The Proteome Browser - A Software Resource for the Chromosome-Centric Human Proteome Project

Robert J. A. Goode¹, Anthony Beitz², Jeffrey Christiansen³, Edouard C. Nice⁴, A. Ian Smith⁵

¹Monash University, Melbourne, Australia, robert.goode@monash.edu
²Monash University, Melbourne, Australia, anthony.beitz@monash.edu
³Australian National Data Service (ANDS), Melbourne, Australia, jeff.christiansen@ands.org.au
⁴Monash University, Melbourne, Australia, ed.nice@monash.edu
⁵Monash University, Melbourne, Australia, ian.smith@monash.edu

DESCRIPTION
In 2010 the Human Proteome Organisation launched the Human Proteomic Project (HPP), aimed at cataloguing the protein information arising from the plethora of worldwide proteomic based studies. To support complete coverage, one arm of the project will take a gene- or chromosomal-centric strategy (C-HPP). The approach to dividing labour in this international effort has been to assign each of the 24 human chromosomes to one or more countries. In this scheme, the Australian/New Zealand consortium has been assigned Chromosome 7, as this chromosome contains various genetic markers associated with diseases relevant to the Australian population.

Despite multiple large international biological databases housing genomic and protein data, there is currently no single system that integrates up-to-date pertinent information from each of these data repositories and assembles the information into a format suitable for a global proteomics effort of the type proposed by the C-HPP.

We have undertaken to produce a data integration and analysis software system for the C-HPP effort and to make data collections from this resource discoverable through ANDS's Research Data Australia. Whilst the software is being designed to be ultimately species and chromosome independent, the initial focus is on the development of a resource for Human Chromosome 7.

During the first phase, up to 4 widely used data sources are being integrated into a web browser interface designed to display an overview of the current evidence supporting the identification of various gene products across the chromosome, such as protein expression, modification and disease association, with the ability to drill down to the original data. Our design allows for easy addition of both new data sources and data categories which would be incorporated in the second phase.

We will present our vision and progress towards the goal of developing a comprehensive data integration and analysis software tool that provides a snapshot of our current proteomic knowledge and will ultimately assist analysing biological function, and the study of human disease.

REFERENCES
ABOUT THE AUTHORS

**Mr Robert Goode** works under Professor Ed Nice in the Clinical Biomarker Discovery and Validation laboratory at Monash University. He has over 10 years experience working in the proteomics field, studying and working under Prof. Richard Simpson at Ludwig Institute for Cancer Research. As a self-taught Java programmer, he is assisting the development team by liaising with the wider proteomics community and translating their needs into software design terms to aid the development of the browser.

**Mr Anthony Beitz** is Manager of Monash University’s e-Research Centre. He has extensive experience in the selection, reuse, development, and deployment of eResearch software infrastructure. He has been project and product manager of major eResearch projects, e.g. DART integration and the ARCHER portal and dataset, and extensive consultancy experience with both technical & non-technical groups. Prior to his appointment at Monash he had over a decade in research and development with Telstra, managing the development and deployment of a range of software tools and technologies.

**Dr Jeff Christiansen** is a Senior Business Analyst at the Australian National Data Service (ANDS), and is responsible for liaison between ANDS and the software development team at Monash. Prior to his role at ANDS he was based in Edinburgh for 9 years as the Project Manager and Senior Curator of the EMAGE mouse gene expression database, and before this he conducted biological research for over 10 years - completing post-doctoral research into early brain development in London, and a PhD in Brisbane, Australia.

**Professor Edouard Nice** is Ed Nice is currently the Director of the Monash Antibody Technologies Facility (MATF) and Head of Clinical Biomarker Discovery and Validation at Monash University. Ed’s research has centred on the purification and characterisation of biologically significant proteins and peptides. In particular he has developed a number of techniques for measuring protein-protein interactions including antibody characterisation. Ed is internationally recognised in the fields of HPLC, micro-purification, biosensor analysis and proteomics. He has published over 170 peer-reviewed articles in these fields.

**Professor Ian Smith** Ian is Pro-Vice Chancellor (Research & Research Infrastructure) at Monash University and as such he has responsibility for the oversight and management of the university’s research infrastructure as well as developing and implementing strategies to meet current future university infrastructure needs. Professor Smith is the Director of the Victorian node of the NCRIS funded Proteomics Australia Consortium. His research, which is well recognised internationally, applies proteomics technologies to study the proteases involved in the generation and metabolism of peptide regulators of both brain and cardiovascular function. He serves on seven international editorial boards and has published over 200 papers. Ian collaborates and consults widely with pharmaceutical and biotechnology industries and has filed many patents. He is a co-founder of a proteomics-based, publicly listed, biotechnology company and sits on many national and international boards and committees. Over the years he has held office-bearing positions in a many national and international societies and has been actively involved in the organisation of numerous scientific meetings.