

BIOINFORMATICS RESOURCES FOR THE STUDY OF VIRUSES AT THE VIRGINIA BIOINFORMATICS INSTITUTE

Anjan Purkayastha, Oswald Crasta, J Dana Eckart, Michael Czar, X.J. Meng, Joao Setubal, and Bruno Sobral

INTRODUCTION

This chapter gives an overview of two bioinformatics resources that are being developed at the Virginia Bioinformatics Institute (VBI) for the study of pathogenic microorganisms, with a special focus on pathogens listed as biodefense priority agents by the Centers for Disease Control (CDC) and the National Institute of Allergy and Infectious Diseases (NIAID). The Pathogen Resource Integration Center (PATRIC) has been formed to provide a comprehensive and centralized annotation resource for eight classes of human pathogens, including all sequenced species, subspecies, strains and isolates. The Pathogen Portal was established to provide the scientific community with an integrated biological data analysis platform that includes multiple and diverse data sets on microbial pathogens and a set of analysis tools.

Advances in genomics and computer technologies are fueling the development of a knowledgebase from large data sets created by high-throughput genomic technologies using digital computation, data, information, and networks. The integration of knowledge from various fields such as computer science, chemistry, and biology has created a vast opportunity by creating new research environments based upon Cyberinfrastructure (Atkins et al., 2003). Cyberinfrastructure refers to the infrastructure required to develop a knowledgebase based upon distributed computing, information, and communication technology. One of the major challenges during the post-genomic era is the integration of diverse data sets (Stein, 2003). As described by Kanehisa and Bork (2003), the main goal of bioinformatics during the 1990s was to create primary databases of genes and proteins. Currently the focus is to extend the databases for quantitative data from transcriptomes and proteomes and to provide interoperability between multiple disciplines. The bioinformatics resources developed under such endeavors are expected to assist the scientific community in making use of the large datasets produced by high-throughput technologies in elucidation of the mechanisms of phenotypic expression. Understanding of these mechanisms may result in the identification of specific targets and products relevant to public health. Both PATRIC and PathPort fall under the auspices of the Cyberinfrastructure Group at VBI whose goal is to provide a scalable, flexible and interoperable technology framework that facilitates biodiscovery.