

## Computational Analysis of the Redox stress response in organisms

Thanura Elvitigala<sup>1,\*</sup>, Abhay Singh<sup>2</sup>, Abha Khandelwal<sup>2</sup>, Maitrayee Bhattacharya-Pakrasi<sup>2</sup>,  
Rajeev Aurora<sup>3</sup>, Himadri Pakrasi<sup>2,4</sup>, Ralph Quatrano<sup>2</sup> and Bijoy K. Ghosh<sup>5</sup>

<sup>1</sup>Department of Electrical & Systems Engineering, Washington University, St. Louis, MO 63130

<sup>2</sup>Department of Biology, Washington University, St. Louis, MO 63130

<sup>3</sup>School of Medicine, Saint Louis University, St. Louis, MO 63104

<sup>4</sup>Department of Energy, Environmental & Chemical Engineering, Washington University,  
St. Louis, MO 63130

<sup>5</sup>Department of Mathematics and Statistics, Texas Tech University, Broadway and Boston,  
Lubbock, TX 79409-1042

\*tre1@ese.wustl.edu

Maintaining Redox homeostasis inside the cells is vital for the survival of any living organisms. Living organisms respond to redox stresses caused by various external conditions by adjusting their metabolic processes. Time course DNA-microarray experiments are considered as one of the best ways to gain an insight on this complex biological phenomenon. In order to extract the hidden information contained in large amounts of raw data generated by the Microarray experiments, a systematic application of various computational techniques is considered essential.

### Introduction to Microarray Experiments and Data Analysis

Two different organisms namely *Synechocystis* 6803, a unicellular cyanobacteria and *Arabidopsis thaliana* a flowering plant, were subjected to number of stress conditions including light quantity (HL), DCMU and light quality (Excitation using Blue or Red light-BR) known to disturb the redox homeostasis of the cells. Samples obtained at different time points over the experiments were used to generate Microarray data. Though Microarrays are considered as a fairly established technique for obtaining transcriptome level information of a cell, data analysis still remains to be a challenging task. We will be discussing different techniques used in various steps including quality assessment, data normalization, clustering, co-expression network generation and pathway analysis.

### Response of Individual Organisms to Redox Stresses

Analysis of Microarray data from three experiments on *Synechocystis* 6803 confirmed that large number of gene were under transcriptional regulation under these stresses. Around 23%, 35%

and 50% of total genes were identified as differentially regulated in HL, BR and DCMU experiments respectively while a common set of 11 % of the genes were differentially expressed in all three experiments. These genes belong to 60% of metabolic pathways, indicating wide range of responses across the cell. Mostly affected pathways include carbon fixation, oxidative phosphorylation, photosynthesis, and ribosomal proteins.

CyanoBase Pathway ID	Total Genes in the pathway	Differentially Expressed in			
		All	HL	DCMU	BR
ATP synthase	10	80%	90%	90%	90%
Chaperones	16	38%	44%	69%	56%
CO2 fixation	15	40%	53%	73%	53%
Membranes, lipoproteins, and porins	12	33%	67%	50%	83%
NADH dehydrogenase	23	30%	43%	70%	78%
Nucleoproteins	7	43%	71%	71%	71%
Phosphorus compounds	3	0%	67%	0%	100%
Photosystem I	16	63%	81%	75%	88%
Photosystem II	27	48%	63%	78%	74%
Phycobilisome	18	67%	78%	72%	100%
Pyruvate dehydrogenase	4	0%	75%	0%	75%
Ribosomal proteins	63	60%	73%	81%	76%
RNA synthesis, modification, and DNA transcription	23	22%	35%	78%	61%

Table1: Summary of the significantly affected pathways in *Synechocystis* 6803 in different experiments. Pathways are based on CyanoBase categorization.

In the case of *Arabidopsis* 20% and 8% of total genes were differentially regulated in the HL and DCMU treatments respectively while around 5% of the genes were differentially expressed in both experiments. Here also genes represented around 65% of the metabolic pathways present in *Arabidopsis* indicating that the redox response is system wide in the higher plants also.

## Extracting Biological Significance of the Microarray Data

Clustering techniques and Co-expression networks were used to group genes based on their behavioural patterns. Over represented pathways in different clusters were identified and used to derive general conclusions on cell response to redox stress conditions. Expressions patterns of different genes were associated with the corresponding metabolic pathways in order to understand the time course response of individual pathways. In addition to identifying system wide response of individual organisms, comparison between two organisms was also done highlighting similarities and differences in their responses. For example, Alanine, Aspartate, Arginine metabolisms are significantly affected in both organisms while Tryptophan and Tyrosine metabolisms are only affected in *Arabidopsis*.

KEGG Pathway ID	Differentially Expressed in Arabidopsis	Differentially Expressed in Synechocystis
Alanine and aspartate metabolism	42%	50%
Arginine and proline metabolism	42%	46%
Carbon fixation	37%	50%
Glutamate metabolism	40%	37%
Glutathione metabolism	43%	18%
Oxidative phosphorylation	47%	6%
Photosynthesis	24%	73%
Ribosome	52%	85%
Sulfur metabolism	29%	14%
Tryptophan metabolism	52%	12%
Tyrosine metabolism	44%	9%

Table 2: Some of the significantly affected pathways in Highlight experiments based on KEGG. Some are significant only in one organism while others are significantly affected in both.

This work was supported by funding from the NSF-FIBR program (EF0425749)