Second SICSA Workshop on Biological Networks: Theory and Applications

21. October 2011
Informatics Forum, Room 4.31/33, University of Edinburgh

Workshop programme

09:00-09:30 **Registration and coffee**

09:30-09:50 Ian Overton, MRC Human Genetics Unit: **Integrative Network Analysis of Gene Function and Drug Mode of Action in Methicillin-Resistant Staphylococcus aureus**
09:50-10:10 Derek Gatherer, MRC Centre for Virus Research: **Computing the Incomputable: Rosen's M,R system in Bio-PEPA**
10:10-10:30 J. Michael Herrmann, IPAB, Informatics, Edinburgh: **Effects of Dynamical Feedback and Network Topology on Criticality in Neural Systems**
10:30-10:40 Tools and resources talk 1: Ian Simpson, Douglas Armstrong, IANC, Informatics, Edinburgh: **The SynSys project and Brainwave Discovery Ltd.**

10:40-11:20 **Coffee break**

11:20-11:40 Nils Christian, Physics, Aberdeen: **Structural and functional analysis of large-scale metabolic networks with the method of network expansion**
11:40-12:00 David Willshaw, IANC, Informatics, Edinburgh: **Measuring topographically ordered connectivity patterns in mouse visual system**
12:00-12:20 Pedro Coutinho, MRC Human Genetics Unit: **Discovery and assessment of conserved Pax6 target genes and enhancers**
12:20-12:30 Tools and resources talk 2: Jacek Grzebyta, Rothamsted Research: **Towards a computational representation of host-pathogen interaction networks.**

12:30-14:40 **Lunch, then coffee and posters**

14:40-15:00 Mark van Rossum, IANC, Informatics, Edinburgh: **Synaptic plasticity and memory retention**
15:00-15:20 Wen-Hua Wei, MRC Human Genetics Unit: **Can genome-wide association based interaction analysis provide new insights into biological network underlying obesity?**
15:20-15:30 Tools and resources talk 3: Mark Shovman, Institute of Arts, Media and Computer Games, Abertay Dundee: **Visualising Biological Networks**

15:30-16:00 **Coffee break**

16:00-17:00 Invited Speaker: Vito Latora, University of Catania: **Complex networks in biology**

Organisers: Matthias Hennig, University of Edinburgh, Pierluigi Frisco, Heriot-Watt University

Poster

Gordon Govan, MACS. Heriot-Watt University

Dynamics of neural networks with different motif distributions

In a network a 3-node motif is a group of three nodes. The 3-node motif distribution indicates the percentages of 3-node motifs present in the network. We study the dynamics of Random Recurrent Neural Networks having two specific 3-node motif distributions. We find that one of these motif distributions is more likely to have regular dynamics before any stimulus is applied and that few nodes need to be influenced in order to have a consistent regular dynamics.

Dennis Lee, Electronic and Electrical Engineers, University of Strathclyde

A candidate for a TNF-α controlled stimulation, oscillatory MAPK, p53, NF-kB network with optional UV modulation

Oscillations have been recognised as occurring in the concentration gene expression of a number of elements integral to the dynamic processes of the control and regulation of biological cellular mechanisms, such as the immune system and the development of cancer and chronic inflammatory diseases. Investigation of the same has indicated potential for the detection of drug targets for chemotherapeutic techniques. An examination of the transcriptional modulation of cellular responses to a plurality of complex collective extracellular and intracellular stimuli by protein complexes such as the tumor suppressing gene p53, the signal cascade mechanism brought about by Mitogen-activated protein kinase (MAPK) and the immune-system, growth, differentiation and apoptosis modulating Nuclear factor kappa-light-chain-enhancer of activated B cell (NF-kB) is conducted. A candidate for a tumor necrosis factor-α (TNF-α) controlled stimulation, oscillating MAPK, p53, NF-kB network with optional ultraviolet light (UV) modulation is demonstrated, providing a platform for future biological and engineering analysis with a view to better understanding the oscillatory nature of the components and enabling their use in drug design and chemotherapeutic techniques.

Nick Schurch, College of Life Sciences, University of Dundee

Systems Biology of PTEN mutants

Carron Shankland, Computing Science and Mathematics, University of Stirling

Scaling from Individuals to Populations: A Process Algebra Approach to Infectious Diseases

Given information about individual behaviour observed in the field, how can population behaviour be extracted? The theoretical problem of how to change between scales is relevant to many fields. As modellers we do this in our heads all the time, but a rigorous method is preferable. Our work is focussed around a rigorous method for changing scales, based on process algebra. We apply this in particular to models of infectious disease, such as measles, and bubonic plague.

Alexandros Xenos, Computer Science, Heriot Watt University

Evolution program for network model

Recently Pierluigi Frisco proposed a new network model. In this model nodes are represented as words over an alphabet and edges between nodes are added according to some distance measures. A program that enables the user to create networks that follow this model has also been implemented. In order to achieve networks with certain topologies, the user has to tune various properties which are provided as input to the algorithm. In this research an evolutionary algorithm has been developed, which automates the properties' tuning procedure. This procedure constitutes a multiobjective optimization problem. In general, evolutionary algorithms have been widely used for solving optimization problems and show significantly good performance in multiobjective optimization problems. Among other parameters,the implemented algorithm requires from the user three network characteristics that need to be reached. These characteristics are the average degree, the average path length and the average cluster coefficient. These values form the objectives that the algorithm tries to optimize. Once the algorithm terminates,the user receives as output, those network's properties that were found to be optimal for the certain input. The program was tested by running simulations of the Escherichia coli protein-protein interaction network. The test – runs were conducted on a computer cluster and the results improved previous attempts to replicate this network.

Talks

Nils Christian
Physics, University of Aberdeen

Structural and functional analysis of large-scale metabolic networks with the method of network expansion

We present the method of network expansion as an approach to analyse genome-scale metabolic networks. The algorithm of network expansion defines the 'scope' of a set of available metabolites, called the seed, as the set of metabolites which the network is in principle capable to produce. While this method, in contrast to constraint-based approaches, cannot produce quantitative predictions, it is directly applicable to networks derived from databases without the need for extensive curation. Further, the calculation of a scope is extremely simple and computationally cheap. We show how this advantage can be exploited to investigate and compare properties of hundreds of organism-specific network. We show how this method can be applied to infer minimal nutrient combinations which are required to produce all essential precursors for biomass production. Our results indicate that it is possible to distinguish between specialist and generalist species based on their metabolic network structure alone. As a second application, we demonstrate how gaps in metabolic networks can be identified and filled. For this, we identify minimal sets of reactions which need to be added to a draft network derived from a genome sequence such that the network is consistent with experimental data. Consistency is assumed if the network can produce all experimentally observed metabolites from the applied nutrient medium.

Pedro Coutinho
MRC Human Genetics Unit

Discovery and assessment of conserved Pax6 target genes and enhancers

The characterisation of transcriptional networks (TNs) is essential for understanding complex biological phenomena such as development, disease and evolution. In this study, we have designed and implemented a procedure that combines in silico target screens with zebrafish and mouse validation, in order to identify cis-elements and genes directly regulated by Pax6. We chose Pax6 as the paradigm because of its crucial roles in organogenesis and human disease. We identified over 600 putative Pax6 binding sites and more than 200 predicted direct target genes, conserved in evolution from zebrafish to human and to mouse. This was accomplished using Hidden Markov Models (HMMs) generated from experimentally validated Pax6 binding sites. A small sample of genes, expressed in the neural lineage, was chosen from the predictions for RNA in situ validation using zebrafish and mouse models. Validation of DNA binding to some predicted cis-elements was also carried out using chromatin immunoprecipitation (ChIP) and zebrafish reporter transgenic studies. The results show that this combined procedure is a highly efficient tool to investigate the architecture of TNs and constitutes a useful complementary resource to ChIP and expression datasets because of its inherent spatiotemporal independence. We have identified several novel direct targets, including some putative disease genes, among them Foxp2; these will allow further dissection of Pax6 function in development and disease.

Derek Gatherer
MRC Centre for Virus Research, University of Glasgow

Computing the Incomputable: Rosen's M,R system in Bio-PEPA

Systems Biology has a mirror image it often fails to recognise: its name is Relational Biology. Relational biologists are interested in the same things as systems biologists, and study them using the same theoretical techniques. However, they do not go to Systems Biology conferences, and have a fundamentally different perspective on biological complexity. To a certain extent they maintain that Systems Biology is a wild goose chase. Their central and most challenging proposition for systems biologists is that there exists a certain network topology, Rosen's M,R system [1, 2], which cannot be computed, owing to its self-referential structure. From this they extrapolate to the more general conclusion that no biological network of any genuinely interesting size can be computed. This talk will outline the arguments used by Rosen and his followers to reach this conclusion and will describe our recent attempts to falsify this hypothesis on an empirical level, by describing the simplest possible M,R system in Bio-PEPA [3] with Michaelis-Menten kinetics and using the Bio-PEPA Eclipse plug-into obtain simulations of the system. The fact that our simulation runs at all would appear to be a prima facie falsification of the relational biologists' stance on M,R systems. However, there are several assumptions involved in our model that may prevent us from fully satisfying Rosen's special definition of simulation. Nevertheless, Bio-PEPA provides at the very least a good approximation of a solution to what has until now been presented by relational biologists as an intractable problem.

1.Rosen R: Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. New York: Columbia University Press; 1991.
2.Cardenas ML, Letelier JC, Gutierrez C, Cornish-Bowden A, Soto-Andrade J: Closure to efficient causation, computability and artificial life. J Theor Biol 2010, 263(1):79-92.
3.Ciocchetta F, Hillston J: Bio-PEPA: a Framework for the Modelling and Analysis of Biochemical Networks. Theoretical Computer Science 2009, 410(33-34):3065-3084.

Jacek Grzebyta, A. Splendiani, M. Urban, K.E. Hammond-Kosack, C.J. Rawlings, M. Saqi
Rothamsted Research

Towards a computational representation of host-pathogen interaction networks.

PHI-base, the Pathogen Host Interaction database (http://www.phibase.org/) is an open access internet resource which provides information on pathogenicity, virulence and effector genes from different pathogens, where the contribution of the genes to pathogenicity has been experimentally tested. It is based on manually curated information retrieved from the peer-reviewed scientific literature, and went on-line in 2005. PHI-base is a database which brings together genetic molecular and phenomic information and potentially has application to many different kinds of studies. Within PHI-base version 4.0, information is represented in a network-based model which is exposed to users through a multi-layer user interface. All information is available in a user readable way as web pages. At a lower level, these pages are enriched by semantic markup (RDFa) which describe each feature in terms of ontology entities. All information is then accessible to other databases/computers through an RDF access point using SPARQL queries. PHI-base represents a complex knowledge domain which connects different types of information pertaining to different scales of observations of a biological system. Our RDF-based representation significantly improves the ability to easily connect across data sources and to characterise in a semantic way the information provided.

Acknowledgements Rothamsted Research receives funding the BBSRC. In addition, specific improvements to PHI-base are currently supported by a BBSRC BBR award called PhytoPath: an Integrated resource for comparative phytopathogen genomics (BB/I000488/1).

J. Michael Herrmann
Informatics, University of Edinburgh
Effects of Dynamical Feedback and Network Topology on Criticality in Neural Systems
After a brief description of a self-organised critical neural network with dynamical couplings, we will ask how the network topology, synaptic homeostasis, neural leakage and long-term learning affect the critical behaviour in this network. We demonstrate that for several typical topologies (random, small-world, scale-free) critical avalanches are typical, while more structured topologies, such as those introduced by structured memory patterns, appear to be adverse to criticality which can, however, this can be alleviated by homeostatic regulation and STDP learning. In the light of these studies we will discuss the benefits and drawbacks of short-term plasticity as a mechanism for the generation of self-organised criticality in neural networks.

Ian Overton, MRC Human Genetics Unit
Integrative Network Analysis of Gene Function and Drug Mode of Action in Methicillin-Resistant Staphylococcus aureus
Staphylococcus aureus is a major pathogen and drug-resistant strains continue to emerge; development of novel treatments is therefore important. Antimicrobial peptides offer a source of potential novel antibiotics to combat resistant bacteria such as Methicillin-Resistant Staphylococcus aureus (MRSA). A promising antimicrobial peptide is ranalexin, which has potent activity against Gram-positive bacteria, and particularly S. aureus. Understanding mode of action is a key component of drug discovery and network biology approaches enable a global, integrated view of microbial physiology, including mechanisms of antibiotic killing.
A Bayesian logistic regression approach was taken to generate a systems-wide network of MRSA gene function. This provided a framework to integrate protein and RNA profiles of ranalexin stress, enabling study of drug resistance and mode of action. Ranalexin response signatures revealed multiple killing mechanisms, including cell wall activity. These cell wall effects were supported by gene disruption and osmotic fragility experiments. Furthermore, twenty-two novel virulence factors were inferred, while the VraRS two-component system and PhoU-mediated persister formation were implicated in MRSA tolerance to cationic antimicrobial peptides. This work demonstrates a powerful integrative approach to study drug resistance and mode of action, informing the development of novel therapeutic strategies against Staphylococcus aureus.

Mark van Rossum
Informatics, University of Edinburgh
Synaptic plasticity and memory retention
The strength of the synapses in the brain are presumably continuously subject to increases and decreases as the result of ongoing learning processes. This realization allows one to approximate the synaptic weight evolution as a stochastic process. This has been used to find fundamental limits of storage. Recently we introduced a synaptic information capacity measure based on Shannon information (Barrett and van Rossum '08). We use this to find the optimal weight dependent learning rules. We find that soft-bound learning rules are better than hard bound rules, although the improvement is quite small. Furthermore, we show how feedforward inhibition further increases storage.

Mark Shovman
Institute of Arts, Media and Computer Games, University of Abertay Dundee
Visualising Biological Networks
Visualisation of biological network data is a common tool in generation of novel insights as well as in communication of these insights to different target audiences. There are many options for visualisation of a network, not all of them equally efficient for a given task. We present various options of visualising biological networks, and compare them in terms of their efficiency in generating and conveying relevant information.

Ian Simpson, Douglas Armstrong
Informatics, University of Edinburgh

The SynSys project and Brainwave Discovery Ltd.

We have adopted a number of network modelling approaches to describe the molecular complexity of the synaptic proteome ranging from static protein interaction networks (PINs) to dynamical simulations of protein complexes including post-translational modifications such as phosphorylation. In this talk I will focus on one of the approaches we have added to our PIN construction methodology, that of Interolog prediction. Interologs are putative protein-protein interactions (PPI) that are inferred from known interactions involving orthologues of our network proteins. We mine existing PPI resources in dozens of species and project these interactions back to our species of interest. Interologs have proved useful in augmenting sparse networks and consolidating both existing and novel PPI data into PINs. We have developed a Perl package Bio::Homology::InterologWalk which uses the Ensembl-Compara database in concert with PSIQUIC enabled PPI web services to automatically retrieve and score Intrerologs for PIN generation. I will illustrate applications of Interologs for novel data validation and PIN augmentation and demonstrate their utility during predictive modelling.

Wen-Hua Wei
MRC Human Genetics Unit

Can genome-wide association based interaction analysis provide new insights into biological network underlying obesity?

We are developing a platform for high-throughput genome-wide association (GWA) based interaction analysis for complex disorders or quantitative traits. Such an analysis will produce signals of gene-gene and gene-environment interactions which can be used to construct interaction network and provide new insights into the genetic aetiology of the trait studied. We perform such analyses on large GWA datasets prepared for studying obesity-related traits. I will present some of the results to address two questions: a) what values can GWA interaction analysis add to our understanding of biological network underlying obesity and b) what helps we need to maximise the values.

David Willshaw
Informatics, University of Edinburgh

Measuring topographically ordered connectivity patterns in mouse visual system

It is generally thought that a combination of molecular guidance cues and electrical activity results in the formation of the specific connections between neurons which are essential for the proper functioning of the nervous system. A paradigm example for how such neural networks are formed is the topographically ordered map of the vertebrate retina onto the superior colliculus or optic tectum. A challenge has been to find sufficiently precise information about the order in these neural maps and to develop methods for their analysis. I have been examining Fourier-based imaging data taken from mouse colliculus in both wild type and mutant animals in which genes specifying some putative guidance molecules had been knocked out. A detailed analysis of the data reveals that in wild types, points on the colliculus that are 50 microns apart project correctly in the map. Knockout of the genes for the ephrinA ligands, that are thought to specify order along one dimension of the map, still results in a map with a high degree of order. This points to a new interpretation of the role of these genes in forming maps of connections.