The utility of predictivity in drug discovery and use ▼

In their article on molecular dynamics networks that recently appeared in *Drug Discovery Today*, Somogyi and Greller [1] describe network models that connect gene expression with disease states. Usually the connection is not monogenic but instead involves a cluster of genes whose expression pattern is meaningful. Thus, the expression patterns can be employed in a predictive sense to: (1) select drug regimens; (2) diagnose disease or disease susceptibility; and (3) plan a strategy for therapeutic intervention and target a pathway or the gene itself using reverse engineering.

A recent analysis of the pharmaceutical industry concluded that companies not relying heavily on predictive techniques in 2005 would not be in business [2]. It further concluded that the pharmaceutical industry must evolve 'from an interactive, intuitive way of doing things to systematic, predictive processes that will rapidly identify disease targets in commercially viable disease segments.'

There exist two distinct opportunities for predictive technologies in the gene-to-drug continuum, the first of which is nicely exemplified by Somogyi and Greller [1]. Network models predict biological activities for the following specific applications:

- oncology drug regimen selection on the basis of gene expression sets from tumor biopsies;
- differentiation of cancerous from non-cancerous tissue;
- stratification of cardiovascular disease and determination of disease susceptibility;
- prioritization of anti-coagulation drugs through modeling of the coagulation cascade; and
- identification of novel gene function using reverse engineering.

These are front-end opportunities, which derive from gene identification and expression information. This knowledge is highly valuable when unequivocal correlates have been established, and we will certainly arrive at this point. Individual gene profiling of patients will then be crucial medical history information and protein expression maps from diseased tissues of individuals will guide acute treatment.

Additional connections between expression genomics and disease states can be made. For instance, positional cloning is a process that involves genome-wide linkage scanning to identify a chromosomal region linked to a disease phenotype followed by mutational and association studies on genomic DNA within the linkage region. Linkage studies have identified susceptibility genes for non-insulin-dependent diabetes mellitus [3] and inflammatory bowel disease [4]. New connections between single nucleotide polymorphisms (SNPs) and specific disease states are constantly being discovered. For example, it has recently been shown that a mutation in the endothelial-leukocyte adhesion molecule 1 (E-selectin) gene (S128R) might be related to the amount of calcification in arteries, and this genetic polymorphism could provide an early detection signal for atherosclerosis [5].

The second predictive opportunity relates to genetic profiles that we all share, and specifically relates to how we process drugs with cytochrome P450 enzymes. For instance, with some exceptions, we all produce and rely on our liver cytochrome P450 enzymes for metabolizing drugs and other substances. At the heart of this predictive opportunity, and providing a bridge between the two predictive opportunities, is pharmacogenomics. Broadly defined, pharmacogenomics is 'the study of the impact of genetic variation on the efficacy and toxicity of drugs or the study of how a patient's genetic makeup determines the response to a therapeutic intervention' [6]. As an example of this point, we can use a pharmacogenomic analysis to determine a patient's allelic profile of cytochrome P450 enzymes. For the CYP2D6 isozyme, this analysis is useful because drugs can be activated or inactivated by this enzyme and 7% of the Caucasian population carry a poor metabolizer allele of CYP2D6 [7]. In staging a clinical trial with a drug that is processed by the CYP2D6 isozyme, this analysis is necessary. However, with CYP3A4, which is responsible for the majority of drug metabolism, pharmacogenomic analysis is not as helpful.

Although multiple alleles of CYP3A4 have been found, none apparently are completely inactivating [8,9]. Thus, with an enzymatic activity that is consistently present and necessary for chemical processing and drug metabolism, it is useful and pertinent to apply computational methods for predicting which drugs will be processed by the CYP3A4 isozyme, and which metabolites will be produced. Using prodrug strategies, predictive methods can be useful for the design of agents that will efficiently produce the desired drug or avoid accumulation of an undesired compound *in situ*.

It should also be noted that some disease states correlate with attenuations of specific cytochrome P450 enzyme expression. Individuals deficient in expression of CYP21B, an enzyme responsible for steroid 21-hydroxylation, suffer from congenital adrenal hyperplasia (CAH), arising from their inability to synthesize cortisol [10]. Interestingly, induction of CYP1A1 enhances the metabolism of polycyclic aromatic hydrocarbons (PAHs) following chemical exposure in mammalian cells, and the mechanism for this adaptive response is understood [11]. Furthermore, it appears that the genotype predicted to confer only 50% of the ability to metabolize nicotine via the CYP2A6 pathway correlates with twice the likelihood of smoking cessation

  PII: S1359-6446(02)02194-3

[12]. The reason for this association, however, is not clearly understood.

In summary, it is clear that at least two types of predictivity will be crucial to future therapeutic drug discovery. First, gene expression patterns will be used to select drug treatment strategies, diagnose disease and plan new strategies for intervention. For cytochrome P450 enzyme expression, it is useful to know of specific instances where expression varies from normal. However, a second type of predictivity involves the use of models to predict the specific sites of metabolism of drugs that undergo oxidation by the major CYP450 isozymes as well as relative rates. The combination of these two types of predictive models will allow us to accelerate the gene-to-drug discovery process.

References

1 Somogyi, R. and Greller, D. (2001) The dynamics of molecular networks: applications to therapeutic discovery. *Drug Discov. Today* 6, 1267–1277

2 PricewaterhouseCoopers (1998) *Pharma 2005: An Industrial Revolution in R&D* (available online at http://www.pwcglobal.com)

3 Horikawa, Y. *et al.* (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat. Genet.* 26, 163–175

4 Hugot, J.P. *et al.* (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411, 599–603

5 Ellsworth, D.L. (2001) Gender- and age-dependent relationships between the E-selectin S128R polymorphism and coronary artery calcification. *J. Mol. Med.* 79, 390–398

6 Peet, N.P. and Bey, P. (2001) Pharmacogenomics: challenges and opportunities. *Drug Discov. Today* 6, 495–498

7 Pettipher, R. (2000) The role of genomics in the development of new and improved therapies. *Innovat. Pharma. Tech.* 5, 62–67

8 Rebbeck, T.R. *et al.* (1998) Modification of clinical presentation of prostate tumors by novel genetic variant in CYP3A4. *J. Natl. Cancer Inst.* 90, 1225–1229

9 Felix, C.A. *et al.* (1998) Association of CYP3A4 genotype with treatment-related leukemia. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13176–13181

10 White, P.C. (2000) Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Endocr. Rev.* 21, 245–291

11 Ma, Q. (2001) Induction of CYP1A1. The Ahr/DRE paradigm: transcription, receptor regulation, and expanding biological roles. *Curr. Drug Metab.* 2, 149–164

12 Gu, D.F. *et al.* (2000) The use of long PCR to confirm three common alleles at the CYP2A6 locus and the relationship between genotype and smoking habit. *Ann. Hum. Genet.* 64, 383–390

Norton P. Peet
Vice President of Discovery Alliances
ArQule
19 Presidential Way
Woburn, MA 01801, USA