

DISCOVERY RESEARCH IN THE CO-GENOMIC ERA — biology goes industrial

The parallel between the search for gold in the last century and drug discovery in the co-genome era is very clear. As with all gold rushes the rewards will go to the most creative entrepreneurs who can pull the technologies and knowledge together well enough to identify the rich prizes within.

**By Dr Timothy Wells
and Dr Georg Feger**

The sequencing of the human genome and the resultant renaissance in biology has been described by many as the new gold rush. In the biotech industry, the hottest debates always centre around who will actually derive the most value from the genome. In many ways there are parallels with the gold mining analogy. The successful people during the gold rushes of the last century were clearly marked out by a number of traits. They had experience of finding gold. The prospectors worked often alone, away from the crowds, to maximise their advantages. They were only interested in the gold that was easy to obtain (often just lying around waiting to be picked up).

Finally, poor communications meant that if they were successful, it was a long time before rumours of their success meant that they had competition. In many ways there are parallels with early drug discovery. A generation ago, most drug discovery was done on targets which used defined small molecules such as histamine, or nicotine as starting points. Generally, the work could be done by only a medicinal chemist and a pharmacologist. They harvested the 'low hanging fruit' of drug discovery, and relatively poor communications meant that the competition might take several years to enter the race.

At the beginning of the 21st century, gold mining is a completely different industry. A tonne of ore has to be crushed to enable the extraction of a few grammes of gold. The range of technologies that have to be brought together is much larger. There is no longer any idea of low hang-

ing fruit. Good communications mean that any competitive advantage is easily lost. However, what makes a successful organisation is not who has the most impressive range of technologies, but one that bring all these technologies together. Success means identifying and extracting gold from rich seams of rock, managing to get 10 grammes of gold per tonne compared to the competitor's five grammes. The parallel with drug discovery in the co-genome era is very clear. The genome is our equivalent of the geologic map of the world. The question is who can read this map well enough to identify the rich sources of gold. Clearly this means accessing a wide variety of technologies: informatics, genomics, proteomics, HTS, high density cell biology, rational drug design, combinatorial and medicinal chemistry, transgenic models of disease and high throughput pharmacokinetics. However, the prize is for the team that can bring the most value, while (to use the analogy) crushing the least rock. What is exciting about the genome is that we are confident there will be value in the genome for a new generation of pharmaceutical and biotechnology products. What is tantalising, is that at this stage nobody knows how many new products will emerge.

For most of us in the biotech sector, our goal is to find new targets. Targets can be broadly defined as points of intervention in metabolic, signalling and hormonal cascades, where we know that intervening with a therapeutic molecule will lead to a change in the disease pathology. Confirmation that proposed proteins are really

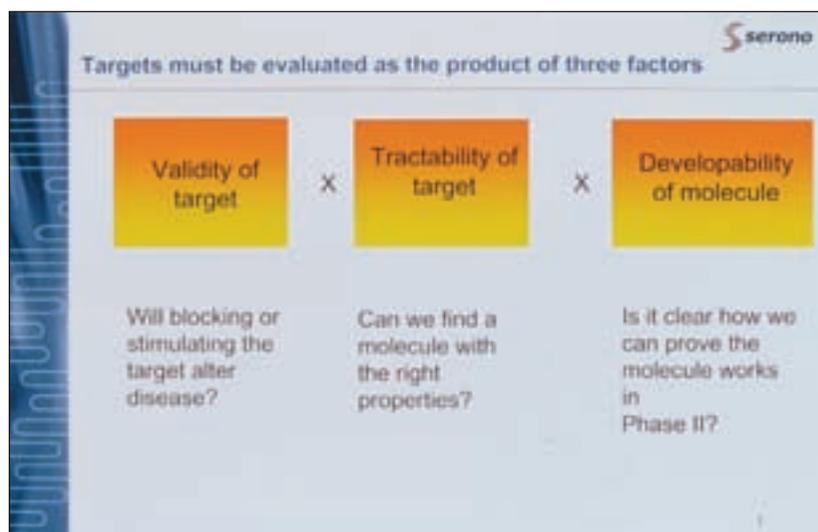


Figure 1
The mechanics of drug discovery. For a target to yield innovative new medicines, it must satisfy three criteria. It must be a validated intervention point, human genetics has massively increased our ability to relate potential targets to human disease. Second, it must be a tractable target, in other words one that is amenable to finding a drug, or where a protein factor can be found. The genome is already helping us to find all the members of known classes of tractable or 'drug-able' targets, with the hope that the technology that served us well in the past will help us to find lots of new molecules in the future. Since only valid and tractable targets are useful, any approach which relies too heavily on one method, will be problematical. Finally, the question as to what makes a developable molecule is the least well understood. For the moment, the genome studies do little to address this

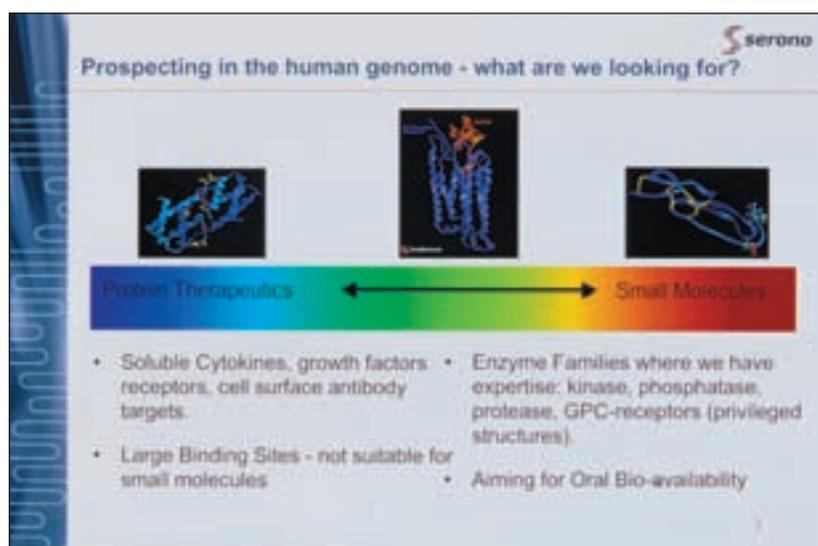


Figure 2
The continuum of molecular interactions. Historically, the pharmaceutical industry has been very good at finding molecules which bind in highly defined small pockets in receptors or enzymes. The biotech industry on the other hand has focused on making cytokines or growth factors, which bind to their receptors using much larger surface areas, and with less spatial definition. The genome will enable us to find more of these two extreme cases. After the genome, the technological challenge will be to find ways of addressing targets that lie in between these extremes

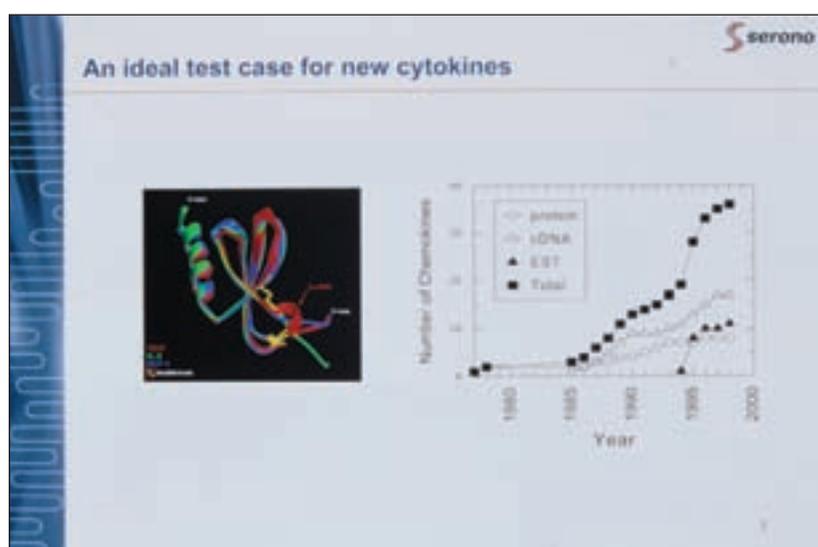


Figure 3
Finding new cytokines. The chemokine superfamily represents a good example of how the human genome sequence (in this case in the form of Expressed Sequence Tags or ESTs), enabled scientists to rapidly identify a large number of new molecules. This was because the chemokines themselves are relatively small proteins, and also because their structural features (disulphide bridges, beta-sheets and an alpha helix) are easy to spot. However, despite the fact that they were identified as much as seven years ago, a pharmacological role has not been clearly assigned for many of them, reinforcing that finding sequences is much easier than assigning biological relevance

Genomics

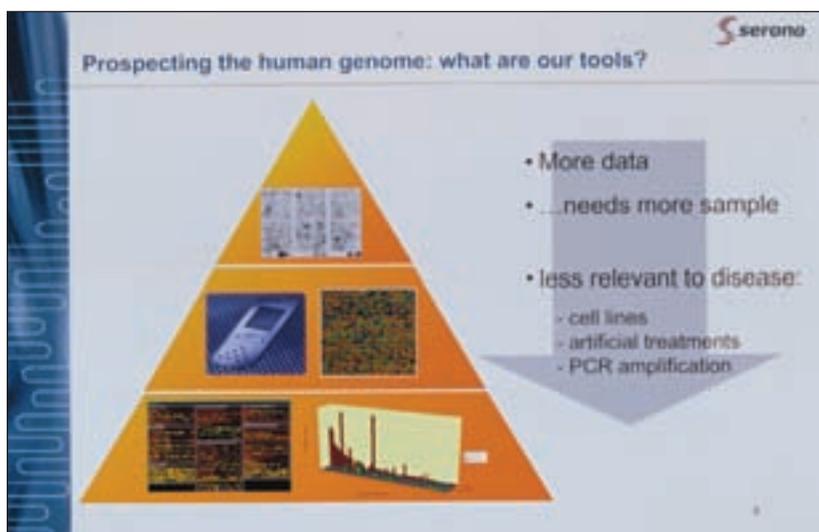


Figure 4
New technologies allow us to measure gene expression. Thousands of genes can be studied with filters, tens of thousands with gene chip technologies and the totality (hundred of thousands) with techniques such as SAGE. However, with increasing data, there is an increase in cost, and often an increase in the amount of material required for analysis. This leads people to use 'model systems' which take us one step further away from real human pathology

targets, (known as validation) means that we can show a role as a checkpoint in disease progression, rather than simply being in the right place at the right time. Often in the past, many projects have been successful in the preclinical stage, but failed in the clinic, simply because the targets chosen were not the ones which actually controlled the disease progression, they were not validated. Given the enormous expense associated with late stage clinical failures, it is clear that if our development of understanding of the human genome and human genetics can focus our discovery efforts on validated targets, there will be a massive gain in efficiency.

One way to enrich our chances of finding new targets linked to human disease is to look at the genetics of disease, only to work on proteins where some modification of the gene is correlated with the disease. Medical genetics has enabled us to find genes associated with the cause, onset and incidence of diseases, as well as the resistance to infection. However, in most cases the genes themselves are not good drug targets (unless one is in the business of replacing the lost gene with gene therapy). The reasons these genes may not be good targets vary, but often it is because there is no obvious way to modify their action with a small molecule.

Consider the classic case of the ApoE4 association with Alzheimer's, a small molecule to convert the ApoE4 protein to the physiological function of ApoE2 is not easy to imagine. On the positive side, the genetics studies highlight potential pathways that can be inhibited. This is especially true

if more than one gene defect is seen to cluster in the same metabolic or physiological process. The study of human genetics is useful as a starting point for getting new ideas on diseases, but we must be able to put these gene products in a physiological context if we are to harness these genetic observations to find new targets. In our drug discovery equation in Figure 1, these approaches are maximising our chances, by maximising the validity of targets.

The alternative approach that can be taken to discovering new targets in the genome era is to examine the alteration of gene/protein expression in disease. By looking at differences in the expression of cDNA (transcriptomics) or proteins (proteomics), between normal and disease tissues, over the time course of the disease, it is easy to identify new genes or proteins related to a disease. Obtaining such data is not difficult; the issue is how to validate the relevance of these genes in human disease.

Within these expression studies, we often find old friends, well-known receptors or enzymes. One degree of novelty can come from finding new members of old families. Medicinal chemists have been good at inhibiting such families of targets as proteases, kinases and G-protein coupled receptors. The biotechnology industry on the other hand has developed protein therapeutics, which are often structurally related growth factors, four helix bundle cytokines, and soluble cytokine receptors (Figure 2). New members that are identified are initially termed orphans, since their substrate/ligand and their function is unknown. Once the substrate/ligand has been identified (and despite the many new technologies that make this task much easier, this is not trivial), we still need to understand their function. Often the identification of function can only be done in vivo (many signal cascades involve messengers made in one tissue causing a response in another). Obviously, technologies such as knockout mice and transgenics may help guide us to a function for a gene product, but this is often a long and painful process.

Clearly, the combination of target identification and validation in the genome era generates a lot of data. One challenge is to integrate this: the genomic identification of a potential target, confirmation that the target may be linked to a physiological or metabolic pathway known to be related to the disease by genetic studies, and validation that the target is tractable. This is the goal for bioinformatics: not just integrating the data but getting it into the hands of the experimental scientist where it can be used to develop new scientific hypotheses.

Genomics

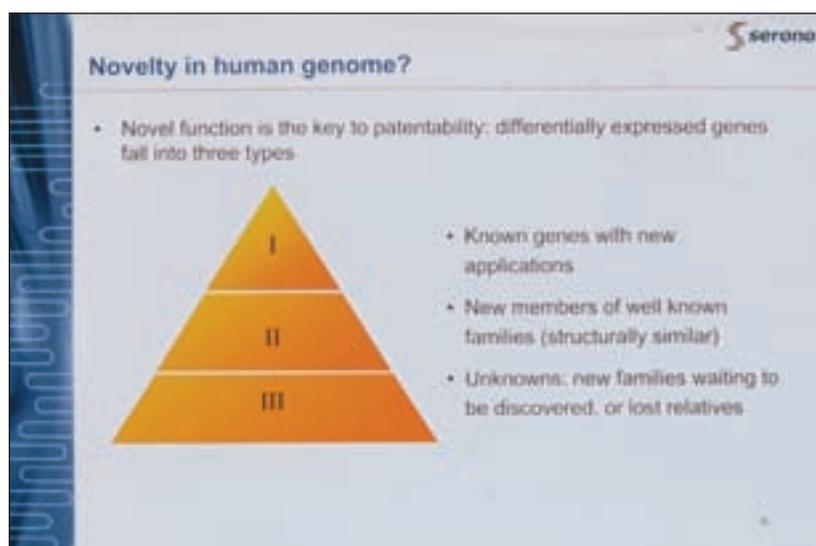
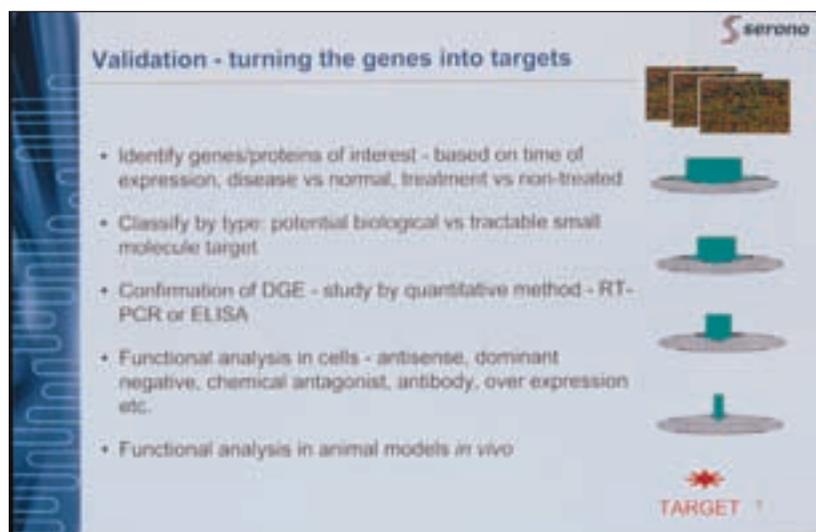


Figure 5

Having lots of over, or under expressed genes is not our goal. We need to know that they are validated, that there is evidence modulating activity of the gene or protein will lead to a modification of the human disease. Final validation can only take place in the clinic, but a number of filters can be put in place to make sure that as many of our targets are as highly validated as possible before entering clinical studies

Figure 6

There is value to genes in all three classes, even the Type I genes can give us use patents. Most people are hoping that there is a rich untapped seam of useful targets and therapeutic proteins in the 'unknowns'. This hope is similar to gold prospecting in Antarctica, the only things we know for sure is it will be difficult and expensive; but the rewards may be massive; and we have to have a better strategy than looking everything gene by gene

One of the key issues in the genomic era is that we risk overstating our understanding of the biology of these newly-found gene products, enzymes and receptors. Just because we know one activity of a protein (and thus give it a name) does not mean we understand the true physiology, or indeed the most likely role in pathology. Using the chemokine area of cytokine biology (Figure 3) as a case study, it is clear that the identification of novel members was a relatively rapid process, completed by several groups in 1994 using expressed sequence databases. What is interesting is that in many cases it took several years to identify a biological activity for the proteins. What is more interesting still is that it is far from clear that we have identified the correct (or physiological) role for each protein. Biology often progresses, by one group identifying a gene product and another (sometimes unrelated, sometimes competitor) discovering the function. The arrival of the genome sequence will not change this paradigm. Despite extravagant claims, no one company or university can hope to cover all of biology in the genome era.

With the majority of the human genome sequenced we are now well into the era of functional genomics. To give some context, functional genomics is no more and no less than biology (physiology and pathology) in the genome era. The consequence of our analogy with gold mining means that we need to be able to study large numbers of potential targets using high throughput cell biology. Our cell biology must also grow up – we need to look at target genes in cellular assays relevant to disease or biology, in vitro, and in more complex systems that are currently studied. As in ore mining we will extract from thousands of potential targets a few gold grains represented by the genes and their corresponding proteins that change the biochemical, phenotypic and pathogenic appearance of a cells interrogated. By adding in data about known polymorphisms in the genes of interest, we can now ensure that we maximise the likelihood of a connection with disease right from the beginning of our projects.

The key to who will get the most value from it is who can use the information to enhance the efficiency of their drug discovery operations. Although there is talk of thousands of new drug targets, the real prizes go to those who can distil out the few that are really important in disease and focus on those. The development of technology in the last 10 years has been astounding, however the rules for drug discovery have not

changed. We need the toolbox that is provided by the genome technologies and move functional biology to the industrial scale. However, in all this talk of high throughput biology, we must not lose track of the real root of success in searching for new targets. The data will provide us with a map, but we still need the prospectors who will use this map, and go hunting for targets, based on integrating all the disease understanding, pathology, pharmacology and genome sciences. In the end, as in all gold rushes, the prizes will go to the creative entrepreneurs, who can take the data and interpret it in new ways. **DDW**

Dr Timothy Wells is currently Vice-President Research, Head of Discovery Ares Serono, based in Geneva. He has worldwide responsibility for the research organisation focusing on providing new candidate molecules for clinical development in the key therapeutic areas of infertility, neurology, autoimmunity/inflammation and wasting. Prior to joining Serono, Dr Wells was Head of Biochemistry and Immunology at Glaxo Wellcome and worked at SmithKline Beecham on the molecular enzymology of Atherosclerosis.

*Dr Georg Feger is Head of the Molecular Biology Department at the Serono Pharmaceutical Research Institute (SPRI) in Geneva, which is part of Serono International SA. The Molecular Biology department consists of research groups in genomics, Tyrosine phosphatases and scientific computing with the common goal of discovering new drug targets. Dr Feger did his PhD in Molecular Biology at the EMBL in Heidelberg. In 1991 he moved for a postdoctoral position to the laboratory of YN Jan at the University of California in San Francisco where he worked on neurogenesis, limb development and DNA replication in *Drosophila*. In 1995 he became Head of R&D at GATC in Konstanz, Germany. In 1996, Dr Feger accepted a position as Group leader for microbial genome sequencing at GBRI (GlaxoWellcome), Geneva.*

Drug Discovery World Winter 2000/1

They say **less is more...**

**... with only four issues a year –
can you afford to miss one?**

**Written by experts within the
industry and academia, *Drug
Discovery World* discusses the issues
that matter in the fast-changing
world of drug discovery and
development.**

**To guarantee your own personal
copy every quarter make sure you
complete the **FREE** subscription
form found in this issue.**

drug discovery world

DDW

the quarterly business review of drug discovery & development