Thomas Triplet, J Comput Sci Syst Biol, 5:1
doi: http://dx.doi.org/10.4172/0974-7230.S1.07

OMICS Group Conferences
Accelerating Scientific Discovery

International Conference and Exhibition on
Metabolomics & Systems Biology

20-22 February 2012 San Francisco Airport Marriott Waterfront, USA

Systems biology warehousing: A benchmark of frameworks for effective data integration

Thomas Triplet
Concordia University, Canada

The rapid development of -omics techniques have provided an unprece-dented amount of data, enabling system-wide biological research. Although information integration has been well investigated in database theory research, biological data present numerous challenges from the lack of standard formats to data inconsistencies resulting from experimental data variations. However, the success of systems biology is contingent on the ability to integrate and utilize a wide variety of types of data. It also relies on computational techniques to automatically predict and assign functional annotations of proteins as effective integration of biological data should enable scientists to perform comparative analyses, modeling and inference of protein functions.To date, no biological data warehouse meets all the requirements for effective integration of system-wide data. BioXRT offers a exible and extensible database structure, BioMart provides advanced data-mining tools although they may not be extended by users. PROFESS features a exible and modular user interface and tools for clustering and statistical analysis of large datasets. InterMine also features a customizable userinterface and is helpful to track the provenance of data. The Open Genome Resource is an open source system for the storage, visualization and analysis of prokaryotic genome data that can be automatically download annotations from relevant databases. GeWare is a laboratory information management system, featuring tools for the integrated analysis of clinical data from large biomedical research studies. However, there now exists a variety of resources that may be helpful in accommodating data inaccuracies, such as approximate string matching or similarity-based algorithms that may be implemented within database management systems for the next generation of biological data warehouses.

Biography

Dr. Thomas Triplet obtained an engineer diploma and MSc. in Computer Science and Engineering in the French Grande Ecole ENSICAEN with distinctions and completed his Ph.D in Bioinformatics at the age of 24 years from the University of Nebraska-Lincoln. He has published more than 10 refereed papers and has frequently served as referee for international journals and conferences. His research, rewarded in 2008/2009 by the Milton E. Mohr fellowship, is primarily on biological data warehousing and analysis using machine learning techniques. He is currently a post-doctoral fellow with the Centre for Structural and Functional Genomics at Concordia University, Montreal, Canada.