Kenyon College
Digital Kenyon: Research, Scholarship, and Creative Exchange

Kenyon Summer Science Scholars Program | Summer Student Research Scholarship

Summer 2004

Microarray Analysis of pH-Dependent Gene Expression in Anaerobically Grown E. coli-K12

Everett Hayes

Follow this and additional works at: https://digital.kenyon.edu/summerscienceprogram

 Part of the Biology Commons

Recommended Citation

Hayes, Everett, "Microarray Analysis of pH-Dependent Gene Expression in Anaerobically Grown E. coli-K12" (2004). *Kenyon Summer Science Scholars Program.* Paper 303.
https://digital.kenyon.edu/summerscienceprogram/303

This Poster is brought to you for free and open access by the Summer Student Research Scholarship at Digital Kenyon: Research, Scholarship, and Creative Exchange. It has been accepted for inclusion in Kenyon Summer Science Scholars Program by an authorized administrator of Digital Kenyon: Research, Scholarship, and Creative Exchange. For more information, please contact noltj@kenyon.edu.

Microarray Analysis of pH-Dependent Gene Expression in Anaerobically Grown *Escherichia coli* - K12

Everett Hayes '05

With E. Yohannes, S. BonDurant, M. Radmacher, and J. L. Slonczewski*, Kenyon College, Gambier OH

Abstract

Microarray techniques were used to look at the effect of pH on *Escherichia coli* strain W3110. Cultures were grown anaerobically under different pH conditions (pH 5.7, 7.0, and 8.5) on LBK medium, buffered with Homopipes. RNA was isolated from five replicate cultures for each condition and purified using an RNeasy kit. cDNA was synthesized and fragmented at the University of Madison-Wisconsin and then hybridized to microarray chips. The arrays were analyzed using Dchip and BRB array tools. 2002 genes were found to be differentially expressed using an ANOVA test. A cluster analysis revealed six clusters of genes based on their level of expression under each of the three pH conditions. A number of interesting genes were identified, including the gadA and gadB genes, which are known to be involved in acid resistance as well as many flagellar genes such as flhC and flhD.

Introduction

·What is *Escherichia coli*?

- ·*Escherichia coli* is a bacterium in the Enteric family
- ·They are found in a number of environments, including being part of the natural flora of the human digestive tract
- ·While many strains of *E. coli* are harmless, some strains of *E. coli* can cause serious infections such as *E. coli* O157:H7, which is associated with food-born illness (Russell et al. 1999)

·What am I studying?

- ·*E. coli* can survive a number of different conditions met in the human digestive tract
- ·Resistance to changes in pH plays an important role in pathogenicity
- ·*E. coli* are facultive anaerobes, meaning they can survive in limited oxygen environments that are met in the digestive system
- ·It is important to look at how anaerobiosis affects resistance to different pH conditions
- ·*E. coli* have a number of known acid and base resistance systems
- ·The role of different genes on acid resistance has been widely studied (Slonczewski et al. 2002; Maurer et al 2004; Slonczewski et al. 1996)
- ·Microarrays allow for exploration of the entire *E. coli* genome
- ·Comparison of all the genes will allow for a genome-wide perspective of how genes operate under different pH conditions

·Why are *E. coli* are a good model organism?

- ·Are non-pathogenic in wild-type strains such as W3110
- ·*E. coli* is closely related to other pathogenic bacteria in the Enteric family (Shigella, Salmonella, etc.)
- ·*E. coli* is widely studied and there is a wealth of information about the organism
- ·The entire *E. coli* genome has been sequenced, allowing construction of microarrays for all open reading frames and many intergenic regions

Methods

Growth

·*E. coli* overnight cultures were grown anaerobically at 37°C
·The overnight was diluted 1:1000
·LBK media buffered with Homopipes (pH 5.7, pH 7.0, pH 8.5) was used
·Eight replicates were made for each condition (24 samples total)
·Cells were grown to log phase growth (~O.D. = 0.2 at 600nm)
·To extract RNA, the sample was poured into a 50 mL conical vial with 2 volumes of RNA protect (16 mL)
·Samples were mixed by vortexing for 5 seconds and then they sat for 5 minutes at room temperature
·Samples were then centrifuged for 10 minutes at room temperature (>5000 x g).
·Supernatant was removed and then the samples were centrifuged for another 2 minutes at room temperature (>5000 x g).
·Remaining supernatant was removed and the vials were placed upside down on a paper towel for 10 seconds to dry.
·Pelleted RNA was then stored at -80°C.

RNA isolation and purification

RNA isolation was performed using a Qiagen RNeasy kit which consisted of a series of specific washing and rinsing steps using the RNeasy column. After the wash steps, the column was spun down and the supernatant was discarded.

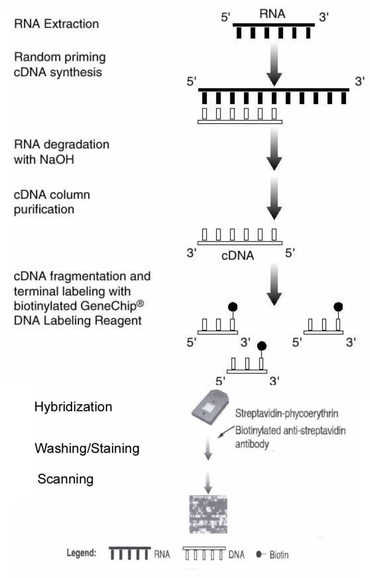
Quality checks

·A 1 μg sample was run on a 1% agarose gel
·RNA samples were also quantified using a spectrophotometer by comparing A_{260} an A_{280} absorbances.

Storage

For storage, 2 volumes of 100% ethanol an 0.1 volume of 3M NaOAcetate were added to the samples and they were stored at -20 °C for short term storage, or -80 °C for long term storage. Samples were shipped on dry ice to the University of Wisconsin-Madison for further analysis.

cDNA synthesis and chip hybridization



Analysis

·Gene chips were analyzed using Dchip and BRB array tools
·Dchip was used to normalize the data and to perform model based expression analysis
·ANOVA analysis was performed on the log2 model based expression indices to determine significance among conditions
·An automated Tukey test was used to do a pairwise comparison on the significant genes
·In BRB array tools, a Multidimensional scaling analysis was performed to compare the different pH classes
·A cluster analysis of all of the significant genes was also made in BRB array tools

Results

Based on the ANOVA test, 2002 genes were considered to be differentially expressed (with 95% confidence). 609 of these genes were from intergenic regions of the genome. The genes were grouped into six distinct clusters bases on the cluster analysis (Figure 1, Figure 2). Cluster 1 contained 266 genes and contained genes with high acid expression and low base expression. Cluster 2 contained 509 genes which had a higher response in acid than in base or neutral conditions. In cluster 3 there were 227 genes with highest expression in neutral conditions. An acid-low trend (lowest expression in acid) was seen in cluster 4, which contained 635 genes. Cluster 5 had 60 genes with a neutral-low response and cluster 6 had 305 genes that had the highest response in base. The Multidimensional scaling, using centered correlation showed distinct grouping of the three classes in separate groups (Figure 3).

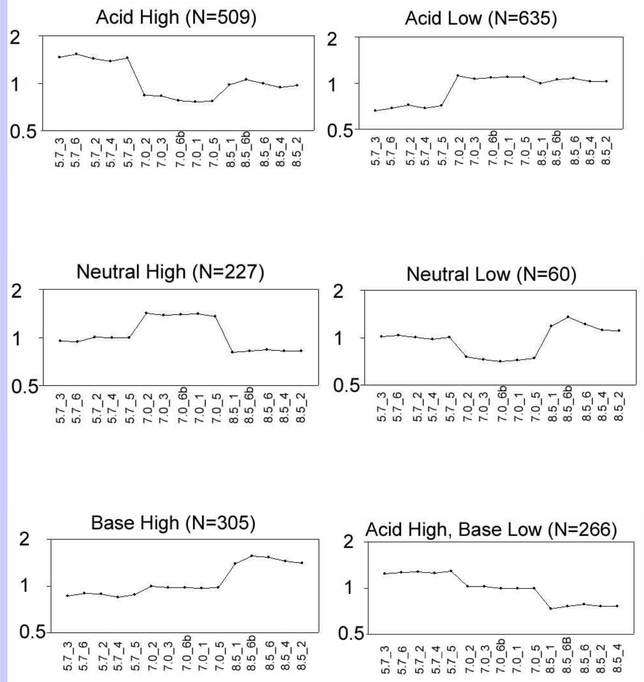


Figure 2. Line plot of the median intensities of each cluster. Each point represents the median intensity of each replicate for the genes in a particular cluster.

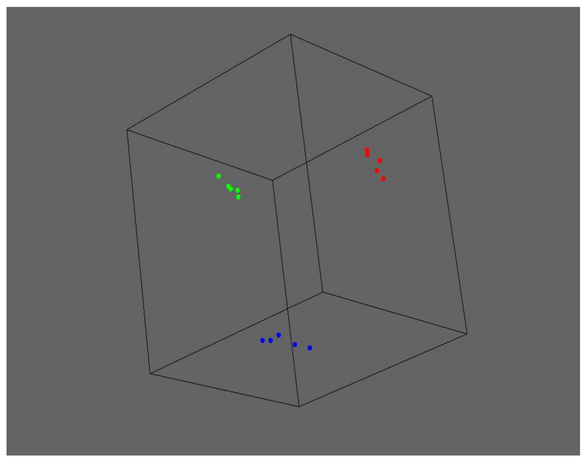


Figure 3. Multidimensional Scaling plot. Distance between samples is shown in 3-dimensional space (red = acid, blue = base, green = neutral pH)

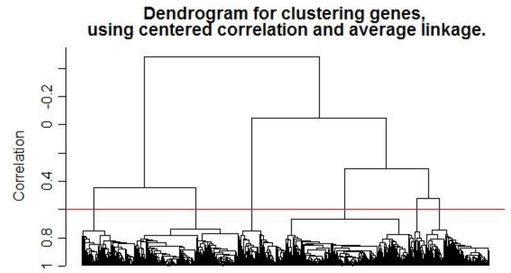


Figure 1. Cluster analysis of the genes in the *E. coli* genome. The clusters are cut at the line shown, giving six distinct clusters (numbered 1-6 from left to right).

Discussion

This project identifies patterns of pH-dependent gene expression in the *E. coli* genome that require or are enhanced by anaerobic conditions. A number of known acid resistance genes were shown to be significantly expressed such as gadA and gadB, as well as a number of genes of unknown function such as many intergenic genes, yabN, yciC, and yieP. Whether or not these unknown genes play a role in acid or base resistance, or if the differences are caused by a reaction to the environmental conditions is not yet known. Many genes that showed significant pH dependence also have known functions in the cell other than pH resistance. Whether or not these genes also contribute to acid or base resistance is an interesting avenue for further studies. Other work could be done comparing how lab strains such as W3110 used in this experiment compare to clinical isolates of *E. coli* in their gene regulation in response to pH. A further comparison of data with that of Maurer et al (2004) may shed light on how anaerobiosis affects pH regulation in *E. coli*. It is the nature of microarray analysis to generate a lot of information that can then be directed towards more hypothesis-driven research. Some genes of interest that were found to have significant differential expression in this experiment include the flhC and flhD genes. These genes are of particular interest because they are part of the flagellar operon and are also known to be involved in anaerobic respiration and the Entner-Doudoroff pathway (Prub et al. 2003).

References

Hersh, B. M., F. T. Farooq, D. N. Barstad, D. L. Blankenhorn, J. L. Slonczewski. 1996. A glutamate-dependent acid resistance gene in *Escherichia coli*. J. Bacteriology. 178: 3978-3981.

Maurer, L., E. Yohannes, S. S. Bondurant, M. Radmacher, J. L. Slonczewski. 2004. pH Regulates Genes for Flagellar Motility, Catabolism, and Oxidative Stress in *Escherichia coli* K-12. J. Bacteriology. In press.

Prub, B. M., J. W. Campbell, T. K. Van Dyk, C. Zhu, Y. Kogan, P. Matsumura. 2003. FlhD/FlhC Is a Regulator of Anaerobic Respiration and the Entner-Doudoroff Pathway through Induction of the Methyl-Accepting Chemotaxis Protein Aer. J. Bacteriology. 185: 534-543.

Russell, J. B., F. Diez-Gonzalez, G. N. Jarvis. 1999. Effects of Diet Shifts on *Escherichia coli* in Cattle. J Dairy Sci. 83: 863-873.

Stancik, L. M., D. M. Stancik, B. Schmidt, D. M. Barnhart, Y. N. Yoncheva, J. L. Slonczewski. 2002. pH-Dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia coli*. J. Bacteriology. 184: 4246-4258.

Acknowledgements

This Project was funded by the Kenyon College Summer Science Program and the Howard Hughes Grant. I would like to thank Joan Slonczewski, Elizabeth Yohannes, Sandra BonDurant (at UW Madison), Michael Radmacher (at OSU), and Wade Powell for all their help. The cDNA synthesis is from Affymetrix protocols that can be found at www.affymetrix.com.