The UAB Proteomics Database

Aubrey Hill¹,* and Helen Kim²

¹UAB Comprehensive Cancer Center, ²Department of Pharmacology and Toxicology, UAB Mass Spectrometry and Proteomics Shared Facility, University of Alabama at Birmingham, 1530 3rd Ave South, Kaul 602A, Birmingham, AL 35294, USA

Received on February 21, 2003; revised on April 25, 2003; accepted on May 8, 2003

ABSTRACT

Summary: The University of Alabama at Birmingham (UAB) Proteomics Database (UPD) (http://www.uab.edu/proteinmenu) was created to provide a repository for the storage and linkage of two-dimensional (2D) gel images and the associated information obtained through mass spectrometry analysis of the proteins excised from the 2D gels in a manner similar to the SWISS-2DPAGE database and the Stanford Microarray Database. This was accomplished through the development of a web interface, a relational database, image maps and hyperlinks stored in the database. In addition to the internally generated data, UPD provides links to the National Center for Biotechnology Information via accession number hyperlinks. UPD currently contains information on 44 individual proteins derived from four experiments conducted by four UAB faculty members. Images of the gels from which each of these proteins was isolated are accessed by hyperlinks embedded in the database.

Availability: The UAB Proteomics Database can be accessed at http://www.uab.edu/proteinmenu

Contact: ahill@uab.edu

IMPLEMENTATION

UPD is implemented as a 3-tier software architecture consisting of a web interface, a relational database, image maps and hyperlinks stored in the database. This was accomplished through the development of a website, a relational database, image maps and hyperlinks stored in the database. In addition to the internally generated data, UPD provides links to the National Center for Biotechnology Information via accession number hyperlinks. UPD currently contains information on 44 individual proteins derived from four experiments conducted by four UAB faculty members. Images of the gels from which each of these proteins was isolated are accessed by hyperlinks embedded in the database.

Availability: The UAB Proteomics Database can be accessed at http://www.uab.edu/proteinmenu

Contact: ahill@uab.edu

INTRODUCTION

Proteomics is an increasingly important area of research with great potential to yield both basic and medically useful information. The separation and enumeration by 2-D gel electrophoresis of proteins associated with various conditions and their subsequent identification by mass spec generates massive amounts of data which must then be processed, stored and made accessible in a manner similar to the SWISS-2DPAGE database (Sanchez et al. 1995; Hoogland et al. 2000) and the Stanford Microarray Database (Sherlock et al. 2001).

The UPD provides a solution to these needs. The design of the UPD assumes that biomedical researchers are not interested in the raw data at its point of acquisition (such as mass spec raw data), but instead expect processed data which is meaningful at a biological level. A second and equally important design principle was to provide a very structured, yet flexible and robust query facility. These two objectives were accomplished by storing proteomics data from a spot/protein perspective and allowing relational queries to be constructed from any combination of attributes and values. The result of the queries include hyperlinks to both an image map of the original gels as well as a hyperlink to a corresponding mass spec data sheet. The gel image map in turn provides a hyperlink from each protein spot back to the mass spec data.

*To whom correspondence should be addressed.
Fig. 1. Query menu.

Fig. 2. Query result.
The UAB proteomics database

record at NCBI. Also shown in Figure 2 is a hyperlink to the original gel from which this protein was isolated. In this case, clicking on human-breast-cell displays the gel image, where Spot 1 is human cytoskeletal keratin-8.

FUTURE ENHANCEMENTS

Future enhancements will include a hyperlinked description of the experimental objectives and results for each set of proteins identified from a given set of gels, as well as a means to perform quantitative comparisons between samples using statistical programs such as SAS. Additional hyperlinks will be added to point to other databases such as the Online Mendelian Inheritance in Man.

REFERENCES