Chapter I
Pathway Biology Approach to Medicine

Peter Ghazal
University of Edinburgh Medical School, Scotland,
and Centre for Systems Biology at Edinburgh, Scotland

ABSTRACT

Systems biology provides a new approach to studying, analyzing, and ultimately controlling biological processes. Biological pathways represent a key sub-system level of organization that seamlessly perform complex information processing and control tasks. The aim of pathway biology is to map and understand the cause-effect relationships and dependencies associated with the complex interactions of biological networks and systems. Drugs that therapeutically modulate the biological processes of disease are often developed with limited knowledge of the underlying complexity of their specific targets. Considering the combinatorial complexity from the outset might help identify potential causal relationships that could lead to a better understanding of the drug-target biology as well as provide new biomarkers for modelling diagnosis and treatment response in patients. This chapter discusses the use of a pathway biology approach to modelling biological processes and providing a new framework for experimental medicine in the post-genomic era.

INTRODUCTION

An increasing number of biological experiments and more recently clinical based studies are being conducted using large-scale genomic, proteomic and metabolomic techniques which generate high-dimensional data sets. Such approaches require the adoption of both hypothesis and data driven strategies in the analysis and interpretation of results. In particular, data-mining and pattern recognition methodologies have proven particularly useful in this field. The increasing amount of information available from high-throughput experiments has initiated a move from focussed, single gene and protein investigations

to the study of multiple component interactions. Vitally, when the output from high dimensional data is integrated with the wealth of information from previously published investigations, the assembly of known and novel network characteristics is possible. The cause-effect association and annotation of multiple genes and gene products by such methods can aptly be described as pathway biology.

Pathway data is variously categorised in terms of metabolic pathways, molecular interactions, gene regulatory networks, signalling pathways and which, are often represented differently and in isolation. In recent years there has been an increasing effort in representing biological pathways using computer science based methodologies. These efforts include databases that aim to curate pathways, such as KEGG (Kanehisa et al., 2006), Reactome (Joshi-Tope et al., 2005), aMAZE (Lemer et al., 2004) or PATIKA (Dogrusoz et al., 2005); databases of experimentally and computationally derived protein interactions, such as MPPI (Pagel et al., 2005) and DIP (Salwinski et al., 2004); or tools that aim to extract pathway information from the scientific literature, examples include PathwayAssist (Nikitin et al., 2003) and Ingenuity Pathway Analysis (Ingenuity Systems). However, a common factor in all these resources is the need to describe pathways visually in order to understand their complexity

In this regard a biological interaction network based on electronic circuitry diagrams provided an informative approach and aims to provide a compact description of the entity relationships in a pathway (Kohn, 1999; Kohn et al., 2006). Kitano et al have extended this approach by introducing a simple notational system for state transitions (Funahashi et al 2003) for the representation of process flow in signalling pathways (Kitano, 2003, Kitano et al 2005). While considerable advances are being made there will be an increasing need to ensure that such pathway descriptions remain intuitive to biologists in general, involving logic and, in particular, an integrative view of molecular interactions, gene regulatory networks and signalling pathways while maintaining compliance with a more formal description of biological processes. This has been the primary motivation behind the Edinburgh Process Notation which uses and exends core aspects of the Kitano notation (Moodie et al, 2006). Collectively, these efforts have been and are part of a community wide effort to develop graphical notation standards (www.sbgn.org). To date much of this work has focussed on experimental cell signalling based systems and little work has been conducted regarding physiological systems and using clinical data.

Here, we discuss a practical guide to a pathway biology approach in medicine. The intended aim is to help translate systems biology from bench science to medical research and ultimately toward clinical use. We particularly put forward the use of logic as a strategy for studying pathways and present arguments for the suitability of logical models for the analysis of clinically derived data.

LITERATURE AND DATA MINING: STAMP COLLECTING

The first task at hand is toward the acquisition of pathway information. In this regard information relating to the components and interactions of pathways need to be compiled, integrated and visualised using research synthesis methodology (Cooper and Hedges 1994). This generally follows a four stage process:

i. A literature review should be undertaken to identify relevant pathway components and interactions. This can be performed using standard Entrez PubMed queries involving keywords, author searches and the use of Boolean operators. A variety of tools can also be used to facilitate this process, e.g. PDQ wizard (Grimes et al 2006). A manual review of the resultant articles is an es-