

## **Integrative study of *H.influenzae*- Host interactions (Project III & Core B2)**

**Principal Investigator: Munson, Robert**

### **Overview**

In aim 1 of project 3, we proposed to describe, at the transcriptome level, the role of cAMP and the cAMP binding protein (CRP) in *H. influenzae* gene regulation. The first goal was the construction and characterization of a strain with a mutation in the adenylate cyclase gene (*cya*). Previously, Redfield and coworkers characterized a *cya* mutant constructed by transposon mutagenesis. Their mutant was characterized with respect to the role of cAMP in carbohydrate transport and competence for transformation. The data clearly indicate that there are cAMP-dependent and cAMP-independent responses. We have constructed a new strain and verified by Southern blot, that it contains a nonpolar mutation in the *cya* gene. This mutant has been used in a number of microarray experiments, several of which are described below.

With respect to Core B, we proposed to use the Operon 70-mer ORF set for *H. influenzae* Rd. This is a unique set of 70-mer oligos, one for each gene. The Operon set became available in early 2003. We have performed a number of experiments using strain Rd and the *cya* mutant to optimize our microarray conditions. We have optimized our RNA isolation procedure, determined that it is not necessary to remove the rRNAs when working with the Operon oligo set and have increased the reproducibility of our analyses by optimizing our hybridizations on a Genomic Systems GENETAC Hybridization Station. Thus, the short-term goals of refining our methodologies to produce the best and most consistent RNA possible and then defining conditions that will minimize the slide-to-slide (or even within a slide) variation in hybridization have largely been completed. We are currently working to optimize data normalization protocols.

Thus, we have demonstrated that we can construct *H. influenzae* mutants, perform transcriptional analysis employing slide-based microarrays and are therefore well positioned to begin investigation of the transcriptome of *H. influenzae*.

### **Current Results**

Competence, the ability to take up DNA from the media and recombine a portion of that DNA into the chromosome by homologous recombination, has been studied extensively in *H. influenzae*. Competence occurs spontaneously near the end of the exponential phase of growth or can be induced by a number of laboratory procedures. Commonly, exponentially growing *H. influenzae* cells are shifted to a nutrient poor growth medium, designated MIV. After incubation for 60-100 minutes in MIV, the majority of cells are competent for transformation. Protein synthesis continues after transfer of *H. influenzae* to MIV but the cells are unable to divide. cAMP concentration is increased during competence development and it is known that *cya* mutants do not become competent or induce the expression of some genes involved in competence development. For example, it is known that the *com* genes are up-regulated during competence development.

We have completed a number of experiments comparing the transcriptome of competent *H. influenzae* to *H. influenzae* cells growing exponentially in supplemented brain heart infusion medium. RNA was isolated from competent cells prepared by incubation in MIV medium and from exponentially growing cells. Microarray analysis was performed as described below. In Fig. 1, left panel, one block of spots for a typical slide is shown. The spots are of high quality with little background. Data were acquired with GenePix software, and uploaded into the Gene Traffic. The data from one slide are represented in the plot of M vs. A (Fig. 1, right panel). The spots corresponding to the genes that were most increased with respect to transcript levels are those in the

upper quadrant of the figure. A subset of these highly up-regulated genes is shown in Fig. 2 as well as a several other representative genes. The strongly up-regulated genes include genes whose products are known to be important for competence development as well as genes that have not previously been considered with respect to competence development. It is important to note that this experimental protocol cannot distinguish between genes that are related directly to competence development and genes that are up-regulated in expression simply due the shift to the nutrient poor medium.

Experiments were performed to define the cAMP regulon and determine whether some competence genes might be regulated independently of cAMP concentration. In these experiments, all strains were grown in supplemented brain heart infusion medium, then transferred to MIV medium for 100 minutes. MIV medium was supplemented with 1mM cAMP where indicated. In Fig. 3, data are presented for 3 comparisons. First, we compared the partial transcriptome of competent strain Rd cells with *cya* mutant cells treated under the same conditions. Many of the genes in strain Rd that were up-regulated during competence development (see Fig. 2) are also up-regulated when expression in strain Rd was compared to the expression in the *cya* mutant (Fig. 3, blue bars). Up-regulation of many of these genes is also observed when transcripts from mutant cells incubated with 1 mM cAMP are compared to transcripts from mutant cells that were not treated with cAMP indicating that 1mM cAMP can reverse the effects of the *cya* mutation for this set of genes (Fig. 3, grey bars). Another comparison is consistent with this conclusion. Expression ratios for many of these same genes are similar when transcripts from competent strain Rd cells (Fig. 3, blue bars) are compared to the transcripts from the *cya* mutant treated with cAMP (Fig. 3, red bars). Purine starvation is also thought to be important in competence development. It is interesting to note that expression ratios of the subset of genes including *purD* and *purM* and *purN* were not dramatically different in the *cya* mutant compared to strain Rd under any of the tested conditions indicating that expression of these genes is not dependent on cAMP levels.

Several genes that were not up-regulated during competence development in strain Rd including HI1677, HI0017, *glpQ* and HI0105 (Fig. 2) are relatively down-regulated in expression in strain Rd when compared to expression in the *cya* mutant (Fig. 3, blue bars). In contrast, when transcripts of these genes in the *cya* mutant treated with cAMP are compared to the transcripts from the *cya* mutant, there is an increased level of relative expression in the cAMP-treated cells, indicating a clear effect of cAMP and an additional yet to be defined level of regulation.

In summary, we have made considerable progress in establishing the 70-mer slide based arrays in our facility. We have demonstrated up-regulation of known competence genes during the development of competence and identified new genes that were not previously associated with competence development. We have also demonstrated that up-regulation of competence genes is partially dependent on cAMP levels. These data confirm and extend data in the literature based largely on *lacZ* transcriptional fusions and will provide significant new information that is more global in nature. These experiments will all be repeated, confidence levels determined, and slide-to-slide variation factored in. MIAME standards for microarray data will be met for all experiments.

## **Materials and Methods**

### Array description

The *H. influenzae Rd* array set was designed and synthesized by Operon Technologies (Qiagen). Each unique 70-mer oligos was designed using Operon's proprietary software using published open reading frame sequences for the *H. influenzae Rd* genome. Probe design criteria parameters were: Tm range 76+/- 5°C, stem length in potential hairpin <=8, cross-hybridization identity against all other genes <=70%, and poly(N) tract repeats <=7. Each oligo was resuspended to a final

concentration of 40  $\mu$ M in 3X SSC and arrayed in triplicate to UltraGAPSII slides (Corning, Corning, NY). After arraying, slides were re-hydrated for 10 sec over a 42°C water bath and then snap-dried on a 100°C hot plate (DNA slide up). Oligos were immobilized on slides using a UV crosslinker set at 600 mJoules.

#### Sample preparation and slide hybridizations

RNA was isolated from a 50 ml culture of actively growing cells in sBHI medium (log-phase) or cells shifted into MIV minimal medium using the Trizol reagent (Invitrogen, Rockville, MD) as described in the detailed materials and methods in the proposal. RNA pellets were resuspended in 400  $\mu$ l of RNase-free ddH<sub>2</sub>O, DNase treated, and then further purified using QIAquick spin columns (Qiagen). Typically, 100  $\mu$ g of total RNA was obtained from 50 ml of culture ( $A_{600} = 0.4$ ). RNA samples (20  $\mu$ g) were labeled using indirect-labeling method based on Brown Lab protocol (<http://cmgm.stanford.edu/pbrown/protocols/amino-allyl.htm>) with some modifications. Briefly, cDNA was synthesized in a 30 $\mu$ l reaction by combining: RNA annealed with 3  $\mu$ g of random-primers (Invitrogen, Rockville, MD), 300U SuperscriptII (Invitrogen, Rockville, MD), 10 mM DTT, 40U RNaseOUT (Invitrogen, Rockville, MD), 500  $\mu$ M dATP, 500  $\mu$ M dGTP, 500  $\mu$ M dCTP, 100  $\mu$ M dTTP, and 400  $\mu$ M AA-dUTP. After 1 hour at 42°C, 200U of SuperscriptII was added to the reactions and synthesis allowed to continue for 1 additional hour. RNA was removed from newly synthesized cDNA by alkaline hydrolysis. Five microliters of 2 N NaOH and 10  $\mu$ l of 500 mM EDTA were added to each reaction and the mixtures incubated at 70°C for 10 min. After hydrolysis, the mixture was neutralized with the addition of 10  $\mu$ l of 1 N HCl. cDNA was recovered and purified using a Microcon YM-30 concentrator (Millipore, Bedford, MA). The neutralized cDNA was combined with 500  $\mu$ l of nuclease-free ddH<sub>2</sub>O and applied to the concentrator. The concentrator was centrifuged at 11,000Xg for 10 min. The flow-thru was discarded and 500  $\mu$ l of nuclease-free ddH<sub>2</sub>O added to the filtrate cup and the concentrator centrifuged again for 11,000X g for 10 min. Flow-thru was discarded and the cDNA washed again with 500  $\mu$ l nuclease-free ddH<sub>2</sub>O as described above. cDNA was collected in a clean 1.7 ml centrifuge tube at 1000X g for 1 min. The filtrate cup was washed with two sequential 25  $\mu$ l aliquots of nuclease-free ddH<sub>2</sub>O for a combined volume of ~52  $\mu$ l. Typically, 10  $\mu$ g of cDNA per reaction was recovered.

One half of the cDNA from each reaction (equivalent to cDNA synthesized from 10  $\mu$ g of total RNA) was labeled with either Cy3 or Cy5 monofunctional dye (Amersham Pharmacia Biotech, Arlington Heights, IL). Labeled cDNAs were purified using QIAquick columns, and concentrated in a vacuum dessicator without heat. Labeled cDNAs were resuspended in 10 mM EDTA pH 8.0 (65 pmol of each in 14  $\mu$ l total), heated to 90°C for two minutes, and combined with 126  $\mu$ l of SlideHyb #1 hybridization buffer (Ambion). Cy3 and Cy5-labeled analytes were hybridized to microarray slides overnight using the GeneTac hybridization station. The automated hybridization protocol used included the following steps: 65°C for 2 hours, 60°C for 2 hours, 55°C for 2 hours, 48°C for 12 hours. Slides were machine washed in low stringency buffer (1.0X SSC, 0.2% SDS, 0.1 mM DTT; flow rate 10 sec; 32°C; hold 15 sec; 2 cycles), in medium stringency buffer (0.1X SSC, 0.2% SDS, 0.1 mM DTT; flow rate 10 sec; 20°C; hold 20 sec; 2 cycles), and in medium stringency buffer without SDS (0.1X SSC, 0.1 mM DTT; flow rate 10 sec; 20°C; hold 20 sec; 2 cycles). Slides were removed from the hybridization cassettes and washed manually for 1 min in medium stringency buffer without SDS. Slides were dried by centrifugation at 1000 rpm for 1 min.

### Hybridization signal acquisition and data filtering

Slides were scanned in an Affymetrix 427 scanner. Gain settings for the scanner were set so that > 95% of the spots on each array were below saturation level to yield maximum dynamic range within an experiment. GenePix software v.4.0 (Axon Instruments Inc.) was used to extract signal intensities (foreground and background) in both channels for each feature. GeneTraffic Duo (Iobion Informatics) analysis software was used to examine significance of gene expression. Spots were rejected as bad or missing by filtering any gene in which the signal was less than 1X local spot background intensity, less than 1X the average background intensity, and less than 100 fluorescent units in either Cy3 or Cy5 channel. We applied the Locally Weight Scatter Plot Smoother function (Lowess-global normalization) to all slide experiments.  $\log_2(\text{ratios})$  intensities were calculated based on the average of the (experimental/reference) intensity ratios of each qualified spot. Average mean ratios were calculated for each gene by averaging the three spot values ( $\log_2\text{ratio}$ ) within a slide experiment.

Figure 1. *H. influenzae* microarray data and Scatter plot

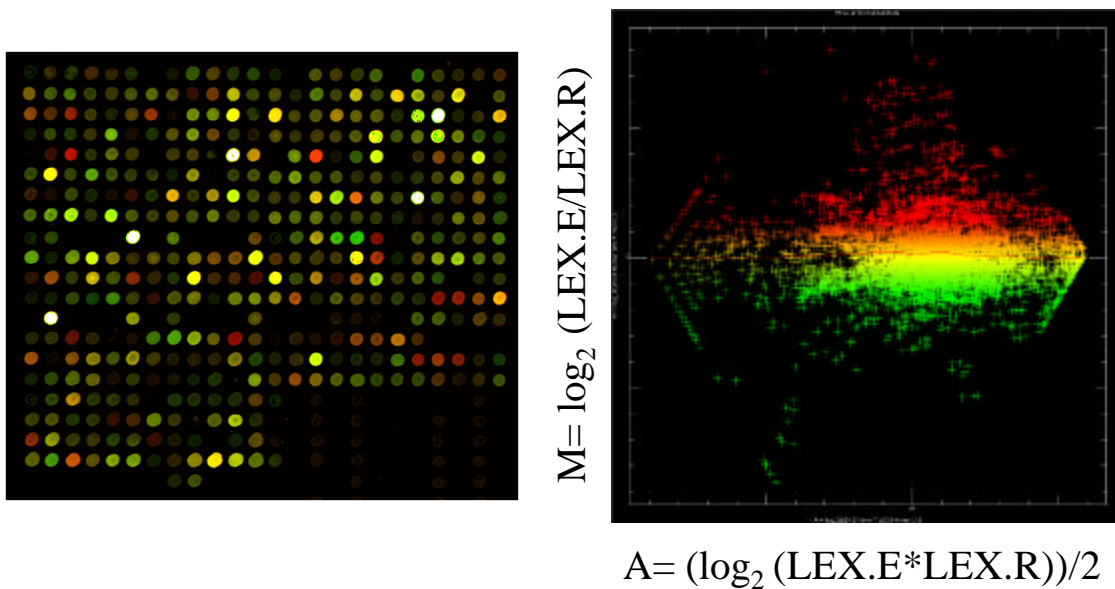


Figure 2. Strain Rd, competent vs. log phase cells

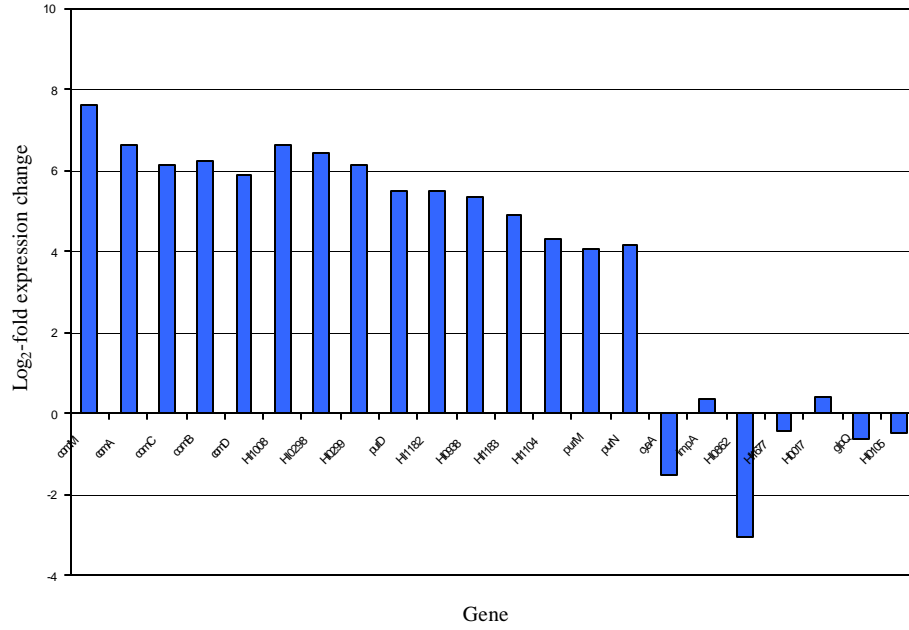


Figure 3. Comparative gene expression in *H. influenzae*

