Q-342: Towards High-Throughput and High-Sensitivity Approaches for Uncovering Total Environmental Gene Expression Patterns

Christopher W. Schadt [1,3], Zamin Yang [1,3], Amudhan Venkateswaran [1,3], Terry Hazen [2,3] and Martin Keller[1,3]

[1]*Biosciences Division, Oak Ridge National Laboratory,* [2] *Earth Sciences Division, Lawrence Berkeley National Laboratory,* [3]*Virtual Institute for Microbial Stress and Survival*

http://vimss.lbl.gov/

INTRODUCTION

Tremendous progress has been made in understanding microbial communities due to emergence of newly developed genomics-based technologies. Current technologies that have been applied to environmental samples for RNA transcriptional profiling include RT-PCR and functional gene microarrays using total RNAs. While these methods have provided considerable insights, they bear significant limitations that prevent their application in a high throughput manner to *de novo* communities. Both methods require background genomic information to allow for design of specific primers and/or microarray probes. Consequently, these methods can only reveal transcription activity of targeted conserved genes that are first surveyed by PCR and sequencing methods, or those that are obtained through comprehensive metagenomic shotgun sequencing. Thus each of these methods have high upfront costs in time, effort and materials. To circumvent this limitation, we are developing a method involving direct sequencing cDNA from the environment samples utilizing a high throughput sequence analysis system such as Bio454. These types of tools will be especially useful in understanding the basis for microbial survival under extreme environmental stressors which is a primary goal of the VIMSS-ESPP project.

GOALS & APPROACH

Our primary goal is to fully develop and implement methods that will allow for high-sensitivity, high-throughput and high-content analysis of global gene expression patterns from total environmental microbial communities.

To achieve this goal we are:

- Optimizing methods for efficient removal of ribosomal and other structural RNAs, thus enriching mRNA's for cDNA shotgun sequencing (Completed).
- Developing and Testing methods for unbiased amplification of cDNA from low biomass / low activity environments (in progress).
- Conducting a thorough evaluation of HT sequencing techniques in comparison to established microarray methods (in progress).
- Developing bioinformatics based methods similar to proteomics for analyzing cDNA sequence fragments derived from shotgun sequencing
- Conducting systematic comparisons of the developed techniques in single species laboratory cultures, mixed species cultures, and simple environmental communities.

COMPARISON OF THREE DIFFERENT METHODS FOR mRNA ENRICHMENT

Since upwards of 80% of total RNA is rRNA, strategies for rRNA removal and mRNA enrichment are a critical first step for implementing high throughput sequencing methods for environmental gene expression profiling. Using 10ug of total RNA harvested from a common pool of *Desulfovibrio vulgaris* cells as starting material, we tested three different strategies for enrichment of mRNA and examined their effects on microarray-based gene expression profiles.

Subtractive Hybridization: This method utilizes biotin modified oligos complementary to conserved regions in 16S & 23S rRNA and subtractive hybridization with streptavidin-coated magnetic beads.

Selective Reverse Transcription: This strategy uses two rounds of reverse transcription. rRNAs are first reverse transcribed with a set of universal primers; RNA/DNA hybrids and cDNA are removed by digestion with RNaseH & DNaseI; then enriched mRNAs are reverse transcribed using random primers

Exonuclease Digestion: This method uses a newly available exonulclease that specifically digests rRNAs that bear a 5'monophosphate group (intact, structural RNAs)

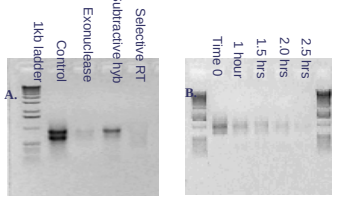

Figure 1. Methods for rRNA removal from total RNA. A) *D. vulgaris* RNA treated with each method and B) Effect of treatment time on exonuclease digestion of 5ug of total RNA

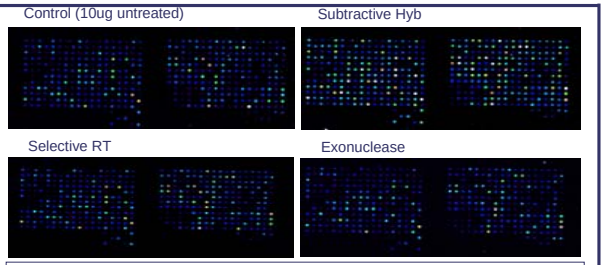

Figure 2. Microarray hybridizations of enriched *D. vulgaris* mRNA derived from each method. Each of methods give comparable results to 10ug of unenriched starting material using our standard labeling and hybridization procedures.

Figure 3. Normalized Signal Intensity data derived from three replicate microarray hybridizations of enriched *D. vulgaris* mRNA derived from each method (Cy5 – X axis) versus unenriched control samples (Cy3 – Y axis). Each of the tested methods gave fairly comparable results to the unenriched starting material. Selective RT and exonuclease digestion results exhibited considerably less variation between enriched and unenriched and across replicates.

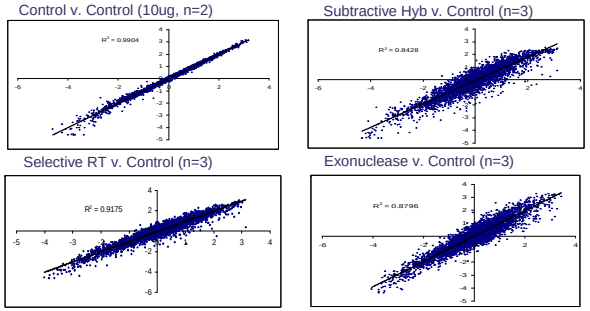

Table 1. Comparison of three methods for rRNA removal from total RNA and their effects on global gene expression patterns in *D. vulgaris.*

	Control	Subtractive Hyb	Selective RT	Exonuclease
Average # genes detected	3106 ± 77	3263 ± 7	3295 ± 18	2943 ± 183
Average Signal Intensity Difference (vs. control)	-137	+3395	-540	+399
Total # gene > 2 fold Increase	5	1797	19	238
Total # gene > 2 fold Decrease	5	45	184	173

SUMMARY OF RESULTS

- Each of the three methods was able to efficiently remove rRNAs, thus significantly enriching the cDNAs
- We were able to successfully apply each method towards microarray hybridizations, reducing the total hybridized amounts about ~5 fold, without significant compromising the quality of the array results.
- Both the subtractive hybridization and exonuclease based methods increased the signal intensity of microarray hybridizations given the same amount of total RNA starting material. No systematic biases were observed in any of the methods.

CONCLUSIONS & NEXT STEPS

- In collaboration with the JGI we are currently repeating these experiments and preparing cDNA for sequencing on the Bio454.
- We are evaluating using each of these methods in combination with new and existing (Gao et al., AEM, 2007) approaches for RNA amplification
- Strategies for rapid *in silico* analysis & quantification and Bio454 data methods for comparison to microarray data will be developed.
- We will further evaluate and apply the developed technique using dual culture (e.g. DvH and Methanococcus) and simple environmental communities.

ACKNOWLEDGEMENT

ESPP is part of the Virtual Institute for Microbial Stress and Survival supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics:GTL Program through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.