ESF Exploratory Workshop on Computational disease Modeling

A joint workshop ESF-CRM

Dates: September 24 to 26, 2008

Place:

Institut d'Estudis Catalans (IEC)
C. del Carme, 47
08001 Barcelona
Tel.: + 34 93 270 16 20
http://www.iec.cat

Organiser: Centre de Recerca Matemàtica & IDIBAPS
Facultat de Ciències, UAB
08193 Bellaterra
Tel.: + 34 93 581 40 86
http://www.crm.cat
Scientific organiser: Albert Compte, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona

SCIENTIFIC REPORT

Executive summary (MAX 2 pages)

Biomedical research is changing due to the rapid accumulation of experimental data at an unprecedented scale revealing increasing degrees of complexity of biological processes. Life Sciences are facing a transition from a descriptive approach towards uncovering the principles underlying biological function and dysfunction. This development entails a major effort on computationally driven data-integration to reveal principles of cellular networks, modules, cells, organs and their interactions across several spatial and temporal scales.

There are two conceptual traditions in computational modelling in Biology. The bottom-up approach emphasizes complex intracellular molecular models and it is well represented within the systems biology community. On the other hand, the physics-inspired top-down modelling strategy selects essential features of relevance for the phenomena of interest and combines the available data in a model of modest complexity.

The workshop (http://www.crm.cat/dismod), supported by the European Science Foundation and the Center for Mathematical Research of Catalonia (CRM, http://www.crm.cat) examined the challenges that computational modelling faces in contributing to the understanding and treatment of complex multi-factorial diseases. Computational approaches might be effective because, although the pathophysiology of a given disease might involve partially described mechanisms at very different levels, from molecular interactions to whole system dysfunction, fundamental understanding of the disease and guidance of experiments might be at reach without having all molecular networks characterized.

The workshop in Barcelona (Sept. 24-26, 2008) brought together modellers, experimentalists and clinicians to discuss how multi-factorial human diseases

(including multiple sclerosis, cancer, cardiovascular disease, diabetes, sepsis, allergy, schizophrenia and addiction) can be modelled given the current available knowledge and data. On the modelling side experts covered molecular network modelling, computational neuroscience, PKPD modelling, hierarchical modelling and agent-based modelling.

Participants at the meeting agreed on two general conclusions. On the one hand, we identified the critical importance of developing analytical tools for dealing with model and parameter uncertainty. On the other hand, the development of hierarchical models spanning several scales beyond intracellular molecular networks was detected as a major objective, which would allow the propagation of experimental constraints between the different levels. This contrasts with current focus within the systems biology community on complex molecular modelling. Through the workshop it became clear that the different scientific modelling cultures (from computational neuroscience, theory, agent based modelling and molecular approaches) would benefit from more intense cross-talk on shared theoretical issues in order to make progress on clinically relevant problems.

Scientific content (Min 1 page)

A necessary condition for successful modelling of diseases is to recognize the need for standards on data-collection and storage, interoperable representation as well as computational tools and standards enabling pattern and network analysis and modelling. There are several important initiatives in this direction such as the FP7 program ELIXIR (http://www.elixir-europe.org) on providing sustainable storage infrastructure for biomedical data in Europe. Similar initiatives are in progress in the US and Asia. Yet these efforts are not sufficient, as the predictive understanding of complex diseases requires computational modelling and representation of these data. However, there are deep and unsolved conceptual issues regarding how to do this, which is reflected in the choice of research questions, data and modeling approach. By focusing on the issue of complex multi-factorial disease modeling we were able to uncover a few core problems that have not been sufficiently recognized, but must be addressed when trying to leverage the available and growing amounts of relevant biological information.

Model selection and parameter uncertainty

Across different application areas a key question concerned the handling of model uncertainty. This refers to the fact that for any biological system there is a verbal model that is most likely partially erroneous and definitely incomplete. Computational model selection has to cope systematically with the fact that there could be additional relevant interactions and components beyond the verbal model. Moreover, since the kinetics is as a rule insufficiently characterized there is a serious indetermination of parameters, even for mechanisms contemplated in the verbal model. Hence, biological models, unlike models describing physical laws, are as a rule highly over-

parameterized with respect to the available data. Thus, large regions of the parameter space contain acceptable parameters, which describe the available data equally well from a statistical point-of-view.

A successful strategy in computational neuroscience has been to identify minimal models that adequately describe and predict the biology. This comes at the price of selecting a too narrow and incomplete model. This approach is justified if there exists good knowledge on the physiological mechanisms involved in a given condition. In situations where biology is less well characterized, then one must consider and compare several plausible model structures. An alternative approach, recently employed within the systems biology and computational neuroscience field, is to search for parameter dimensions that are important for model performance, as opposed to individual parameter sets. There seemed to be general agreement that the concept of model ensembles therefore represents a promising approach. The process of characterizing parameter values must therefore be applied to each model structure. The resulting ensemble is the collection of model structures and their associated probabilistic parameter distributions. Stochastic search of parameter space using a variety of techniques (MCMC-based) seems to be state of the art. Multi-start convex optimizations or particle swarm optimization (PSO) algorithms locate a potentially large number of local minima of a user-defined, biologically relevant objective function, but do not offer assurance of adequate coverage of parameter space, nor do they have the asymptotic property of resulting in a probability density function in parameter space. It is not yet clear under what conditions a optimization is most useful. Furthermore, once model parameters have been selected for competing model structures, there are no clear ways to combine these models to create the ultimate ensemble used to formulate predictions: choosing the “best” structure or weighing competing structures based on their relative

fitness. Model-guided experimental design appears a promising venue as a possible means of clarifying model structure.

Model selection is therefore important to prevent overfitting and to distinguish between competing explanations. Bayesian model selection is becoming standard in some areas of computational neuroimaging (e.g. Penny et al. 2004, NeuroImage 22, 1157-1172; Stephan et al. 2007, NeuroImage 38, 387-401) and may be useful for many areas of systems biology. Moreover, interpreting parameter estimates is difficult because (i) they are conditional on the model chosen, and (ii) they may exhibit interdependencies with other parameters in the model. There are good reasons to believe that such interdependencies are unavoidable (and to some degree even desirable, to increase robustness against lesions) in biological systems (see also Gutenkunst PloS Comput Biol, 3:e189, 2007). There is also a bias towards mechanistic and molecular models in systems biology. However, whenever possible, models should not only be mechanistic, but also allow for experimental validation of the mechanisms they propose. This means that their components should be at a level of description which allows for experimental perturbations, given the biological techniques we currently have available. More generally, a mechanistic model is not very helpful unless we have experimental means to assess its predictive validity (over and beyond its face validity and construct validity; these different types of validity are not always distinguished, although the distinction is very important).

Hierarchical models

The second major theme concerned the development of hierarchical models, spanning several levels of biological organisation such as intracellular networks, cell-to-cell interactions up to interacting organs or the whole body.

Much attention was devoted to organ-level models or diseases (multiple

sclerosis, allergic rhinitis and sepsis), and on the Virtual Physiological Human Project in particular. Here methodologies like co-simulation, which allows for parallel simulation at different time-scales in different modules and levels were presented and discussed. Likewise, tools and modelling environments for handling of such integrative models including different types of equations (ODEs, PDEs, SDEs etc) were discussed and evaluated. It was evident that much work on the systems biology community targets intracellular networks whereas computational neuroscience tends to perform top-down modelling. Presentations on schizophrenia and the presentation of nicotine addition focused on the use of very simple, top-down models to explain complex phenomena and offered useful predictions. By characterizing systems properties these constraints can propagate to lower levels and therefore reduce the number of consistent solutions for a given data-set. This aspect has not been sufficiently recognized when trying to model large-scale high-throughput data. On the other hand top-down modelling runs the risk of having a too narrow model selection. There appears to be a lack of theory for how to integrate model selection with constraint propagation across several layers of biological organization, for which only sparse data is available. This may prove to be a useful path towards modelling complex diseases even though data is incomplete. One useful practical first approximation may be the notion of disease networks - the interaction between different diseases and their respective molecular components. This approach may provide both bottom-up and top-down constraints for understanding complex diseases. There are multiple well-known examples for these, including obesity-diabetes, Gaucher disease-Parkinson disease, etc. At the same time, alteration of the physiome by a given disease can also lower the chance of developing another disease, e.g., sickle cell dis. (Hgb S) and malaria infection.

Assessment of the results - the future

An outcome from the workshop was the establishment of new contacts and generation of new ideas when different fields of computational biology met together with clinical researchers. This was an important accomplishment in itself. Overall, seeking broader expertise towards theoretical advances seems a high priority to make progress on model selection and hierarchical modelling for complex disease research. There is a need for a scientific forum where control theory, physics and applied mathematics can stimulate method development across different areas of computational biology. The participants agreed on finding an appropriate funding opportunity for such an initiative. More theory can therefore potentially nurture current modelling efforts to make use of “simpler” models and therefore be more useful to the clinic. Here, several participants expressed the belief that the notion of groups of patients with common patterns was a useful concept and probably a more accessible target than truly understanding individualized dynamics.

FINAL PROGRAMME

Wednesday, September 24

14:00-14:30	Albert Compte	Welcome and introductory remarks. Brief presentation of ESF activities.
14:30-15:10	Mikael Benson	What do clinical researchers want from modelling and systems biology?
15:10-15:50	Jesper Tegnér	Bridging the gap – challenges and possibilities
15:50-16:20	Coffee Break	
16:20-17:00	Randall Thomas (replacing Jean-Pierre Boissel)	Systemic Physiopathology: why and how?
17:00-17:40	Gunnar Cedersund	Progress and challenges in systems biology studies of type II diabetes
17:40-18:20	Discussion: General objectives and challenges – Round table	
18:20-19:00	Cheese and wine reception	
20:00	Dinner	*Detailed information on the social activities is given below*

Thursday, September 25

9:00-9:40	Marta Cascante	A Systems Biology approach to multifactorial diseases
9:40-10:20	Zoltan Oltvai	Disease networks
10:20-10:50	Coffee Break	
10:50-11:30	Fazoil Ataullakhanov	Mathematical modeling of the metabolism and viability of red blood cells as a tool to study the mechanisms of hereditary hemolytic anemia.
11:30-12:10	Jörg Stelling	Robustness and intervention in cellular networks
12:10-12:40	Discussion: Focus on molecular pathways	
12:40-14:30	Lunch	
14:30-15:10	Charles Auffray	Combining transcriptome analysis, functional annotation and systemic modeling to decipher the cellular states of innate tumor drug responses
15:10-15:50	Gary An	An Agent-Based framework for Integrative Dynamic Representation of Biomedical Knowledge: Towards an Ecological Paradigm for Collaborative Research

15:50-16:20	Coffee Break	
16:20-17:00	Jesper Tegnér	The role of multiple organs in complex metabolic diseases - towards genetics of gene expression
17:00-17:40	Randall Thomas	Collaborative multi-scale modeling of blood pressure regulation and fluid homeostasis
17:40-18:10	Discussion: Multilevel integration	
19:00-21:00	Guided tour around the Barcelona Gothic Quarter	
21:00	Dinner (aproximate time as dinner will begin when participants arrive from their tour)	

Friday, September 26

9:00-9:40	Pablo Villoslada	Computational modelling of the immune system for understanding autoimmune diseases and immunotherapies
9:40-10:20	Boris Gutkin	Computational Models of nicotine Addicition: from circuit dynamics to behavior
10:20-10:50	Coffee Break	
10:50-11:30	Klaas Enno Stephan	Towards neurocomputational models for investigating and diagnosing psychiatric diseases
11:30-12:10	Albert Compte	A systemic modeling approach to the pathophysiology of atherosclerosis
12:10-12:40	Discussion: Focus on cellular population modeling	
12:40-14:30	Lunch	
14:30-15:10	Gilles Clermont	Biological variability and in silico design of interventional clinical trials
15:10-15:50	Gustavo Deco	Computational Neuropsychiatry: Neuronal Fluctuations, Dynamics and Schizophrenia
15:50-16:20	Coffee Break	
16:20-17:00	Mats Gyllenberg	Evolutionary aspects of human diseases
17:00-17:30	Discussion: Focus on systemic and abstract modeling	
17:30-18:30	Final discussions and follow up activities	
21:00	Dinner	

Statistical information

Age structure: mean 45.93 yo, standard deviation 7.27 yo, range: 30-62 yo

Gender repartition: 16 male, 1 female
Countries of origin: Spain (4), Sweden (3), France (3), USA (3), Switzerland (2), Russia (1), Finland (1)

Final list of participants

Name	Age	Institution Country	University / Institution
An, Gary	43	USA	Northwestern University Feinberg School of Medicine
Ataullakhanov, Fazoil I.	62	Russia	National Hematology Research Center
Auffray, Charles	57	France	CNRS
Benson, Mikael	54	Sweden	Göteborgs University
Cascante, Marta	48	Spain	Universitat de Barcelona
Cedersund, Gunnar	30	Sweden	Fraunhofer-Chalmers Centre for Industrial Mathematics
Clermont, Gilles		USA	University of Pittsburgh
Compte, Albert	37	Spain	Hospital Clínic - IDIBAPS
Deco, Gustavo	46	Spain	Universitat Pompeu Fabra
Gutkin, Boris S.	42	France	Départment d'Etudes Cognitives (DEC) at Ecole Normale Supérieure (ENS)
Gyllenberg, Mats	52	Finland	University of Helsinki
Oltvai, Zoltan N.		USA	SCAIF
Stelling, Jörg	39	Switzerland	ETH Zürich
Stephan, Klaas Enno	35	Switzerland	University of Zürich
Tegner, Jesper	46	Sweden	Linköping University
Thomas, Randall S.	57	France	IBISC
Villoslada, Pablo	41	Spain	Universidad de Navarra