

## Back to Darwin?

At a recent meeting on genomics in Berlin, one of the speakers said, 'In ten years time, those with a hypothesis-based approach to science will be equivalent to those who believe in the Flat Earth theory.' Strong sentiments! But it was said with a conviction that chilled those among us who come from the 'old-fashioned' style of research. Indeed, the Human Genome Project is undoubtedly a huge success and it has set an example of how brute-force strategies in the life sciences can succeed. This is a message that is understood by all science policy makers. Ultimately, the Human Genome Project is the result of an industrial approach to amassing data. Cleverly coordinated robots and machines churned out sequences at a predictable rate. With a fixed charge per sequenced base, the total cost and progress towards the defined goal can be predicted and monitored. How much less amenable to project managers is the traditional way of performing research, where, by definition, the result is often a surprise. With the success of platform technologies, it will become increasingly difficult to convince assessment boards seeking simple management solutions that there is a continued need for hypothesis-driven research.

All over the world, plans are afoot to establish platform technology units that will carry out 'blazed earth' assaults on the jungles of information that to date have only been hacked at by the machetes of individual research projects. Proteomics will tell us which proteins are expressed in each cell under all conditions. Transcriptomics will provide information on every RNA that is transcribed in the cell. Functional genomics will plot the passage from the DNA through RNA to protein. Structural genomics will provide the structure of each domain. Post-translational genomics will plot the dynamic changes in the modification of proteins. Lipidomics, glycosylomics—and who knows what else—will all arrive in the near future on the desks of committees that decide on

grant proposals. And each of these approaches will require an enormous level of funding and coordination.

The need for these data is undeniable and acquiring them should be supported. However, the question will have to be asked, 'Who will be able to use this new encyclopaedia of life?' The philosophy of those who practice platform technologies is that having all the facts is sufficient. Clearly it is not. Any mission description for biological research must include the ambition to understand nature itself. Databases alone do not provide such insights—no more than a random collection of a sonnet's words will enrich the mind of its reader.

Ironically, the modern fashion of systematically collecting data without a theoretical framework brings us back to the days of Darwin. He collected the facts first and then worked out the full meaning of his observations. It was impossible to formulate the theory of evolution without first generating a large body of data from which the truth was extracted. The same holds true today. The problem, however, is that both data collection and hypothesis testing must be funded and funding is *always* limited. The decisions that will be made on the future of biological research will ultimately reflect policy makers' views on the value of both approaches to research. Already I hear the assertion, 'Of course hypothesis-driven research has to be funded, but...'. We will likely arrive at a point where hypothesis-driven research will receive the same second-class treatment as 'basic' research does today. In fact, the two will become synonymous for grant givers.

Perhaps it is time to re-examine and reassess the value of the research that we performed prior to the arrival of the systematic industrial approach. In reality, there is very often no hypothesis to guide us through our 'old-fashioned' research. Many scientists select a given gene—perhaps simply because their post-doc. or graduate studies led them to it—and

analyse it and its function as long as the grants allow. A trail of publications is laid down but very often it never pronounces a theory on the role of the gene or its product in a biological system.

Strangely, we are neither encouraged nor challenged by the scientific system to speculate on how the final picture will look. There is a difference between arguing that a protein is important in a biological system and postulating exactly why it is important. The standard publication and refereeing procedures militate against such speculations. How often have we read in the referees' comments on the discussion, 'This theory is not supported by the data presented'? Of course, speculation must have lacunae of data. We should encourage and challenge authors to speculate on the outcome of their research. This journal invites authors to add a brief speculation section to their discussion. Unfortunately, nobody has taken up this offer to date.

In the absence of hypotheses, scientists carrying out traditional research will be vulnerable to the accusation that they are collecting data inefficiently by outmoded means. If there were a stronger theoretical framework for the research then it would improve in quality and would be better suited to integrating data that come from large-scale projects. All practitioners of science will benefit from such a mixed approach. Those who have the knowledge of biological systems and who know the questions they raise will obtain their answers more quickly, and those who organise the industrial-scale approaches will see an immediate value added to their work. If, on the other hand, a wedge is driven between these two communities through comments such as the one cited at the start of this editorial, then there will be no winner. This situation must be avoided.

*Frank Gannon*

DOI: 10.1093/embo-reports/kvd094