Systems Biology: Cellular Weather Forecasting?

I would like to first outline the background for our research, then give a definition of ***systems biology***, before concluding with some remarks on interdisciplinary research initiatives such as the one we are here for today.

For me, the most amazing example for the wonderful complexity and beauty in nature is the life cycle and total metamorphosis of the butterfly. During its morphological development the organism undergoes a dramatic transformation. Nowadays we argue that the information to enable this fascinating process is rooted in the ***genome*** of the organism; whereby the genome is understood as the entirety of the genetic information, encoded in a physical structure known as the DNA molecule.

With the total metamorphosis of the butterfly we then have an example for two organisms in one. This example also demonstrates the only relative difference between living things that otherwise appear and behave quite differently. I will return to this question of differences later on. What the life sciences are concerned with are ways to explain such observations at the ***physiological and phentotypic level***, related to the appearance and function of an organism, ***but*** explained from the ***cell-*** and ***molecular level*** upwards. I did deliberately avoid the concept of a ***gene*** as a causal agent, since I believe that this concept can mislead us.

Understanding the natural world, including butterflies, is not profitable and thus the declared aim of our research has to be human health. Considering then a human body, it consists of ***organs***, which in turn are composed of ***cells***. For example, in case of the liver, it consists of some three-hundred billion cells. Zooming in from the organ- to the cell-level, we are covering an enormous scale of magnitude, from a human cell with a diameter of only ten micrometer to a human body consisting of 10^{13} cells. It is from time to time helpful to remind ourselves of the intellectual cosmos we are trying to travel here: from dynamic interactions of molecules at the nanometre level to the physiology of an entire organism, that can be as large as 2.11m…

The earth has about six billion inhabitants; a single human body consist of ten-thousand billion cells. Have we not all tried to make sense of the conflicts happening around the world?

While different kinds cells, say in the brain or the liver, have distinct structures and functions, all cells contain a complete copy of the genome. The continuing excitement in the life sciences, described as the ***"post-genome era"***, derives from the fact that we have now available technologies to read the genome sequence and generally to characterise the components of the cell.

The “Omics-family”, that is, scientists working in the fields of genomics, transcriptomics, proteomics, supported by physicists and engineers, providing the technology, together with bioinformaticians, helping to manage the information generated, have been able to identify, characterise, catalogue many of the components that are involved in developmental and disease mechanisms.

The initial hope was then that there would be a simple “gene/disease” relationship, and to this day we find newspaper headlines that report “Gene *x* for disease *y* has been discovered”. Upon reading the text, the headline is usually refined by a statement that “the discovery is taking us *a step* towards an understanding of the disease”. The truth is, that in most cases where we have suspected a simple answer we realise that we are dealing with ***complex, regulated networks of interacting dynamic processes***. I say ***“in most cases”*** since biology is so difficult because for every rule we seem to discover, there are a multitude of exceptions. And so there are indeed cases where a single erroneous gene or even a single mutated base pair can have fatal consequences for the health or fate of the organism. We know this situation from our experience at University. Most things are very robust to change and yet in some cases a single person can cause tremendous damage for colleagues and the university as a whole.

Having discussed the biological background and motivation, I like to move now on to a definition of systems biology. The word “systems” in ‘systems biology’ refers to the ‘systems sciences’, ‘systems and control theory’; which in practical terms means ***mathematical modelling*** and ***simulation***.

Regardless of whether you are a mathematician or a biologist, a ***system*** is considered as a relation on variables/indicators/items defined in set theoretic terms,

$$S \subset I_1 \otimes \ldots \otimes I_n$$

In the same way, a ***complex system*** is a relation on systems/subsystems, i.e.,

$$S \subset S_1 \otimes \ldots \otimes S_n$$

such that there is a distinct behaviour of the complex system while the integrity of the subsystems is preserved.

The rise of systems biology stems from the realisation that molecular characterisation and cataloguing an ever increasing pool of information will necessarily be limited as a process to gain a better understanding of biological function and disease mechanisms. The central dogma of systems biology is, that it is not primarily the information in the genome but the dynamic interactions of the molecules generated from this information that gives rise to biological function. An appropriate definition of systems biology is then the following:

Systems biology investigates inter- and intra-cellular dynamic processes, using signal- and systems-oriented approaches.

The key question in systems biology is then “How do cells, genes and their products ***act*** and ***react*** in response to environmental changes?”. The message is then that, instead of cataloguing genes as causal agents for some protein, function, or phenotype, we should relate observations to ***sequences of events***: it is ***system dynamics*** that gives rise to biological function! Let me explain this emphasis on dynamics by demonstrating the limits of comparative or mining approaches.

We all read about the close similarity of the human genome with that of much “simpler” organisms. For instance, our genome apparently shares a 75% similarity with the nematode worm and is 98% similar to that of a chimpanzee. One of our own projects is concerned with *Mycobacterium bovis* in cattle, and the human form *Mycobacterium tuberculosis*, where the similarity at DNA sequence level is estimated at 99.9%. How is it possible that the entirety of the genetic material is nearly identical and yet at the phenotypic or physiological level, the organisms appear and function very differently? This apparent contradiction can be solved by a simple Gedankenexperiment. Imagine two organisms that consist of only eight genes, each of which can only either be switched “on” or “off”. A comparative study of these two systems would be limited to the discovery of 2^8 or 256 different pattern. If we however consider gene expression as a dynamic process, allowing for only three time points, there are already 256^3, that is, more than 16 million ways to encode information.

This then demonstrates the importance of signal- and systems-oriented approaches to the study of gene expression and signal transduction. One wonders why it is only now that systems biology receives attention? There are two main reasons: only now we are beginning to get our hands on technology that is going to allow us to ***quantify*** gene activity and protein concentrations in time course experiments. The second reason is that we are entrenched in the static, comparative approaches of molecular characterisation and data mining: we need to change the way we think about these processes and how we design experiments. ***Systems theory is not a collection of facts, but a way of thinking***.

For reasons I do not have time for to explain them, the systems biological approach is more demanding on the quality, accuracy and richness of data sets. While we all associate ***bioinformatics*** with ***vast amounts of data***, in ***systems biology*** it is a ***lack of quantitative data*** that will continue to pose a major technical challenge for the years to come. However, since the technological progress has been tremendous in recent years, we can be optimistic in that respect. Probably the greatest hurdle to the success of systems biology is of another, more human, nature.

There are now in most countries funding programmes focussing on systems biology, and in many countries entirely new buildings are built to support this kind of interdisciplinary research. One may hope that this building activity is simply an expression of the importance of the discipline, but the truth is that it is also recognition of the difficulties we face. Interdisciplinarity can take various forms,

and although bioinformatics would by definition fall into this category, a proper measure for ***interdisciplinarity*** is the ***dependence*** of the careers of researchers involved. In bioinformatics, the relationship between the biologist and the bioinformatician is more often of a supporting role than one on equal terms. On the other hand, in systems biology, there are already several examples, where the advances of biological understanding were possible only due to mathematical modelling. In future there will be more situations where biologist can only succeed if they are supported by a statistical and mathematical analysis, while at the same time the modeller or data analyst is unable to devise models without understanding the context in which the data are generated. This research is then of high risk to the career of the scientists involved: it does not only depend on the availability of technologies and tools, but on whether scientists get on; whether ***both*** are prepared to go through the lengthy process of learning about each others' work before any recognisable results are achieved.

The importance of this truly interdisciplinary research is widely recognised, and the best way to promote such synergy is to create opportunities for scientists of the engineering, physical and mathematical sciences to meet biological and medical researchers. This is then also the main reason behind the buildings and centres that are created around the world.

My worry is that Germany is going to struggle with the changing nature of research in these disciplines. The reasons for my concerns are not the lack of money for new buildings, but rather that the attitudes that our university structures create in us may hinder this process. My suggestion is therefore for our regional government, Universities, faculties and departments, to invest in a doctoral ***training*** programme. I have just returned from a presentation at an interdisciplinary research school in Göteborg, Sweden. There, professors ***or group leaders*** can apply, together with a partner from another discipline, for a project that supports the doctoral training of two researchers from different disciplines. I am convinced that we are more likely to succeed if we bring together young researchers from different disciplines, rather than funding a talking shop for professors. I should add that one of the conditions for the Swedish group leaders to be allowed to participate is that they are contributing lectures to an interdisciplinary course.

For the reasons I described, any initiative of the Universities of Greifswald and Rostock to create a platform for interdisciplinary research is very important. I sincerely hope that this platform is not becoming a stage for professors to perform their "give me the money act" but that younger actors are encouraged to play a role.

Further information on our research activities can be found on our website at www.sbi.uni-rostock.de.

Olaf Wolkenhauer, Rostock 27 November 2003