
Preface

The concept of proteogenomics, utilizing advances from the fields of proteomics and genomics, was introduced at around the time of the completion of the sequencing of the human genome. The emergence of proteogenomics is mainly due to the rapid development of two key technologies: high-throughput DNA sequencing and mass spectrometry-based proteomics. The ability to determine protein sequences by mass spectrometry has provided a unique tool to the identification and the verification of novel genes, predicted exons, and open reading frames. Consequently, proteogenomics has been used for genome annotation, including the validation of known or annotated protein-coding genes; the improvement of gene annotations assigning correct start sites; the mapping of signal peptides, proteolysis, and other posttranslational modifications (an important element of biological function that is not encoded directly in the genome); as well as the identification of splicing variants and mutant proteoforms often associated with disease progression.

Considering the rapid advancement in the field, it is perhaps appropriate to define proteogenomics as an intensive research area that investigates the correlations between proteomic data and their corresponding genomic and transcriptomic data, keeping the goal to improve our knowledge about life at the molecular level, which is a more complete view that has been initially suggested. The interplay between the two data streams of genomics and proteomics certainly allows for a better understanding of biological functions and molecular mechanisms in health and disease. Today, genome sequencing provides nearly complete coverage, including transcriptome profiling, while targeted proteomics can be focused on specific regions of the proteome and determine predicted proteins.

The goal of this book is to display this extended view on proteogenomics, depicting research areas where proteogenomics is actively playing an essential role and also highlighting some emerging research arenas without pretending to cover all fields of application. The chapters of this book offer the readers a general insight to the integrative analyses of various types of omics data and present advances within specific principles, such as next-generation sequencing of DNA, mRNA sequencing, ribosome profiling, as well as mass spectrometry- and antibody-based proteomics. The applications are selected to exemplify the great potential of proteogenomics to contribute to human disease research, particularly to cancer and personalized medicine.

Importantly, this book attempts to identify some common features that integrate the various fields and areas where intensive efforts should be made to drive research more efficiently in the near future. One of these is certainly bioinformatics, which has shown amazing power and development during the last couple of years and which is anticipated to provide powerful approaches to improve our ability to work with and combine the large data sets that genomics, transcriptomics, and proteomics generate.

At last, I would like to thank all the authors of this book for their exceptional contributions, sharing their expert views of the field, and presenting their original research. Their enthusiasm and timely delivery of their manuscripts helped me tremendously to realize this project. It is my sincere hope that the readers would enjoy this book as much as I enjoyed preparing it.

Galveston, TX, USA
March 1, 2016

Ákos Végvári

Proteogenomics

Végvári, Á. (Ed.)

2016, VIII, 184 p. 57 illus., 55 illus. in color., Hardcover

ISBN: 978-3-319-42314-2