MDEAT - a new databionic evaluation and analysis tool to identify the virulence regulon of *Listeria monocytogenes* as a model system

T. Hain¹, A. Ultsch², S. Chatterjee¹, A. Amend¹ and T. Chakraborty¹

¹Institut für Med. Mikrobiologie, Justus-Liebig Universität, Frankfurter Str. 107, 35392 Giessen, Germany; ²Fachbereich Mathematik und Informatik, Arbeitsgruppe Datenbionik, Philipps-Universität Marburg, Hans-Meerwein-Str., 35032 Marburg, Germany

**Background**

The pathogenic bacterium *Listeria monocytogenes* is able to cross the endothelial, placental and blood-brain barrier causing gastritis, meningitis, neonatal death and other diseases in humans. PrfA (Positive regulatory factor A) is the sole regulator of the virulence genes identified in *Listeria monocytogenes* till date. It regulates genes within the virulence gene cluster (vgc), which harbors *prfA* itself, *plcA*, *hly*, *mpl*, *actA* and *plcB*. PrfA is also known to regulate cell-wall associated internalins (*inlA* and *inlB*), secreted internalins (*inlC*) and activates the transcription of genes involved in hexose phosphate uptake and bile salt hydrolases. We tried to identify additional PrfA regulated genes using a novel bioinformatic approach.

**Bioinformatical approach**

Throughout the last several years, the microarray technology has become widespread within the scientific community. Microarrays allow the analysis of changes in the expression profiles of several thousands genes in a parallel approach. Nevertheless, it is still not clear what number of experiments are sufficient to generate significant results from a minimal set of data. To overcome this problem, we have used a new bioinformatics tool called MDEAT (microarray databionic evaluation and analysis tool), which permits qualitative analysis of the results of a microarray experiment. If the quality is sufficient, only a small number of experiments to reach statistical significance have to be performed. If the quality is insufficient, more experiments have to be performed. The databionic-based approach has its roots in a theoretical background for the well known Pareto-80/20-law. Precise estimations of the distribution of the over- or under-expression of genes could be calculated, called Pareto Density Estimation (PDE). A picture of an experiment’s PDE gives a direct visual feedback of the experiment’s quality. Statistical measures for the quality of an experiment can also be derived from this PDE.

**Microarray experiment**

To evaluate the databionic approach, we selected whole genome microarrays of *Listeria monocytogenes* to investigate the expression profiling of the virulence regulon as a well-studied model system. For this purpose, we compared microarray experimental data collected according to a direct Cy dye labelling protocol with data collected according to an indirect Cy dye post labelling protocol. These results were analyzed comparatively for *Listeria monocytogenes* (wt) and *Listeria monocytogenes* complemented with *prfA* (wt+) versus the *prfA* deletion mutant (ΔprfA).

**Results and Discussion**

Several new PrfA regulated genes were discovered in this work using MDEAT. We identified an endo-1,4-beta-xylanase, a putative peptidoglycan-linked protein with a LPXTG motif and a cytosine desaminase as being positively regulated by PrfA, whereas genes involved in glutamine and glycine/betaine uptake as well two lipoproteins appear to be negatively-regulated. We confirmed the presence of PrfA binding sites for all of the positively regulated candidate genes, but not for any of the negatively regulated genes. These results suggest a more global regulatory role for PrfA, which can act both as an activator and/or repressor of genes.