Chapter 12
Using Functional Linkage Gene Networks to Study Human Diseases

Bolan Linghu
Novartis Institutes for BioMedical Research, USA

Guohui Liu
Millennium Pharmaceuticals Inc, USA

Yu Xia
Boston University, USA

ABSTRACT

A major challenge in the post-genomic era is to understand the specific cellular functions of individual genes and how dysfunctions of these genes lead to different diseases. As an emerging area of systems biology, gene networks have been used to shed light on gene function and human disease. In this chapter, first the existence of functional association for genes working in a common biological process or implicated in a common disease is demonstrated. Next, approaches to construct the functional linkage gene network (FLN) based on genomic and proteomic data integration are reviewed. Finally, two FLN-based applications related to diseases are reviewed: prediction of new disease genes and therapeutic targets, and identification of disease-disease associations at the molecular level. Both of these applications bring new insights into the molecular mechanisms of diseases, and provide new opportunities for drug discovery.

INTRODUCTION AND BACKGROUND

With the development of sequencing technologies, whole genome sequencing has been achieved for diverse species (Flicek et al., 2008). For a fully sequenced organism, most protein-encoding genes can be readily identified by available bioinformatics approaches (Flicek et al., 2008). By contrast, it remains a challenging task to understand the

DOI: 10.4018/978-1-60960-491-2.ch012

specific biological functions of these genes, and how dysfunctions of these genes lead to diverse human disease phenotypes.

A particular cellular function usually requires the collaboration between a specific group of genes or proteins. Rather than acting alone, these genes interact and communicate with each other in diverse ways, but all for the common purpose of maintaining the normal status of a specific biological process (Kanehisa et al., 2004). On the other hand, when one or more genes involved in a particular biological process are dysfunctional, the normal status of the biological process might be perturbed, which might further cause the organism to show abnormal physiological phenotypes referred to as a disease (Goh et al., 2007). Correspondingly, therapeutic drugs aim to target the genes or proteins involved in these perturbed biological processes such that the normal status of the biological processes can be reestablished (Janga & Tzakos, 2009). Therefore, for gene function and human disease research, it is very important to consider individual genes as functional related components within a coherent biological system.

Recent network-based approaches have demonstrated great success in representing functional relationships among genes with applications to understand gene function and human disease (Ahmed & Xing, 2009; Franke *et al.*, 2006; Huttenhower *et al.*, 2009; Kohler *et al.*, 2008; Lage *et al.*, 2007; I. Lee *et al.*, 2008b; Linghu *et al.*, 2008; Linghu *et al.*, 2009; McGary *et al.*, 2007; Oti & Brunner, 2007; Oti *et al.*, 2006; Schadt, 2009). In these networks, nodes represent genes, and edges represent functional associations between linked genes. These networks are referred to as functional linkage gene networks (FLN). In this chapter, we first review the molecular basis for genes working as a functional group, and demonstrate the existence of functional associations between genes implicated in a common disease. Next, we review ways to construct different types of FLNs, as well as two important FLN-based applications related to human diseases: prediction of new disease genes and therapeutic targets, and identification of disease-disease associations at the molecular level (Figure 1).

MAIN FOCUS OF THE CHAPTER

Functional Associations between Genes Underlying the Same or Related Diseases

Genes Work in Groups to Carry Out Particular Cellular Functions

Most genes or proteins do not work alone. Instead, a group of genes or proteins collaborate with each other efficiently as a specific functional module to carry out a particular cellular task (Hartwell et al., 1999; Ravasz et al., 2002). Such functional modules can be represented as specific biological processes or pathways. Genes or proteins within the same biological process or pathway can have multiple types of functional associations. For instance, certain proteins physically interact with each other to form a protein complex and function as a whole unit (Stelzl et al., 2005); certain transcription factors regulate a group of target genes in order to coordinate cellular activities related to a particular biological process such as cell cycle (Chen et al., 2000); certain genes with similar sequences encode members of a protein family such that different members can be used under different cellular conditions (Iwabe et al., 1996). These functional associations can be inferred from various data sources generated by different experimental and computational approaches. For instance, yeast two-hybrid and mass spectrometry experiments can detect protein-protein binary physical and co-complex interactions (Ewing et al., 2007; Rual et al., 2005); microarray experiments can detect co-expression relationships among genes at the transcriptional level (Griffith