Network analysis: a new approach to study endocrine disorders

A Stevens1,2, C De Leonibus1,2, D Hanson1,2, A W Dowsey1,3, A Whatmore1,2, S Meyer4, R P Donn5, P Chatelain6, I Banerjee1,2, K E Cosgrove7, P E Clayton1,2 and M J Dunne7

1Faculty of Medical and Human Sciences, Institute of Human Development, University of Manchester, Manchester, UK
2Manchester Academic Health Science Centre, Royal Manchester Children’s Hospital, Central Manchester University Hospitals NHS Foundation Trust, 5th Floor, Oxford Road, Manchester M13 9WL, UK
3Paediatric and Adolescent Oncology, The University of Manchester, Manchester M13 9WL, UK
4Stem Cell and Leukaemia Proteomics Laboratory, School of Cancer and Imaging Sciences, The University of Manchester, Manchester M20 4BX, UK
5Musculoskeletal Research Group, NIHR BRU, University of Manchester, Manchester M13 9PT, UK
6Department Pediatrie, Hôpital Mère-Enfant, Université Claude Bernard, 69677 Lyon, France
7Faculty of Life Sciences, University of Manchester, Manchester M13 9NT, UK

Correspondence should be addressed to A Stevens
Email adam.stevens@manchester.ac.uk

Abstract

Systems biology is the study of the interactions that occur between the components of individual cells – including genes, proteins, transcription factors, small molecules, and metabolites, and their relationships to complex physiological and pathological processes. The application of systems biology to medicine promises rapid advances in both our understanding of disease and the development of novel treatment options. Network biology has emerged as the primary tool for studying systems biology as it utilises the mathematical analysis of the relationships between connected objects in a biological system and allows the integration of varied ‘omic’ datasets (including genomics, metabolomics, proteomics, etc.). Analysis of network biology generates interactome models to infer and assess function; to understand mechanisms, and to prioritise candidates for further investigation. This review provides an overview of network methods used to support this research and an insight into current applications of network analysis applied to endocrinology. A wide spectrum of endocrine disorders are included ranging from congenital hyperinsulinism in infancy, through childhood developmental and growth disorders, to the development of metabolic diseases in early and late adulthood, such as obesity and obesity-related pathologies. In addition to providing a deeper understanding of diseases processes, network biology is also central to the development of personalised treatment strategies which will integrate pharmacogenomics with systems biology of the individual.

Key Words

- modelling
- microarray
- endocrine disruptors
- secretion
- network biology

Introduction

Network biology has developed as a method to study the many interactions that occur in individual cells, helping to understand the complex biological processes in molecular biology and how they link to integrated biological systems (Pujol et al. 2010). The recent genome sequencing projects have provided a nearly complete list of human gene products and this has been followed by the first drafts of connectivity maps between proteins
List of proteins, genes, or metabolites from omic data sets

Function and prioritisation
and gene expression profiles for many different tissues (Parkinson et al. 2007). Network biology detects the relationships between different biological components, including genes, proteins, transcription factors, small molecules and metabolites (Arrell & Terzic 2013, Tomar & De 2013), and enables the physical or functional relationships between these components to be identified and interrogated, see Fig. 1 and Box 1. Network analysis can therefore be used to dissect the complexities of biological systems and provide information which is directly relevant to understanding the pathobiology of disorders, including the pathways of disease processes and the prioritisation of candidate disease-causing genes. The aim of this review is to explain how the analysis of biological networks can be carried out and to illustrate how systems biology is impacting on our understanding of diseases, with a particular focus on endocrine conditions which by their very nature are complex and impact upon many organs systems.

Network analysis
How do we visualise biological network models?

The commonest representations of biological networks are ‘ball and stick’ diagrams. The ‘balls’ represent proteins or genes and the ‘sticks’ represent the known interactions. These can be directed, with an arrow representing ‘action’ or ‘flow’ of information, or they can be undirected, where connections represent direct associations. Biological networks are ‘graphs of connected nodes’, where nodes represent biological components, which are connected through edges to show their inter-relationships, from the physical or functional associations to metabolic pathways and regulatory interactions (Arrell & Terzic 2013, Tomar & De 2013; Fig. 1). Several models have been used for network analysis, from ‘Power Graphs’ and the ‘Information Flow Model’, to ‘Boolean’ and ‘Bayesian’ networks (Missiuro et al. 2009, Pujol et al. 2010, Youngs et al. 2013). In Bayesian networks, the nodes represent biological variables and directed edges represent causality through conditional probabilities between them. Specifically, they have been used in several fields of systems biology, such as cellular networks inference, pathway modelling, quality assessment of protein–protein interactions and functional annotation of proteins and gene expression analysis (Pujol et al. 2010).

Commerically available software for the analysis of gene networks are available (Thomas & Bonchev 2010) and specialise in the ‘pattern recognition’ of genes clustering within the networks derived from either a previously defined pathway ‘cannon’ (e.g. Kyoto encyclopedia of genes and genomes (KEGG) database, http://www.genome.jp/kegg/) or from a process of network inference algorithmically determined using known biological function (Calvano et al. 2005). Most scientists in the life sciences define network analysis based on commercial software of this type, i.e. one in which connections are primarily based on text-mining algorithms. These types of software are designed to be easy to use, provide further quality control of data and have had a major impact on the way in which systems biology analysis is integrated with omic analysis. However, there are advantages to more advanced statistical/computational approaches for generating and analysing networks, primarily around using network properties to prioritise genes for further analysis – these will be covered in more detail in the section ‘Computational analysis of molecular interaction networks’.

Analysis of omic datasets: towards integrated analysis

To analyse omic data, two or more contrasting groups are normally compared, such as a ‘treated vs control’ or ‘received dose vs therapeutic response’. A major issue in the analysis of omic datasets with hundreds or thousands of variables is that the number of comparisons is so large that it is likely that the total number of false positives becomes unfavourably high. This phenomenon is moderated by the application of a false discovery rate correction to the statistical test, for example Benjamini–Hochberg correction (Benjamini & Hochberg 1995). However,
common false discovery rate correction methods assume that comparisons are independent, which is often not the case. Increasingly, extra confidence in the analysis of omic data sets can be provided by observing whether identified changes coalesce into clusters within biological networks (D’Eustachio 2011, Haw et al. 2011). These two approaches are not mutually exclusive and are increasingly modelled together to allow interpretation of different types of omic datasets.

Biological networks can be inferred using omic datasets by mapping the data to models of all known protein–protein or –genetic interactions and generating a much larger model of the interactome (Stark et al. 2011, Ngounou Wetie et al. 2013; Fig. 1). An interactome represents the whole set of direct and indirect molecular interactions involved with the phenomenon being studied (Fig. 1). The interactome model therefore combines data from different omics datasets/databases to generate network cluster maps correlating with the organisation and priority of processes that provide a particular function(s). These models also allow multivariate analysis of the data with consequent gains in statistical certainty (Choi & Pavelka 2011).

The availability of whole genome transcript data has also aided the development of novel approaches to analyse this type of data. For example, it is now possible to incorporate the influence of upstream regulators such as transcription factors and miRNA into modelling of available data. These strategies, in turn, may assist and strengthen putative mechanisms of disease (Medan et al. 2012). It is also possible to fully integrate genetic and transcriptomic analyses by the generation of expression quantitative trait loci (eQTL). An eQTL is a genomic locus that regulates transcription and is defined by mapping methods that examine the linkage between variation in expression and genetic polymorphisms (Kulp & Jagalur 2006, Villa-Vialaneix et al. 2013).

**Computational analysis of molecular interaction networks**

The study of biological networks has increased exponentially over the past 10 years and has been paralleled with the increase in collection and statistical analysis of network-indexed data. Statistical methods to examine the relationship between the properties of biological networks and function have also been developed and research in this area is ongoing (Choi & Pavelka 2011). There are now several analytic methods that are commonly used for relating the structure/topology of biological networks to the functional properties of cells and systems (Toubiana et al. 2013). It is well recognised that most biological functions are closed systems at low entropy states and can be resolved to a set of ordered processes. Cluster modularity aims to define this order with the recognition that sub-networks of interconnected nodes might represent molecules which are functionally linked and work in a coordinated manner to achieve a
definite function (Fig. 1C; Gavin et al. 2006, Adjari et al. 2012, Szalay-Beko et al. 2012). Important concepts in analysis include node centrality and network robustness (Barzel & Biham 2009; Box 1 and Fig. 1D). Various mathematical operators are useful in pattern analysis. For example, network motif analysis identifies small networks that are over-represented when compared with a randomised version of the same network (Milo et al. 2002), while network alignment and comparison are used to describe similarities between independent networks; this is a type of analysis that has been particularly valuable in understanding the evolutionarily conserved pathways that are common to several organisms (Choi et al. 2013). Current software approaches to the assessment of network structure and its relationship to biological function are summarised in Table 1.

Whilst the variety of available approaches to the software analysis of network properties can seem daunting and at the current time have limited consensus of use, ‘the ability of these approaches to unravel information not made available via traditional analysis of complex data sets’ (as stated by Toubiana et al. (2013)) is increasingly being recognised.

Software-based methods in network analysis

Different software-based methods can be used to generate biological networks using omic data as ‘seeds’, but all use algorithms to infer relationships between omic changes and known interactions in the literature or from open-access databases such as BioDRID (Stark et al. 2011, Chattr-Aryamontri et al. 2013) and Reactome (Haw et al. 2011). Interactome models generated by these methods can be visualised using software such as Cytoscape, which is an open source software platform for the analysis of complex networks (Cline et al. 2007, Smoot et al. 2011).

Network analysis has also been developed as a process to increase confidence in the observations of differential expression in omic datasets by correlation with biological function (Calvano et al. 2005, Werner 2008). This procedure allows the identification of putative key functional elements within the identified networks of differentially expressed genes. Ingenuity Pathway Analysis (IPA) Software is one of the best known examples of this type of approach. Specifically it uses an algorithm which generates inferred networks of omic data by functional similarity (Calvano et al. 2005). Both the power and the

Table 1 Some of the main applications in use for the identification of network properties, highly connected network regions, associated biological function and the functional prioritisation of network elements. The majority (all except DAPPLE) are plugins for Cytoscape (Smoot et al. 2011)

<table>
<thead>
<tr>
<th>Application name</th>
<th>Analytical features</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ModuLand</td>
<td>Uses: analysis of community centrality. An algorithm for determining extensively overlapping network modules. Moreover, it identifies several hierarchical layers of modules, where meta-nodes of the higher hierarchical layer represent modules of the lower layer</td>
<td>Szalay-Beko et al. (2012)</td>
</tr>
<tr>
<td>Cytohubba</td>
<td>Calculation of network topological properties. Two characteristic analysis algorithms, Maximum Neighborhood Component (MNC) and Density of Maximum Neighborhood Component (DMNC) are developed for exploring and identifying hubs/essential nodes from interactome networks</td>
<td>Lin et al. (2008)</td>
</tr>
<tr>
<td>ClusterONE</td>
<td>ClusterONE (Clustering with Overlapping Neighborhood Expansion) is a graph clustering algorithm that is able to handle weighted graphs and readily generates overlapping clusters</td>
<td>Nepusz et al. (2012)</td>
</tr>
<tr>
<td>MCODE</td>
<td>MCODE is an algorithm that finds clusters (highly interconnected regions) in a network</td>
<td>(Bader &amp; Hogue 2003)</td>
</tr>
<tr>
<td>NetworkPrioritizer</td>
<td>NetworkPrioritizer provides multiple features to rank individual nodes based on centrality measures according to their relevance to the rest of the network</td>
<td>Kacprowski et al. (2013)</td>
</tr>
<tr>
<td>Reactome FIs</td>
<td>Reactome FIs was designed to find network patterns related to condition. Uses community clustering</td>
<td>Newman (2006) and Haw et al. (2011)</td>
</tr>
<tr>
<td>DAPPLE</td>
<td>DAPPLE stands for Disease Association Protein–Protein Link Evaluator. DAPPLE looks for significant physical connectivity among proteins encoded for by genes in loci associated to disease according to protein–protein interactions reported in the literature</td>
<td>Rossin et al. (2011) <a href="http://www.broadinstitute.org/mpg/dapple/dapple.php">http://www.broadinstitute.org/mpg/dapple/dapple.php</a></td>
</tr>
</tbody>
</table>
weakness of this approach are that it is based on ‘mining’ an extensively annotated literature database; this allows highly specific associations to be made but these are only as good as the current database used (Khatri et al. 2012). Literature databases have extensive coverage of biological function but can still suffer from distortions in that some interactions may be over- or under-represented (Khatri et al. 2012). Interactome modelling generated from the databases of known protein: protein and protein:genetic interactions such as BioGRID (Chatr-Aryamontri et al. 2013) suffers from this issue to a lesser extent but requires the more mathematical approaches described in this section to provide functional interpretation.

Analysis of network structure can also be determined by the Cytoscape Software platform, which can be used to provide topological analysis. Topology, a branch of mathematics studying shapes and spaces, is utilised in network analysis to describe network properties such as connectivity, which is an inherent network property involving hubs and bottlenecks (Fig. 1D and Box 1; Lin et al. 2008). Using this approach, the top 10–20% of protein–protein interaction network nodes can be ranked for both connectivity and bottleneck scores and then used to generate a minimal essential network (Sun & Zhao 2010, Stevens et al. 2013d). A minimal essential network represents the most functionally relevant element of an interactome model and therefore is used to assess biological function (Stevens et al. 2013a,b,d).

**Disease networks**

During the last decade, much effort has been spent in generating a list of disease–gene associations, and explaining how genetic disorders and the corresponding disease–genes might be related (Goh et al. 2007, Goh & Choi 2012). However, this process has significant challenges. For instance, diseases can arise from mutations in more than a single gene (locus heterogeneity) and different mutations in the same gene may cause different phenotypes and genetic disorders (allelic heterogeneity) (Goh et al. 2007). In order to advance this area, the ‘human diseasome’ has been generated using network biology, which brings together combined datasets of all known disease–gene associations. This has been created on the basis that there is a cohesive relationship between disease-causing genes and their products, and developed on a conceptual framework which systematically links all recorded genetic disorders (the human disease phenome) with the complete list of known disease-causing genes (the human disease genome) (Janjic & Przulj 2012). Through the network-based model for the human diseasome, it has been shown that genes that contribute to a common disorder exhibit an increased propensity for their products to co-interact through protein–protein interactions; have a tendency to be co-expressed in specific tissues and share common cellular and functional characteristics, as annotated in gene ontology (Pujol et al. 2010, Gene Ontology Consortium 2013).

Disease networks have been created also for orphan diseases that have special consideration as they have a low rate of prevalence. Network relationships based on shared genes or shared functional features have been derived and used to construct the orphan disease-causing mutant gene interactome (Zhang et al. 2011). Interestingly, Zhang et al. (2011) found that in contrast to the common disease-causing mutant genes that are predominantly non-essential, orphan disease-causing mutant genes are largely essential and are ubiquitously expressed. However, the authors also found that the connectivity of gene-based and other feature-based orphan disease networks is largely different, suggesting that the relationships between diseases cannot be fully captured by the gene-based network alone. As network analysis is evolving through the addition of information to databases, these approaches are nevertheless likely to provide important added value for improved diagnosis, prevention, and discovery of novel therapeutic strategies for orphan diseases.

Disease-focused networks are needed to ensure the validation of multiple biological markers and the understanding of pathophysiology of many diseases (Vlahou 2013). Several examples of disease-focused networks are now emerging, especially in the area of cancer research where numerous cancer cluster gene networks have been described in different types of tumours, including colorectal, ovarian and breast cancers (Bravata et al. 2013, Kumar et al. 2013b, Yamaga et al. 2013, Zhang et al. 2013). Network analysis is also leading to promising advances in other fields, including neurological diseases (Parkinson’s and Alzheimer’s diseases (Balthazar et al. 2013, Lones et al. 2013, Strafella 2013)) and neuropsychiatric disorders (Kasperek et al. 2013); diseases of the immune system (Ivanov & Anderson 2013) including autoimmune or haematologic diseases (Mani et al. 2008, Kumar et al. 2013a); the pathogenesis of coronary heart diseases (Huan et al. 2013), fatty liver diseases (Sookoian & Pirola 2013) and endocrine/metabolic conditions, which will be covered in more detail in the following sections.
Network pharmacology

An improved knowledge of the complex nature of diseases is also beginning to impact upon drug development and to the therapeutic management of diseases (Murthy et al. 2013). Whilst most prescribed drugs interact with specific cellular targets, in vivo they modify cellular systems through interactions with multiple proteins rather than single entities. Computational modelling of these interactions is now possible and has facilitated a shift towards network-centric strategies in new drug design (Nacher & Schwartz 2008). As a consequence, a novel concept has now emerged in drug discovery for improving treatment responsiveness termed ‘polypharmacology’ (Barabasi et al. 2011) and has been implicated in the design of novel anticancer drugs (Tang & Aittokallio 2013) and antimicrobial agents (Rohde et al. 2012).

Analysis of network biology is also helping to understand how the multifaceted interactions between drugs and their cellular targets contribute to efficacy and adverse reactions/side effects. Network pharmacology therefore provides a framework for analysing the actions of drugs on systems, for identifying targets which will modulate disease phenotypes, and for developing therapies that are less vulnerable to drug resistance and side effects (Takarabe et al. 2011). Another major area of research in drug discovery is predictive toxicology. Network biology can identify putative secondary targets for a given compound or explore potential downstream effects for blocking or enhancing the actions of key nodes.

Networks analysis of endocrine diseases

In this section, we summarise disease areas in endocrine sciences, which are engaging network analysis to better understand the pathogenesis of disease and the responsiveness of patients to therapy (Fig. 2). In many cases, the introduction of network analysis to endocrine disorders has come about, because candidate gene approaches to pathogenesis are limited, the phenotype of a particular disease in individuals may be variable and because natural history of a condition is often not inferred from traditional investigations.

Congenital hyperinsulinism in infancy

Congenital hyperinsulinism (CHI) is a rare disease due to dysregulated, excess insulin secretion from pancreatic β-cells, which manifests as profound hypoglycaemia (De Leon & Stanley 2007, Banerjee et al. 2013). CHI is the commonest cause of persistent hypoglycaemia in infancy. In contrast to other causes of hypoglycaemia such as infection and hypopituitarism, hypoglycaemia in CHI is unpredictable and severe with a significant adverse impact on future neurodevelopment (Senniappan et al. 2012). The condition usually manifests not only in the neonatal period but also later in childhood with milder forms in children who may be able to maintain normoglycaemia on oral feeds. The phenotype of CHI is heterogeneous, with wide variability not only with respect to time of presentation but also the probability of spontaneous resolution vs long-term persistence of disease. Transient CHI resolves spontaneously within weeks to months of diagnosis and accounts for ~50% of all patients (Arnoux et al. 2011). Little is known about the cause(s) of hypersecretion of insulin in these patients or the mechanisms responsible for the resolution of symptoms. In contrast to transient disease, the pathophysiology of persistent CHI is far more widely described. So far, mutations in nine genes – GLUD1, SLC16A1, HADH, UCP2, KCNJ11, ABCC8, HNF1A, GCK and HNF4A, are known to associate with persistent CHI, but these only help to understand the process of disease mechanisms in...
~ 50% of patients (Banerjee et al. 2011, 2013, Senniappan et al. 2012). In our studies, computational biology of disease-causing genes has recently been used to seed an interaction network model. Genes integral or tightly connected to the network have been found and represent new candidates for understanding the exact aetiology of this disorder across the spectrum of severities (Stevens et al. 2013b).

From the CHI interactome model, cluster analysis of the components of the associated pathways can be carried out (Newman 2006). This analytical approach allows the identification and prioritisation of genes/proteins that are related to mechanisms of disease and may therefore act as biomarkers of CHI for the purposes of diagnosis and disease management (Stevens et al. 2013b; Fig. 3).

The largest cluster identified in the CHI interactome model was associated with cellular signalling with ‘tropomyosin receptor kinase signalling’ as the most significantly related pathway ($P = 5.0 \times 10^{-3}$). Tropomyosin is an integral component of actin filaments within the cytoskeleton. The insulin receptor has been shown to have actions on the structural architecture of the cell (Kadowaki et al. 1985) and polymorphism within regulatory regions of the tropomyosin gene has been associated with cytoskeleton-mediated insulin signalling (Savill et al. 2010). The second largest cluster found in the CHI interactome model was associated with nuclear signalling and the most significantly related pathway was ‘regulation of SMAD2/3 signalling’. SMAD2 has been shown to be associated with the development of pancreatic islets (Goto et al. 2007). These observations derived from the network analysis of the CHI interactome model provide prioritisation of network elements for further research.

**Human growth and short stature in childhood**

Network analysis has been used to demonstrate that the pattern of human growth can be linked to changes in the expression of sets of genes within evolutionarily conserved growth pathways (including WNT, TGFβ, VEGF and the NOTCH signalling pathways) and that this occurs in a tissue-independent manner (Stevens et al. 2013d). This phenomenon has also been shown to occur in rodent models (Finkielstein et al. 2009); however, the pattern of human growth and development is distinct from that of other mammals (Bogin 1999, Thompson & Nelson 2011), and is characterised by rapid but decelerating growth in infancy, an extended childhood phase of slow growth and a pubertal growth spurt before final height is reached, termed the infancy–childhood–puberty growth model (Karlberg et al. 2003). Despite extensive experience over the past 25 years in the clinical management of growth failure in growth disorders, including growth hormone deficiency (GHD) and Turner syndrome (TS), predicting treatment efficacy in individual children remains a significant challenge as both age and stage of childhood development are now known to be important predictors of the outcome of treatments (Ranke et al. 1999, Ranke et al. 2000, Ranke & Lindberg 2011, Patel & Clayton 2012, Ranke 2013).

The possibility of stratifying patients based on genetic variation of response to GH therapy led to the development of the PREDICT study, the aim of which was to assess the genomic profiles of GHD and TS children in response to GH therapy by the analysis of both genetic and transcriptomic data (Clayton et al. 2013, Stevens et al. 2013a). Integrating gene expression and genetic data has provided mechanistic insight into growth response and helped to support the identification of genes with predictive value of growth response. Insulin-like growth factor 1 (IGF1) production over the first month of treatment (Stevens et al. 2013a) and first-year growth responses to recombinant human GH (r-hGH) therapy (Clayton et al. 2013) have both been studied using this approach and several growth-related pathways involved in response to GH therapy have been identified beyond the traditional GH and IGF1 signalling pathways. These included the MAPK pathway, a pathway well known to be involved in growth (Yamauchi et al. 1997, Hanson et al. 2012) but considered to have a less prominent role than the GH pathway. In the MAPK pathway, multiple single nucleotide polymorphisms (SNPs) in genes associated with response to GH therapy were located in both GHD and TS and these genes were shown to be located in network clusters that also contained gene expression changes associated with response to GH therapy. A network analysis using prioritisation by the connectivity of gene expression associated with the carriage of CDK4 alleles linked to IGF1 change over the first month of GH treatment also highlighted the role of the MAPK pathway’s response in GHD children (Stevens et al. 2013a). From these models, proof of principal has been established, that it is possible to identify genetic biomarkers of growth response which could have important predictive potential.

**Obesity and adipose tissue**

Adipose tissue is a host to a number of complex regulatory networks, involving growth, metabolism and intracellular signalling pathways. Several systems-led investigations
Figure 3

Human interactome network analysis of CHI-associated genes. (A) Genes with a known association with CHI were used to infer a network from the BioGRID model of the human interactome (http://thebiogrid.org/). (B) Biological pathway ontology associated with each network module (colour coded); hypergeometric test false discovery rate (FDR) \( \leq 0.001 \). The inferred network was generated using the BioGRID plugin for Cytoscape 3.1.93 to identify the primary interactors of CHI genes from all known physical and genetic evidence. Reproduced from Stevens A, Cosgrove KE, Padidela R, Skae MS, Clayton PE, Banerjee I & Dunne MJ 2013b Can network biology unravel the aetiology of congenital hyperinsulinism? Orphanet Journal of Rare Diseases 8 21–28.
have been performed to assess regulatory networks within adipose cells. Specific lipid metabolism networks have also been discovered, which have environmental and genetic influences and form a highly regulated and integrated homeostatic network primarily designed to maintain energy homeostasis and tissue growth/expansion (Sethi & Vidal-Puig, 2007). These pathways are relevant to obesity and obesity-associated metabolic dysregulation as adipose tissue plasticity and expandability play an important role in energy homeostasis in health and disease. Metabolomics has been applied to the study of lipid molecules including free fatty acids and phospholipids (Han & Gross 2003). This lipidomic approach has been used to investigate the metabolic regulation of adipocytes in obesity in order to further understand the cascade of events progressing from adipocyte expansion to immune cell recruitment and initiation of inflammation (Kamei et al. 2006). Lipidomics utilising network analysis has identified metabolic biomarkers for the response of adipose tissue to treatment. A study on the use of rosiglitazone on lipid metabolism in diabetic mice recently suggested that increased de novo synthesis of the fatty acid palmitoleate may have relevance as a biomarker (Watkins et al. 2002). Palmitoleate has been shown to be synthesised during lipogenesis and its production correlates with improved insulin sensitivity (Cao et al. 2008). In humans, the proportion of palmitoleate was increased in obese but otherwise healthy individuals, suggesting that the production of this fatty acid could reflect an adaptive measure to restore insulin sensitivity. Surprisingly, palmitoleic acid production was not increased in morbidly obese subjects (Pietiläinen et al. 2011). As obesity has a direct association with systemic inflammation and the secondary development of diseases including atherosclerosis, obesity-related insulin resistance and type 2 diabetes, network analysis has also been used to explore functional connectivity between adipocytes and other metabolic tissues; see sections below.

**Obesity and atherosclerosis**

Obesity status is often associated with hypercholesterolaemia that leads in turn to an increased risk of atherosclerosis and cardiovascular diseases (Goldstein & Brown 2001). One of the first key events in the development of atherosclerotic plaque is the uptake of oxidised LDLs by macrophages, which leads to local inflammatory reactions and to the formation of lipid-loaded macrophages, termed ‘foam cells’. LXRα, is a ligand-dependent nuclear receptor that regulates cholesterol homeostasis and reverses cholesterol transport; therefore, it represents an important drug target.

A recent study has used integrative genomic analysis to understand LXRα-dependent functional characteristics and transcriptional regulation pathways (Feldmann et al. 2013). Several SNPs within LXRα-dependent binding sites were found to have differentially expressed target genes involved in the biological processes linked to inflammation, cell growth and apoptosis. To analyse the role of LXRα modulation in atherosclerosis and related diseases, a human macrophage and foam cell network model was used to investigate the actions of a synthetic LXR agonist – T0901317, which has a potential atheroprotective role. By using transcriptional network analysis, highly integrated LXRα ligand-dependent transcriptional clusters were detected which contribute to the reversal of cholesterol efflux and the dampening of inflammation processes in the foam cells (Feldmann et al. 2013). An important implication of these data is that network analysis in this field is helping the understanding of molecular mechanisms governing atherogenesis, which has relevance to therapeutic strategies to prevent atherosclerosis.

**Obesity and type 2 diabetes mellitus**

Current studies based on integrating genotypic data have shown that complex traits such as obesity are modulated by compound networks based on multiple genetic loci and environmental factors. Obesity is also well known to be a chronic tissue inflammation contributing to insulin resistance and secondly to the development of pancreatic β-cell failure and type 2 diabetes mellitus (T2DM). In the USA, the current obesity epidemic is running in parallel with a T2DM epidemic and the underlying pathophysiology are closely linked (Osborn & Olefsky 2012). Several mechanisms proposed to induce β-cell failure have been detected, involved in endoplasmic reticulum stress, oxidative stress, lipotoxicity and glucotoxicity which are mostly mediated by tissue inflammation (Hotamisligil & Erbay 2008, Santiago & Potashkin 2013). A genetics and systems-based approach to the identification of susceptibility loci in obesity-induced diabetes in mice has recently been added to our understanding of the complex genetic interactions contributing to this condition (Davis et al. 2012). In this study, loci contributing to diabetes-related traits were identified, and candidate genes for variations in gene expression were prioritised. Co-expression networks for mice at different ages and clinical traits were then generated and several modules found to be tightly associated with obesity-induced diabetes (Davis et al. 2012). In a separate study, gene and protein interactome models were generated from gene...
expression and metabolic profiling data in obese mice. Network analysis of these models was used to identify differentially activated sub-networks that demonstrated immune system gene networks and inflammatory biomarkers were upregulated in the adipose tissue and that this preceded insulin resistance therefore revealing key links between nutrition and genetic factors (Mori et al. 2010).

The recent introduction of genome wide association studies has also led to the identification of several SNPs associated with diseases, including T2DM. Network analysis has now been used to integrate SNPs known to be associated with T2DM with eQTL and this has led to the discovery of several novel candidate genes for T2DM, involved in the regulation and metabolism of insulin, glucose and lipids (Kang et al. 2012). Biological network analysis in humans has also revealed an association between T2DM and Parkinson’s disease through a network model which places insulin resistance and inflammation central to defects in insulin secretion and defective glucose sensing and impaired insulin signalling, leading to cognitive decline and neurodegeneration (Santiago & Potashkin 2013).

Development of drug-induced insulin resistance: network analysis example of glucocorticoid and GH therapies

The development of insulin resistance and T2DM is not only related to obesity, but can also be iatrogenic. Fleuren et al. (2013) have recently reported on the use of gene networks to identify which genes may play a role in the development of insulin resistance induced by glucocorticoids and identified several candidates, including phosphoenolpyruvate carboxykinase and glucose-6-phosphatase. Treatment with r-hGH can present short-term complications including insulin sensitivity which we have studied in pre-pubertal children with GHD. By network analysis clusters of genes have been identified that underlie metabolic and growth responses to GH and are associated with adipocyte differentiation and cell cycle regulation pathways (Stevens et al. 2013c). These approaches can also be applied to a variety of drug-induced disease networks and may help to identify candidate genetic markers for predicting response to treatment and the development of side effects.

Diabetic kidney disease

Diabetic kidney disease (DKD) is a microvascular complication of diabetes mellitus and is usually accompanied by a progressive loss of kidney function. It has been shown that genetic variations and environmental factors are strongly linked to DKD susceptibility and to the progression of kidney dysfunction (Komorowsky et al. 2012, Bhalla et al. 2013). The improvement of standardised high-throughput technologies of tissue sampling combined with network analysis of tissues from patients with DKD has now led to a far better understanding of the pathophysiology of DKD. Using human renal biopsies, transcriptomics in tubulo-interstitial compartments has recently been combined with hypothesis-driven network pathway analysis to document the involvement of several genes related to inflammation in progressive DKD (Komorowsky et al. 2012). Transcriptional network analysis combined with promoter modelling approaches has also been used to identify a specific NF-κB module in the pathogenesis of DKD, effectively segregating progressive forms of DKD from mild diabetic nephropathy and control subjects (Schmid et al. 2006). Other mechanisms have also been shown to be involved in human DKD, including specific gene expression patterns pointing at pathways involving hypoxic gene regulation, vascular plasticity and JAK/STAT signalling (Lindenmeyer et al. 2007, Berthier et al. 2009).

Conclusions

We have an unprecedented capacity to generate datasets that can document the detailed and complex interactions that take place within a cell. These are being mapped in incredible detail through a combination of technological advancements in experimental techniques, including the ready availability of second- and third-generation sequencing platforms; associated methods (such as RNA-seq and ChIP-seq); other high-throughput ‘omics methods, including quantitative proteomics, interactomics and miRNAomics/MiRomics (Szabo et al. 2013) and the application of computational biology to analyse, visualise and model the systems data. Epigenomics, the whole genome study of DNA methylation and histone markers as indicators of transcriptional activity, is also amenable to network analysis, but this has not yet been implemented in an endocrine context. By contrast epigenomic/network biology integration has made important advances to the area of cancer biology (Huang et al. 2013).

These technologies promise rapid advances both in the scale at which biological research is performed and in the scope of the questions that can be tackled. However, they also offer new challenges, owing to the volume of data generated. In the last few years, a new situation has
Review

A STEVENS and others

Endocrine network analysis

52:1 R90

arisen in biology, whereby, in many cases, the rate-limiting step has shifted from data generation to data analysis. Compounding this problem is the fact that many researchers with traditional experimental training lack the computational skills that ‘big data’ projects require. For this reason, the full potential of these new technologies remains to be fully realised. In this review, we have provided examples of network analysis in endocrinology. Current technology and new approaches to the analysis of omic data can now be combined and network biology can be used to reveal cohesiveness and order from a broad array of datasets representing all aspects of cell function enabling systematic prioritisation and interpretation of the underlying biology of a network. The systems biology approach will undoubtedly enhance understanding of the normal physiology and pathophysiology of different diseases. Interactome modelling of pharmacogenomic responses to therapy will also identify putative new drug targets and possible drug-induced side effects, leading to the introduction of new or repositioned therapeutic agents in clinical practice.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the review reported.

Funding

A S was supported by Merck Serono S.A. (Geneva, Switzerland). C L was supported by the University of Manchester and a European Society for Paediatric Endocrinology (ESPE) Research Fellowship, sponsored by Novo Nordisk A/S.

References


Watkins SM, Reifsnyder PR, Pan HJ, German JB & Leiter EH 2002 Lipid metabolome-wide effects of the PPARγ agonist rosiglitazone.


Received in final form 19 August 2013
Accepted 1 October 2013
Accepted Preprint published online 1 October 2013