Genomes, screening, and a surfeit of riches

▼ This editorial is being written after the recent publication of two drafts of the complete human genome[1,2]. Even though neither of these 'complete' genomes are actually complete, they still represent a remarkable achievement, and one which rightly has caught both the public and scientific imagination. Of course, the pharmaceutical and biotechnology industries have been some of the main drivers for the human genome project, and have made major contributions to its completion. In addition, the use of genomic information, both from humans and other organisms, has been *de rigeur* in much of the drug discovery process for some considerable time and so these publications, in one way, are only incremental to what many of us are doing already. However, these publications do give us a much clearer idea of the limits of the human genome, as well as identifying many potential additional therapeutic targets that were not discovered by classical cloning or homology. Though there may be gaps in the genome, and the identification of which genes really are expressed is still not trivial or routine, we now have some limits on what we might conceive of as targets for drug discovery.

The fact that there are only 30,000–35,000 human genes rather than >100,000 as originally thought, will not simplify our choice of possible targets for non-infectious diseases. We now know there are at least 560 G-protein-coupled receptors, 80–90 cation channels, 130 ion transporters, 60 nuclear hormone receptors, possibly 3000+ transcription factors, 580 protein kinases and 120 phosphatases, over 200 proteases, and so on[1]. Even without including other identified signal transduction and enzyme families, this represents a very large number of targets just concentrating on those protein families that have been rich veins for the industry to date, and the problems of deciding which are valid therapeutic targets remains as one of the major challenges of drug discovery.

So, what does 'completion' of the human genome project mean for HTS, as distinct from the whole drug discovery process? It means that we will be getting not only more targets in known classes (which has throughput implications), but also new target classes that have yet to be exploited and where novel assay formats and technologies might be required. The next 'catalogue' stage of exploitation of the knowledge in the human genome will be to identify which proteins are expressed in individual cells and tissues. However, we will then be faced with the much harder, but considerably more interesting task of examining all the interactions that are made by these proteins and their dynamic changes in normal physiology and abnormal pathology. The integrative nature of the latter activity will identify many new possible pathways and targets for therapeutic intervention, but will consequently create major challenges for screening because the complexity of what will be discovered will require more-sophisticated assay formats that go beyond simple single protein–compound interactions.

These newer types of targets, along with an expansion in the range of targets in existing classes, will require considerable ingenuity in the identification and optimization of lead small-molecules in order to have appropriate therapeutic action without side effects. This will lead to a bigger variety of structural classes of compounds, which have the potential to produce side effects that will not be predicted from existing knowledge and databases. Although HTS against primary targets has become the norm in drug discovery, the same is not universally true for secondary assays, which are designed to provide information on, for example, toxicity and metabolism. Several articles in this issue tackle this question, and provide possible solutions to a very important issue, which will become even more important in the future as new information from the human genome project is integrated into drug discovery.

This is an exciting time for drug discovery and HTS, and we should regard the identification of many new potential targets as providing a surfeit of riches. We should not let the sheer numbers paralyze our decision making, but rather view this as a driver to devise novel approaches and technologies for HTS.

Paul England
Aurora Biosciences
11010 Torreyana Road
San Diego
CA, 92121, USA
tel: +1 858 404 6650
fax: +1 858 404 6787
e-mail: englandp@aurorabio.com

References

1 Dennis, C. *et al.* eds (2001) The Human Genome. In *Nature* (Vol. 409) pp. 813–958
2 Roberts, L. *et al.* eds (2001) The Human Genome. In *Science* (Vol. 29) pp. 1177–1351

1359-6446/01/$ – see front matter PII: S1471-1931(01)00148-3