JOURNAL OF PROTEOMICS 100 (2014) 60–67

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Review

The Human Proteome Organization Chromosome 6 Consortium: Integrating chromosome-centric and biology/disease driven strategies☆

C.H. Borchers[a], J. Kast[b], L.J. Foster[c], K.W.M. Siu[d], C.M. Overall[e], T.A. Binkowski[f], W.H. Hildebrand[g], A. Scherer[h], M. Mansoor[i], P.A. Keown[i,j,*], for the Human Proteome Organization Chromosome 6 Consortium

[a]University of Victoria/Genome BC Proteomics Centre, Victoria, BC, Canada
[b]Biomedical Research Centre, University of British Columbia, Vancouver, BC, Canada
[c]Centre for High Throughput Biology, University of British Columbia, BC, Canada
[d]Centre for Research in Mass Spectrometry, York University, Ontario, Canada
[e]Centre for Blood Research, Faculty of Dentistry, University of British Columbia, Canada
[f]Midwest Centre for Structural Genomics, Argonne National Laboratory and Computation Institute, University of Chicago, USA
[g]Department of Microbiology and Immunology, University of Oklahoma, OK, USA
[h]Australian Genome Research Facility, Walter and Eliza Hall Institute, Parkville, Australia
[i]Department Medicine, University of British Columbia, Vancouver, BC, Canada
[j]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

ARTICLE INFO

Available online 8 August 2013

Keywords:
Human Proteome Project
Proteomics
Chromosome 6

ABSTRACT

The Human Proteome Project (HPP) is designed to generate a comprehensive map of the protein-based molecular architecture of the human body, to provide a resource to help elucidate biological and molecular function, and to advance diagnosis and treatment of diseases. Within this framework, the chromosome-based HPP (C-HPP) has allocated responsibility for mapping individual chromosomes by country or region, while the biology/disease HPP (B/D-HPP) coordinates these teams in cross-functional disease-based groups. Chromosome 6 (Ch6) provides an excellent model for integration of these two tasks. This metacentric chromosome has a complement of 1002–1034 genes that code for known, novel or putative proteins. Ch6 is functionally associated with more than 120 major human diseases, many with high population prevalence, devastating clinical impact and profound societal consequences. The unique combination of genomic, proteomic, metabolomic, phenomic and health services data being drawn together within the Ch6 program has enormous potential to advance personalized medicine by promoting robust biomarkers, subunit vaccines and new drug targets. The strong liaison between the clinical and laboratory teams, and the structured framework for technology transfer and health policy decisions within Canada will increase the speed and efficacy of this transition, and the value of this translational research.

☆ This article is part of a Special Issue: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

* *Corresponding author at: Department of Pathology and Laboratory Medicine, University of British Columbia, Room 1559, Immunology, Vancouver General Hospital, Vancouver V5Z1M9, Canada.* Tel.: +1 604 875 4393; fax: +1 604 875 4911.
E-mail address: paul_keown@yahoo.com (P.A. Keown).

http://dx.doi.org/10.1016/j.jprot.2013.08.001

Biological significance
Canada has been selected to play a leading role in the international Human Proteome Project, the global counterpart of the Human Genome Project designed to understand the structure and function of the human proteome in health and disease. Canada will lead an international team focusing on chromosome 6, which is functionally associated with more than 120 major human diseases, including immune and inflammatory disorders affecting the brain, skeletal system, heart and blood vessels, lungs, kidney, liver, gastrointestinal tract and endocrine system. Many of these chronic and persistent diseases have a high population prevalence, devastating clinical impact and profound societal consequences. As a result, they impose a multi-billion dollar economic burden on Canada and on all advanced societies through direct costs of patient care, the loss of health and productivity, and extensive caregiver burden. There is no definitive treatment at the present time for any of these disorders.
The manuscript outlines the research which will involve a systematic assessment of all chromosome 6 genes, development of a knowledge base, and development of assays and reagents for all chromosome 6 proteins. We feel that the informatic infrastructure and MRM assays developed will place the chromosome 6 consortium in an excellent position to be a leading player in this major international research initiative.
This article is part of a Special Issue: Can Proteomics Fill the Gap Between Genomics and Phenotypes?

Contents

1. Introduction

Chromosome 6 (Ch6), a metacentric chromosome 171.11 Mbs in length, contains approximately 6% of the human genome [1]. The first gene map was completed in 2003, and current sequence data identify a total complement of between 2344 and 2780 genes, with an average density of 16.2 genes per Mb [2]. Between 1002 and1034 of these genes code for known, novel or putative proteins, and about 2.2% of the chromosome is occupied by exons with a mean length of 281 Bps. More than 350 other genes code for miRNA, snRNA, snoRNA and miscellaneous transcripts, while a further 700 are processed or unprocessed pseudogenes [2]. Recent studies have identified genes related to critical biological functions throughout the length of Ch6, of which the largest is the PARK2 gene on the q arm (1.4 Mb, 12 exons) [3–5]. These genes code for approximately 3000 known protein transcripts expressed in extracellular, intracellular or membrane compartments, many are involved in immunity, inflammation, neuronal activities and other critical cellular activities, of which key examples are presented in Table 1.

Of the several discrete regions within the chromosome, one of the most prominent is the extended major histocompatibility complex (eMHC). This 7.6 Mb super-region is located on the short arm of Ch6 and extends telomerically from RPL12P1 to HIST1H2AA [2]. The five sub-regions of the eMHC contain 523 genes, of which approximately 260 (50%) are expressed [2,6]. The eMHC is the most gene-rich region of the human genome, with a density of over 68 total genes and 35 protein coding genes per Mb. Several functional gene clusters have been defined within this extended region (six clusters and six superclusters) of which the two largest and potentially overlapping are the histone and tRNA genes. Transcripts of both are highly required in biological regulation and may be under selection pressure to cluster in association with the MHC [6]. The Human Leukocyte Antigen (HLA) genes located within the eMHC at 6p21.3 are critically related to infection, immunity and inflammation. The more than 200 genes within this hypervariable cluster are divided into 3 regions designated as class I, II and III. The HLA genes are typically highly polymorphic and exhibit tight linkage disequilibrium. More than 8000 alleles have now been identified within the HLA genes, coding for an estimated 6800 proteins of which only 2% can be serologically distinguished by current antibody methods.

1.1. *Role in human diseases*

Chromosome 6 is functionally associated with more than 120 major human diseases, including cancer, heart disease,