

**NCI Center for Bioinformatics**  
**Informatics Seminar Series**

**Integrating Gene Expression And Proteomic  
Data With Other Biological Data Types:**

**Human Disease Associations, Literature  
Mining, and Biological Pathways**

—  
**2:00 until 4:30 PM**  
**January 26, 2004**

**6116 Executive Blvd., Conf. Room 3056A/B**

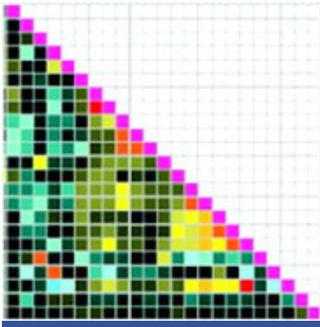
—  
**Kevin G. Becker**  
**Gene Expression and Genomics Unit**  
**National Institute on Aging, NIH**  
**Baltimore MD 21224**

High throughput gene expression and proteomic studies demand equally matched systems for biological analysis of large quantities of data. We have developed three publicly accessible web based tools, which are useful for the integration of high throughput gene expression and proteomic data. These include:

PubMatrix for high throughput semi-automated literature searching <http://pubmatrix.grc.nia.nih.gov/>; the Genetic Association Database for the integration of human genetic data <http://geneticassociationdb.nih.gov> and BBID-The Biological Biochemical Image Database <http://bbid.grc.nia.nih.gov/>

I will discuss the challenges in developing these tools; novel features which aid high throughput data analysis, and applications of these tools in microarray and proteomic data analysis.

Following the presentation there will be an open discussion.



# PubMatrix

[PubMatrix](#) | [Contact Us](#) | [Links](#) | [NIA DNA Array Unit](#)

PubMatrix is a simple way to rapidly and systematically compare any list of terms against any other list of terms in PubMed. It reports back the frequency of co-occurrence between all pairwise comparisons between the two lists as a matrix table. Lists of terms can be anything; gene names, diseases, gene functions, authors... pretty much anything. The user can then quickly sort or browse the frequency matrix table to do individual searches independently.

This allows the user to build up tables of word relationships in PubMed in the context of your experiments or your scientific interests. This is useful for analyzing combinatorial datasets, as found with multiplex experimental systems, such as cDNA microarrays, genomic, proteomic, or other multiplex comparisons. The PubMatrix database is an archive of previous searches on many topics.

## Instructions

- 1) **Register**; After registering, your browser will be redirected to the main page on the secure site where you are required to login with the username and password that you created during the registration process. You only need to register once, on subsequent visits you will go directly to step 2 and login with the username and password that you created during registration.
- 2) If you have registered, go to the **[authenticated site](#)**.