROGRESS OF EACH FACILITY

Transcriptomics (ORNL)

Objectives

- To perform initial microarray expression profiling studies for the model bacterium *S. oneidensis* MR-1 to establish a baseline response to various environmental stresses.

- To create and use whole-genome microarrays for *Desulfovibrio* and *Geobacter* for analysis of stress-induced transcriptomes.

Progress since last report

- Completed detailed analyses of microarray expression data from pH stress experiments comparing growth of *S. oneideinsis* under neutral pH (pH 7) to growth under pH 4 or pH 10. A subset of genes differentially expressed in response to pH stress has been selected for further expression validation using real-time PCR. Primers for this subset of genes have already been synthesized, and we are currently in the process of developing the protocol for real-time PCR. The writing of a manuscript describing this work also is currently underway.

- Initial microarray experiments analyzing the MR-1 transcriptome under heat shock conditions have been completed and the data analyzed using Genespring. *rpoH* and *rpoE* deletion mutant strains of *S. oneidensis* have been created, and the response of these mutant strains to thermal stress will be measured using whole-genome microarrays.

- Preliminary growth studies looking at the response of wt MR-1 to strontium toxicity have been completed. These experiments indicate that MR-1 growth is inhibited strongly at strontium chloride concentrations of about 200 mM in both minimal and rich media.

Strontium speciation will be analyzed using atomic absorption, and microarray experiments are in the process of being designed.

- Microarray experiments analyzing the effect of H_2O_2-induced oxidative stress on global gene expression patterns in wt *S. oneideinsis* have been completed and some additional array experiments are planned. *oxyR* and *dps* deletion mutant strains of *S. oneidensis* have been generated. Phenotype characterization and microarray expression analysis are underway.

- Manuscript describing salt shock response of *S. oneidensis* is in the process of being written.

- The order for synthesis of oligonucleotide probes for *Desulfovibrio* whole-genome microarrays has been sent to MWG.

Future work

- Complete pH stress manuscript. Design more detailed time-series experiments.

- Complete detailed data analysis for oxidative stress microarray experiments.

- Begin microarray expression profiling of *S. oneideinsis* cells exposed to toxic levels of certain metals (e.g., strontium).

- Target more genes for deletion mutagenesis based on expression data.

- Construct *Desulfovibrio* whole-genome microarrays.

Proteomics (Diversa)

Overview: Upon the completion of sample preparation optimization and software development, our activities moved towards the method development in quantitative proteomics apart from remaining analysis of *D. vulgaris* proteome.

Objectives

- Complete proteome analysis of *D. vulgaris*.

- Finish data analysis for SDS fraction (membrane proteins) of *D. vulgaris*.

- Continue to validate chemical/isotope labeling method(s) on *D. vulgaris,* which potentially. quantify the change of protein expression upon cell stress.

Progress since last report

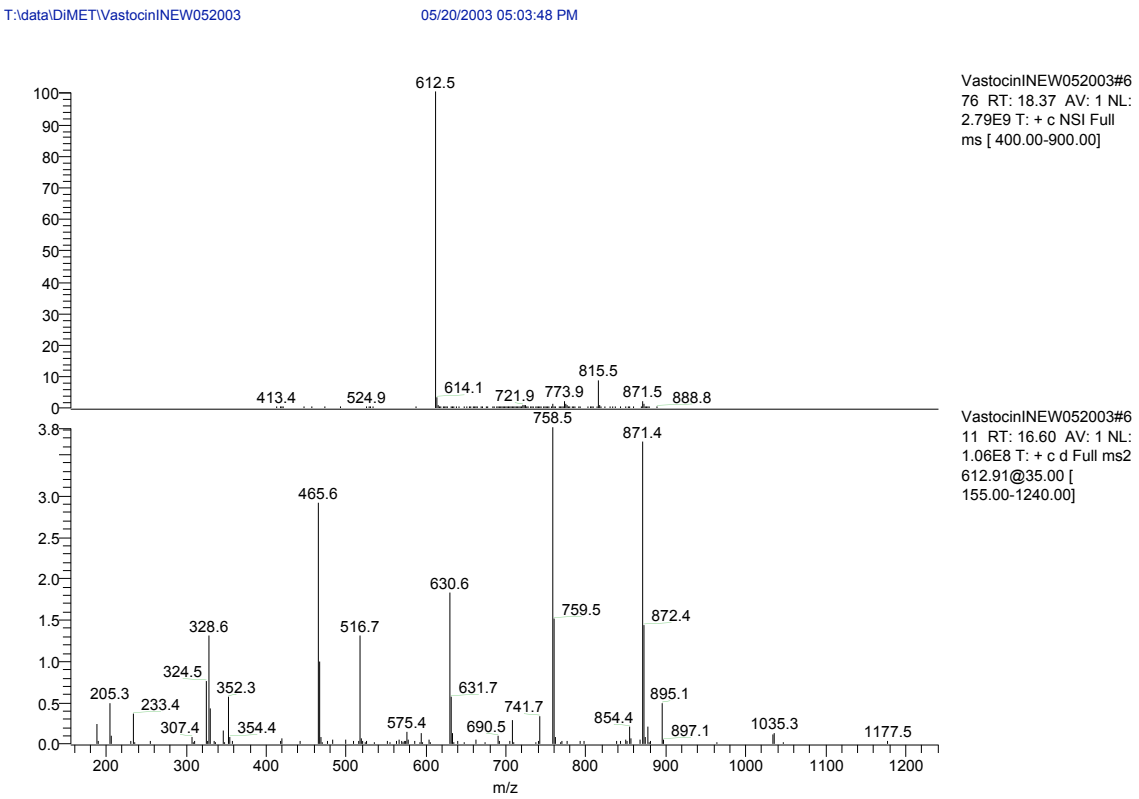
- Proteins identified from SDS fraction of *D. vulgaris*: From the SDS-solubilized proteins, we identified 538 unique proteins with an average coverage of 4 peptides per protein from a database with 3674 genes. More specifically, 132647 raw MS/MS spectra resulted in 1832 unique tryptic peptides. (The results will be deposited in Biofiles)

- Method evaluation in chemical/isotope labeling procedures for differential quantitative analysis.

- *Esterification methods*: it target C-terminal and carboxyl group of peptides. We further optimized the protocols. The yield was approaching 100% on standard peptides. We also applied this on standard protein digests for higher complexity, and data is being analyzed.

- *Sulfohydral labeling reagents*: we have designed several new compounds to label the reduced cysteine side chain to introduce the differential mass. One of the compounds has been synthesized and reacted well with standard peptides and proteins.

Figure 1. MS and MS/MS of $(M+2H)^{2+}$
Peptide [Arg^8]-Vas: Cys**[Cxm]-**Tyr-Ile-Gln--Asn-Cys**[Cxm]-**Pro-Arg-Gly-amide



Protein complexes (Sandia)

Objectives

- Identify target proteins to be tagged for protein complex identification.

- Use proteomics techniques like ICAT and others to identify proteins that are overexpressed under stress conditions of heat, O2, pH and UV radiation.

Progress since last report

D. vulgaris

- Two dimensional differential in gel electrophoresis runs performed on whole cell protein extracts obtained from *D. vulgaris* under heat shock (Cy3) and non heat shock (Cy5) conditions. Currently processing trypsinized gel extracts using MALDI to identify proteins exhibiting differential expression levels.

- Optimization of ICAT procedures:

- Column based separations:
  i. Set up in-house HPLC systems for complex peptide separations using polysulfoethyl columns
  ii. Packed in-house nano columns with a variety of packing materials to test for peptide separations
  iii. Set up in-house nano-LC systems to be used in conjunction with QTOF

E. coli

- Obtained activated magnetic beads (tosyl, amine etc) for protein complex identification and related equipment for bead separation experiments.

- Labeled tosyl activated magnetic beads with RecA antibody for immunoprecipitation experiments on whole cell protein extracts obtained earlier.

Metabolomics and Proteomics (UCB, LBNL)

Objectives

- Develop methods for the identification and quantification of nucleotides and derivatives.

- Plan preliminary experiments to generate gene knockouts of *D. vulgaris* including cloning and allelic exchange strategies using barcodes to identify the mutants.

- Work towards completing experiments for oxygen stress response (integrated activity of the core).

- Develop a list of putative two-component systems as initial gene knockout targets (with computational core).

Progress since last report

Metabolite analysis

- Different HPLC columns (cyano, amino, amide, C-30 and cyclodextrin columns) were evaluated to separate 19 nucleotides. All columns except the amide column can not separate the nucleotides.

- Negative scan mode was found to more sensitive than positive scan mode for nucleotides. A MRM negative mode detection conditions were established to efficiently detect the nucleotides.

- A normal-phase HPLC method using amide-80 column was developed to partially separate 19 nucleotides and then selectively detect them by MS (Figure 2).

- An ion-suppressor was evaluated for replacing sodium ions in solvent with protons to apply mass spectrometry with ion-exchange HPLC. It seems the ion-suppressor works well to exchange sodium ions up to 140 mM of NaOH solution. However, peaks become broader after the ion exchange (tested by AMP), probably caused by increased acidity.

Gene knockouts

- An OD_{600} growth curve for *D. vulgaris* was generated at 37°C. Because the organism generates precipitates as it grows, the experiment will need to be repeated, with the inclusion of a protein assay for each time point.

- In preparation to generate gene knockouts in *D. vulgaris* was cultured on lactate-sulfate plates in the anaerobic chamber.

- A strategy to study global stress response in *D. vulgaris* is being developed. In this a library of knock-out mutants for all the two component systems will be systematically generated in *D. vulgaris* using a high throughput method. Towards this goal, we have designed a generic, ligation free, 2-step PCR strategy for the knock-out mutagenesis of any/all genes in *D. vulgaris*. This proposed strategy was distributed the GTL members for feedback.

- Furthermore, to generate a list of candidate Histidine Sensor Kinases the CRITICA annotated FASTA files uploaded to the Biofiles site were used in addition to the TIGR annotated files. A collaborative effort has been started with Eric Alm to narrow down the list of candidate HSKs to be used in the knock out strategy

Proteomics

- Several of the parameter settings resulting in discrepancy in protein ID lists from the ProICAT and ProID have been resolved. An unresolved software glitch in ProICAT causes incorrect ID of Heavy tagged peptides. Work is in progress to solve this issue.

- The LC-QSTAR instrument has been switched to 2D-nanoLC and the method is now being tested with simple commercial protein mixture and soluble protein fractions.

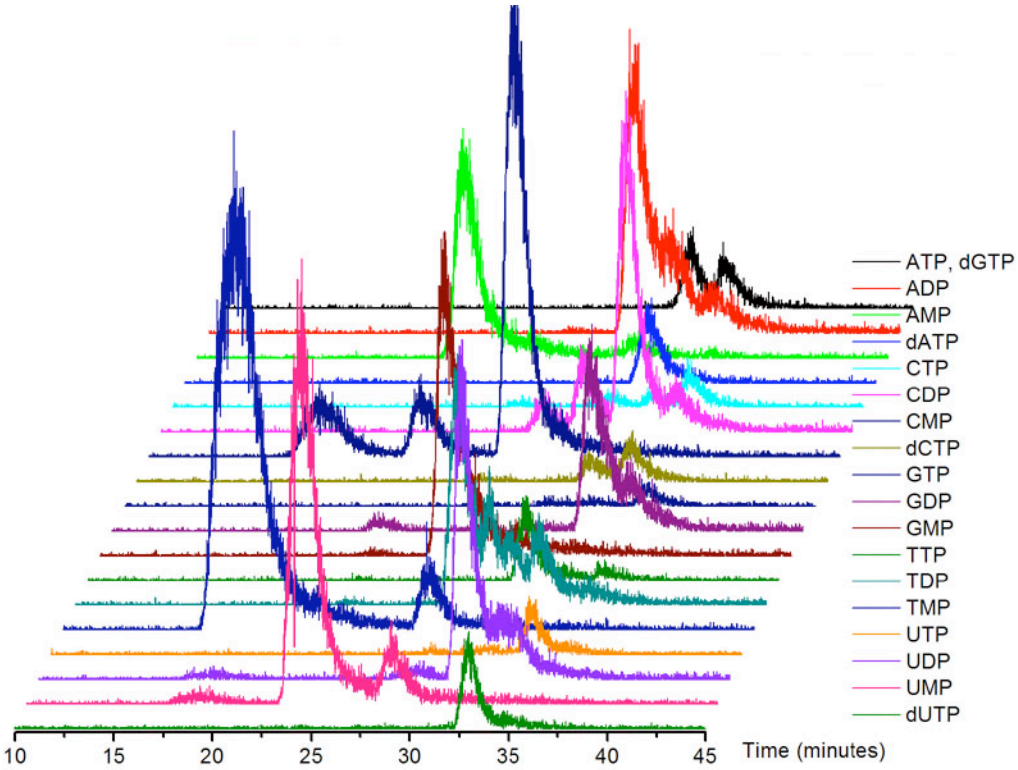


Figure 2. Separation of 19 nucleotides with an amide-80 column. The chromatographic conditions are: 0-5 minutes: 20% buffer, 5-30 minutes: 20-40% buffer; 30-45 minutes: 40% buffer. The buffer is NH4OAc, 5mM, pH 4.95. The organic solvent is acetonitrile. All nucleotides are detected using a QTrap with negative MRM scan mode.

Future work

- Optimization of the normal-phase HPLC separation methods will be performed and quantification of the nucleotides using this method will be conducted.

- A method to efficiently extract metabolites from *D. vulgaris* cell will be tested based on the current protocol for *E. coli* cells.

- Since ion-exchange HPLC is very efficient to separate a variety of metabolites, ion-exchange HPLC coupled with ion-suppressors will be developed to separate nucleotides and glycolysis pathway intermediates.

- A CE-MS has been purchase to further support the metabolomics efforts

- Test the new strategy developed for generation of knockout strains.

- Repeat the growth curve with the addition of the protein assay.

- Complete the first collaborative oxygen stress response ICAT experiment.

IV. Computational Core

LBNL – Arkin

In May, we completed preliminary cis-regulatory element prediction for all three of our target organisms: D. vulgaris, G. metallireducens, and S. oneidensis. These predictions come from a single algorithm based on over-represented pairs of motifs (dimers) found upstream of predicted operons. These predictions represent a first pass toward describing regulatory motifs in these three organisms, and will be followed shortly with results from several complementary methods. In addition, we plan to improve the accuracy of these predictions by adapting them to use our own operon predictions.

For our prokaryotic gene expression database, we made progress parsing results from a large number of microarray data formats into acommon format for use in comparative studies. We are still waiting for responses to our requests for data from a number of laboratories, and plan to send a second round of request letters in the near future. As we complete the initial data entry, we intend to begin analysis of the data sets we have received thus far.

We have completed a working prototype of the graphical engine for our comparative genome browser. It currently allows for two distinct modes of browsing: anchored mode (anchored to a single gene in each genome), and set mode (which displays a common set of genes in several organisms, regardless of genomic position). In the short term, we will integrate the browser into the VIMSS Comparative Microbial Genomics website and allow the rendered genes to link to information pages for each gene including data on that gene from the Functional Genomics Core.

Finally, we have started writing two manuscripts related to our operon prediction algorithm. One paperdescribes the method, and a separate paper investigates the evolution of operons. We hope to submit both of these papers over the next few months.

LBNL – Olken

Biopathways Graph Data Manager

Data Modeling: Contexts

We continue to work on the specification of context annotations and queries. Bioprocess contexts (BCs) include information on when and where a biological process occurs (organism anatomic location, genotype, developmental stage (or cell cycle phase), or environmental conditions). We believe that such context information will be of increasing importance as biopathways database become larger and more detailed. In particular, environmental context annotations will be increasingly important to VIMSS - as specifications of the stressors for microbial organisms.

Contexts would be specified as a compound (disjunctive normal form) Boolean predicate. Elementary predicates must include range (inequality) predicates (e.g., for pH, salinity). We envision encoding such contexts using OWL (an extension of RDF) since this appear to be the most popular knowledge representation languages in use in the bioinformatics community.

Query contexts (QCs) would similarly be specified as Boolean predicates. We anticipate that the most common query would return those bioprocesses which MAY satisfy the query context, i.e., the intersection of the query context with the reaction context is not empty. There will be two common cases: either the QC subsumes the BC, or the QC and BC have overlapping ranges of environmental variables. The former case arises when the bioprocess context is more detailed that the query context, e.g., the query context may be anaerobic conditions, while the bioprocess context might specify anaerobic conditions and marine (salt water) environment.

It would be unrealistic to require users to specify bioprocess contexts for each bioprocess individually. Presumably users will want to specify contexts for collections of bioprocesses (pathways, metabolism of an entire organism, or some taxa of organisms). Thus we will require some sort of inheritance mechanism to propagate such context constraints down to individual bioprocesses. We envision that successively narrowly scoped contexts are to be ANDed together. However, it is not yet clear that this will be sufficient, .e.g., to handle exceptional contexts in phylogenies due to loss of genes or horizontal gene transfer. The details of such an inheritance mechanism remain to be designed. Supporting such inheritance of contexts at run time would make query answering more complex.

We have not addressed query processing techniques for context constraining queries.

Data Modeling: RDF as a graph data model

We continue to investigate the possible use of RDF(Resource Description Framework) as our graph data model. RDF supports labeled directed graphs. We are leaning strongly toward this approach - at least initially. The attraction of this approach is the existence of a research and development community using RDF, as a potential source of tools for editing, drawing, and simple graph queries.

Implementation Planning

We continue to work on implementation planning. Our attention has been increasingly focused on a "data-centric" strategy, which gives primacy to issues of data integration of existing data sets from the BIOSPICE project. A graph navigational API would be provided (in Java) to support the development (or migration) of various applications (e.g., user interface codes, etc.).

We do not envision persistent replication of BIODB. Rather we would construct (and then materialize) a graph-based view of the relational BIODB. A major issue is how to construct such a view. We anticipate mapping each n-ary relation to nodes with n binary relations (graph edges). A automated schema conversion tool and a graph view materialization tool (to create a main memory graph representation of some subset of the database) would be created.

The Java-based graph navigation API would insulate application programmers from the details of the data exchange formats and related parsing issues. This will facilitate adoption of our graph data manager software. In the longer term we will provide a query-based API, however, the navigational API could continue to be used to traverse query results.

At Arie Shoshani's suggestion we are investigating development of a graph-based query tool as part of the BGDM project - to be funded via Sandia GTL funding. We conducted some very

early design studies of such a graph-based query language and are searching for a suitable graph editor / layout toolkit.

Teaching

On May 16, 2003, F. Olken lectured on Biopathways Databases for the Program on Genomic Applications Short Course on Comparative Genomics Tools, held at LBNL and organized by Sylvia J. Spengler and Manfred Zorn.

Collaborations

Jean Loup Faulon, a chemical graph theorist at Sandia Livermore, visited us and discussed his work on the use degree sequences to characterize graph structures. We had a freindly and fruitful discussion. We are hopeful that we may be able to incorporate some of his graph characterization codes into our graph DM.

Publications

The workshop report summary and whitepapers from the NLM workshop we co-organized with H.V. Jagadish (U. Mich.) on data management for biology has appeared as the July 2003 issue of the journal OMICS.

Future Work

- Elaborate design of context annotations and queries for biopathways (sub)graphs.
- Commence design of visual graph-based query language.
- Complete workshop report from NLM Workshop on Data Management for Life Sciences Research.

V. Project Management

Project Schedule

The GTL project schedule is undergoing updates and will be posted to the VIMSS Discussion Board on a regular basis. Any updates/comments/revisions should be sent to Nancy Slater via email (naslater@lbl.gov).

GTL Project Meetings

- The next GTL Executive Committee Meeting will be held on July 17, 2003.

- The 2003 VIMSS GTL Annual Retreat will be held on August 1-3 at the Hotel Nikko in San Francisco. Friday's agenda includes a poster presentation by each of the GTL project team members. The PIs will present on Saturday. See the Retreat website for the registration form and more details (http://vimss.lbl.gov/gtl_ann_retr.html).