

Sector: Information Technologies**Subsector:** Bioinformatics

High-Throughput Proteomics Pipeline

Information Summary:

Reference code:	ROI 06013
Technology Overview:	A bioinformatics pipeline for the high-throughput analysis of data from bottom-up proteomics experiments
Application:	Proteomics research
Validation	Used by the Genome Quebec Proteomics Platform
Inventors:	Kearney Robert et al.
Contact:	Francois Labonte Francois.Labonte@mcgill.ca (514) 398 5786

Benefits

- **Efficient Design** – Relies on data structures and algorithms for the annotations of proteins to assist in inferring biological significance from protein identification. The algorithms allow sorting results into a unified proteomics database for data mining purposes across multiple experiments.
- **Modular Structure** – Compatible with commercial, open source or in-house software to perform various functions.
- **Data Sharing** – Facilitates the overall flow of data between researchers working on different projects in the same or different facilities.
- **Integrated Aspect** – Supports all phases of mass spectrometer proteomics from sample preparation to protein grouping. The software can be tailored to the user's needs whether it is a low volume on-demand environment or high-throughput proteomics.

Technology Description:

We have developed a bioinformatics pipeline for the high-throughput analysis of data from bottom-up proteomics experiments. The system is built on a comprehensive relational database and incorporates components for laboratory information management, peptide and protein identification from mass spectra, inter-sample comparisons and analysis, annotation, and sophisticated report generation. The system is fully operational and in daily production in the Genome Quebec Proteomics Platform. It is distinguished from other systems in two main ways: 1) it is designed to be used in service platforms and to manage multiple projects in parallel; and 2) the analysis and reporting mechanisms are built to handle complex samples; that is experiments involving comparisons of data sets involving multiple conditions, biological samples, or time points.

Opportunity

- **Rapid entry to the market** – Minimal customization and software development efforts required for the market launch of the currently available version of the Pipeline.
- **Support tool for added-value services related to proteomics** – The design of the pipeline as a service platform allows offering added-value services related to proteomics research. Instead of purchasing dedicated equipment and software, a research laboratory would outsource the conduct of its proteomics experiments to a service provider exploiting the proteomics platform.



Dr. Robert E. Kearney

Professor; Departments of Biomedical Engineering.

Fellow of the Engineering Institute of Canada

Fellow of the American Institute of Medical and Biological Engineering

Fellow of the Institute of Electrical and Electronics Engineering.

Research focus:

Bioinformatics for Proteomics: The acquisition, analysis and interpretation of proteomics data require a series of well-integrated, high-throughput technologies to lead the researcher from experimental design to biological insight. Bioinformatics provides the tools to implement individual analyses, link the technologies, maximize throughput, and optimize the information extracted from the data. It also facilitates the overall flow and accessibility of data among scientists. For the past several years Dr. Kearney has led the bioinformatics for Proteomics Platform of the McGill University Genome Quebec Innovation center. His efforts focus on three specific objectives. (1) The deployment and operation of a robust, secure *information technology infrastructure* to support the acquisition, analysis and interpretation of proteomics data; (2) The development of *CellMapBase* a custom, database and web-based application for the distributed acquisition and analysis of proteomics data; and (3) The elucidation and evaluation of new *algorithms and tools* to assess, and improve the efficiency and accuracy of protein identification.

Biomedical System Identification and Signal Analysis: Development of tools and techniques for the analysis of biomedical signals and system. The emphasis is on practical methods for the identification of linear-time-varying and nonlinear systems within a continuous-time, nonparametric context. Related to this are methods for relating the structure and parameters of model estimates to the underlying biomedical mechanisms.

Human Motor Control: Research into the role of the peripheral neuromuscular system in the control of posture and movement. System identification methods are used to address three main questions: (1) What are the mechanical properties of human joints and how do they vary under normal physiological conditions. (2) What mechanisms are responsible for generating the mechanical behavior; what are the relative roles of intrinsic muscle properties and reflex mechanisms. (3) What role do these mechanical properties play in the control of posture and movement.