DDT Vol. 9, No. 22 November 2004

Molecular biomarkers in drug development

David A. Lewin and Michael P. Weiner

Arguably, the most immediately promising reverberation of the genomics era has been the application of biomarkers to drug development. The promise of applying biomarkers to early drug development is that they might aid in preclinical and early clinical decisions such as dose ranging, definition of treatment regimen, or even a preview of efficacy. Later in the clinic, biomarkers could be used to facilitate patient stratification, selection and the description of surrogate endpoints. Information derived from biomarkers should result in a better understanding of preclinical and clinical data, which ultimately benefits patients and drug developers. If the promise of biomarkers is realized, they will become a routine component of drug development and companions to newly discovered therapies.

David A. Lewin
CuraGen Corporation
Pharmacogenomics Services & Biomarkers
555 Long Wharf Drive
New Haven
CT 065011, USA
e-mail: david.lewin.grd.biol@aya.yale.edu
Michael P. Weiner*
The Rothberg Institute for Childhood Diseases
530 Whitfield Street
Guilford
CT 06437, USA
*e-mail: mweiner@childhooddiseases.org

▼ Biomarkers are 'a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention' [1]. For the purposes of this review, we will focus on the subset of biomarkers that might be discovered using genomics or proteomics technologies, 'molecular biomarkers'. Clinical biomarker analysis that relies on imaging technologies (e.g. microscopy, MRI, PET, X-ray) has been a subject of other recent reviews [2,3].

Currently, the greatest impact of biomarkers is through the use of diagnostics at the point of care. In the future molecular biomarkers should have their most lasting impact in preclinical and clinical studies to evaluate the safety and effectiveness of new drugs. Biomarkers that facilitate the development of new therapies might ultimately become companion diagnostics that guide the use and administration of drugs. In this review we will focus on the application of biomarkers to the early stages of drug development.

Beyond the general definition of a biomarker noted above, there are other generally accepted conceptions of biomarkers. Different institutions subscribe to these according to their objectives and perspectives on drug development. For example, some adopt a classification system for clinically applied surrogate markers first proposed by Mildvan *et al.*: Type 0, Type I, and Type II. Type 0 markers are those that are associated with the natural history of the disease. Type I markers are those markers that indicate a known response to therapeutic intervention, and Type II are those that reflect a clinical outcome as predicted by a surrogate endpoint [4]. Others will focus on biomarker definitions that are more analytical in nature, although these biomarkers might be applied in clinical settings as well as preclinical ones. In its recently released draft *Guidance Document For Industry: Pharmacogenomic Data Submissions* [5], the US Food and Drug Administration (FDA) is encouraging the voluntary submission of DNA-based molecular biomarker data for the agency to develop enough experience with this type of data to formulate guidelines for future pharmacogenomic submissions that might have regulatory impact. The guidance document pragmatically defines biomarkers. A 'valid biomarker' is a 'biomarker that is measured in an analytical test system with well-established performance characteristics and for which there is an established scientific framework or body of evidence that elucidates the physiologic, toxicologic, pharmacologic, or clinical significance of the test results.'

The document further refines this definition to include two additional levels: 'known valid biomarkers' that are evaluated using robust analytical methods yielding results that are generally accepted by the medical and scientific community, and 'probable valid biomarkers'. The latter definition differs primarily in terms of general recognition of the results within the expert community.

PII: S1359-6446(04)03272-6

These definitions are highly subject to interpretation and it will be the reduction to practice, initially as Voluntary Genomic Data Submissions (VGDS), which will set the parameters of what is useful to the FDA to make determinations of safety and efficacy. It is important to note that not all such data might be considered a VGDS. The draft guidance indicates classes of pharmacogenomic data that will have to be submitted as part of a drug application. General acceptance of biomarkers will be contingent on consistent demonstration of their use. Such demonstrations will aid in refining the definitions of biomarkers and facilitate their acceptance by regulatory agencies and pharma alike.

Application of biomarker panels to early drug development

A significant inefficiency in the drug development process is the late point at which safety and efficacy are determined. There is a high cost to drug development estimated by some to be in excess of US$800 million for a single agent, US$200 million of which might be attributed to compound failure [6]. The earlier the most common biologically-driven failure modes (safety – particularly hepatotoxicity [7]), efficacy and bioavailability [2,8] can be detected, the more likely it is that the cost of failure can either be reduced or redirected towards other value-generating activities, such as the characterization of back-up compounds.

Although there are well-documented costs associated with bringing a drug to market and with late-stage drug development failures, it is much harder to estimate the costs associated with premature termination of a research program. Should biomarker data play a role in development decisions? Obviously no decision with such profound impact on a company will be based solely on the results of a single preclinical or early clinical observation without a thorough evaluation of many types of data (toxicity being a significant exception). However, biomarker data, particularly if the biomarker is tightly correlated with the mechanism of action, might contribute to making a decision to change dose levels, or regimen, or to suggest an alternative indication. It is too early in the history of biomarkers to know what their long-term impact will be, although a good biomarker should clarify clinical observations and, in doing so, aid in drug development.

Expression pharmacogenomics and toxicogenomics [9,10] are examples of how nucleic acid-based biomarker panels can be applied to make early decisions regarding compound development. Early assessment of hits from HTS using biomarkers must be cost-effective and reliable to facilitate triage of both hits and lead series. Ideally, the cost over the life of a research program of applying early biomarker assessment for safety and efficacy should be offset by the anticipated savings brought about by a relatively greater number of successful animal efficacy and safety studies. Successful animal safety trials, however, are no guarantee of clinical success [8].

Poor therapeutic index (the ratio of therapeutic dose to toxic dose, TI) is a major reason for compound attrition [11]. The criteria used to interpret the results of predictive toxicogenomic and pharmacogenomic screens to anticipate favorable TI are strongly influenced by: the (1) intended indication; (2) tractability of the molecule to medicinal chemistry; and (3) the toxicity profile relative to compounds in the same chemical class and/or clinical application.

How might toxicogenomic and pharmacogenomic lead-stratification technologies be applied to the results of a drug screen to balance the toxicity and efficacy terms of the TI equation? Integration of toxicogenomic and pharmacogenomic data derived using molecular biomarkers combined with predicted ADME characteristics, solubility properties, and tractability for medicinal chemistry, as well as other molecular characteristics [12–14] might prove useful early in preclinical development for lead stratification. Together, these data might be useful for triage of lead series and aid in the ranking of the compounds according to favorable development characteristics (Figure 1).

The outstanding question is whether any one or more of the necessary technologies and data integration methods are sufficiently robust to make reliable early preclinical decisions. The catch-22 is that risks must be taken to test these methods to better understand the perils of using these approaches. Because the risks are ill-defined in terms of the costs of wrong preclinical decisions, there is justifiable hesitancy on the part of investigators to depend on these nascent technologies. To test these methods properly, one would have to run the rather expensive experiment of two parallel development paths. Is the cost of not developing what might have been a good drug comparable to the cost savings of avoiding late-stage failures? The answer to this question will not be clear until the predictive tools for early stage compound evaluation mature sufficiently to justify the risks. A promising new predictive efficacy method is one such tool that is a strong first step towards evaluating the use of early application of predictive methods.

Predictive efficacy

Recently, Gunther and colleagues described a method of predicting the class in which a psychoactive drug might fail with almost 90% confidence using the gene expression 'efficacy profiles' induced by 36 CNS drugs. The profile for

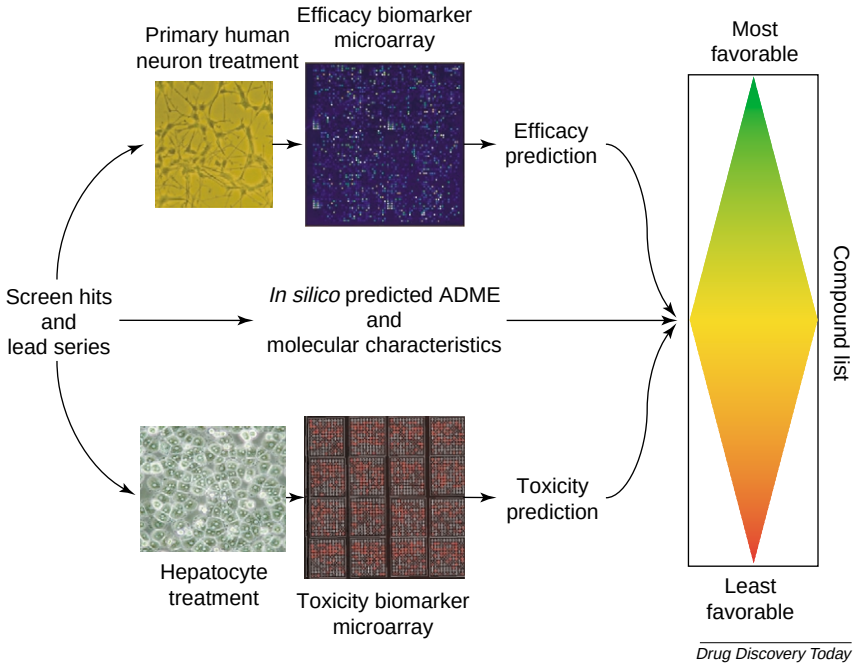

Figure 1. Application of predictive efficacy, toxicogenomics and molecular predictions to early lead selection. Hits and lead series molecules' characteristics might be evaluated using *in vitro* assays of efficacy and toxicity biomarker expression, combined with *in silico* predictions of ADME and drug-like qualities, to prioritize compounds for further development.

each of three indications was developed using a set of predictive efficacy markers derived from a training set of known antidepressants, antipsychotics and opioid drugs treatments to primary human neuronal cultures [14]. With the use of a proprietary microarray, they demonstrated that they could correctly predict the psychoactive class of drugs that had not been included in the training set with accuracies of 83.3–88.9% depending on the analytical method applied. In theory, this method could be applied to a series of lead compounds to identify hits from the screen that genuinely address CNS indications. The results from such an assessment might then be used to prioritize hits on the basis of psychological efficacy on the cellular level.

Although the results of this particular experiment did not yield clinically useful markers because of availability of CNS tissue, the principles can be generalized to other systems for the identification of markers that might have use in preclinical and clinical settings. The principle of efficacy profiling has been successfully applied to a neuronal cell line (E. Gunther, pers. commun.), which serves as a proof of principle that the method should be readily translatable to cell lines in general, as well as to the primary human material used in the original study.

The preclinical use of biomarkers derived using this method could include the comparison of leads to each other as well as to drugs in the same class both *in vitro* and potentially later in preclinical development using *in vivo* models. If the markers include secreted proteins in serum or markers assayable from a clinically tractable sample, then the same markers used in preclinical studies could be considered qualified for development and validation as clinical biomarkers (Figure 2).

Biomarkers in early drug development

Having a dose-response biomarker early in clinical development should be helpful for the setting of dose ranges in Phase II as well as for the development of dose regimens. If a biomarker is dose-responsive, then it should be useful in identifying a dose at which no meaningful increase in biological response is observed. If the dose does not increase the production of the biomarker, and there is no distinguishable change in apparent efficacy or toxicity, then the biomarker data would add to the confidence in setting a maximum dose to test in Phase II. For example, the lowest dose that results in a maximal biological response might be the highest dose that needs to be delivered in a dose-ranging study. Similarly, if multiple doses are given, then one might hope to see biological responses of approximately the same magnitude for each dose. If the biological response is attenuated with subsequent doses, then this might mean that the doses should be separated further or eliminated if the separation interval exceeds a meaningful period for the indication (Figure 3). Having a dose-response biomarker can help to set a dose that has an optimal TI.

Discovering biomarkers

Gene expression profiling and patterns of protein production can be used to identify candidate biomarkers of disease that might have use as surrogate markers in lead optimization, preclinical studies and clinical trials. In some cases biomarkers might also have the potential to be developed as diagnostic or prognostic devices (Figure 2b). In chronic diseases in which there are no reliable early predictors of therapeutic efficacy (for example, atherosclerosis, rheumatoid arthritis, osteoporosis, chronic obstructive pulmonary disease, hepatitis C-related illnesses, or Alzheimer's disease), clinical trials are inevitably long and expensive.

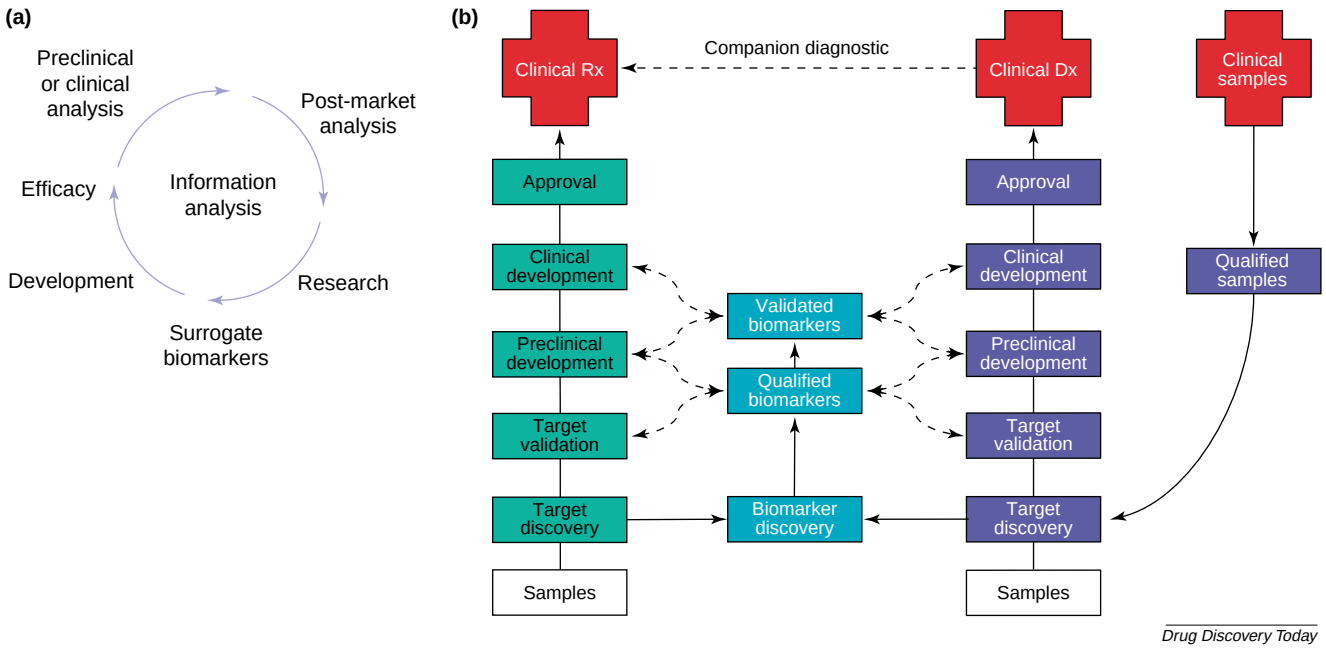

Figure 2. Biomarker discovery and application as an integral component of drug development. (**a**) Disease program approaches should yield biomarker candidates. (**b**) Parallelism of biomarker development in the drug development pipeline. Biomarkers can be applied to many steps in the drug-discovery pipeline. They can also be evaluated for use in both a pre- and post-market follow-up study.

The aim of developing reliable and specific biomarkers that act as early predictors of efficacy or long-term toxicity is to reduce the time, size and cost of clinical trials. The identification of biomarkers as a response to drug treatments in some situations might offer opportunities for development of early prediction tests for individual patient responses to therapy.

In general, regardless of whether animal models or clinical samples are available, it is important to keep in mind the nature of the samples that will be available in a clinical setting. The best biomarker is of no use if it cannot be accessed in the least invasive manner possible. This is essential for gaining the cooperation of the patients and the physicians who participate in a clinical trial. These samples, whether tissue or body fluids, can be examined using genomic and proteomic methods. Briefly described below are a small selection of genomic and proteomic technologies that are being used in the discovery and validation of molecular biomarkers.

Nucleic acid-based molecular biomarkers

Gene expression is a complex process involving coordination of dynamic events that can be regulated at multiple steps. Gene expression is controlled at the transcriptional level (transcription initiation, elongation and termination), at the posttranscription level (RNA translocation, RNA splicing, RNA stability), at the translational level (translation initiation, elongation and termination), and at the posttranslation level (protein splicing, translocation, stability and covalent modifications). Abnormalities in any of these steps can result in changes to the stoichiometry and activities of the various DNA/RNA/protein complexes, thus leading to an altered or disease status. It is the understanding of these functions of proteins and their interplay that is a major task of biomedical researchers.

Expression profiling has not only been successfully applied to identify pathways controlling biological phenomena, but it has also significantly contributed to increasing our understanding of regulatory pathways and the molecular biology of gene expression. Numerous examples of application of expression profiling have been described [15–18].

Methods for biomarker discovery using expression profiling are classified into two types: open and closed. Open systems are those that do not require an advanced knowledge of the sequence of the genome being examined [19]. An example of an open system would be 3′ end sequencing of expressed cDNAs, and its high-throughput derivative, the serial analysis of gene expression [SAGE] [20,21]. Closed systems almost always require some advance knowledge of the genes being examined [22–25]. The closed systems are exemplified by gene chips, including glass slides and the silicon-wafer-based gene chips [26–28].

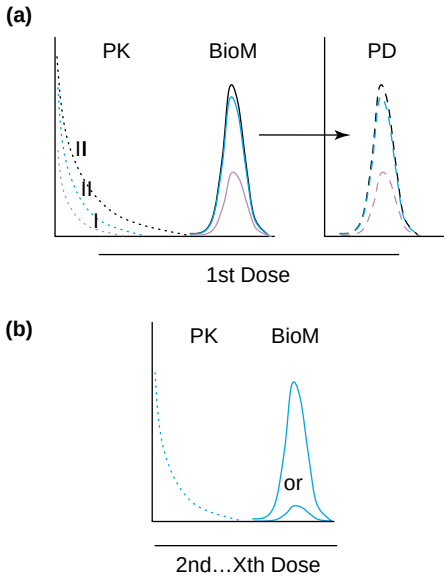

Figure 3. Theoretical use of biomarkers for setting dose and regimen. Dose response biomarkers in animal models might be useful in Phase I to identify minimum and maximum doses. Biomarker dose response in Phase I clinical samples might be useful to demonstrate if increasing dose level results in a corresponding biological response. If there is a correlation with a pharmacodynamic marker, or the biomarker itself is relevant to the mechanism of action, then these data should be helpful in setting a maximum dose at which a biological response occurs. (**a**) In this Figure, there is no significant difference between the response to dose level III and dose level II. Therefore, it might not be appropriate to dose patients in Phase II at the higher dose level. (**b**) Similarly, if multiple doses are being tested, then a diminished response to the subsequent doses might indicate that the time between doses might need to be varied to achieve a maximal biological response to the drug. PK identifies a hypothetical pharmacokinetic curve, BioM indicates the biomarker profile, and PD refers to a pharmacodynamic response.

Confirmation of nucleic acid-based biomarker candidates

Most genes discovered by expression experiments are necessarily confirmed by a separate secondary means. Typically, this is accomplished by the use of a quantitative method for gene analysis, such as either reverse transcriptase PCR (RT-PCR) or a quantitative PCR method such as real-time quantitative PCR (RTQ-PCR, for example TaqMan® – Roche). RT-PCR is a semi-quantitative means of measuring the amount of message in starting material, and although still being improved, has been mostly supplanted by the far more accurate RTQ-PCR method. The TaqMan method measures PCR product accumulation through a dual-labeled fluorogenic probe [29] and provides highly accurate and reproducible quantitation of mRNA. Unlike other quantitative PCR methods, real-time PCR does not require post-PCR sample handling, preventing cross-contamination of sample [30]. A third method, *in situ* hybridization, is sometimes used for tissue validation of markers genes. This method is much less quantitative, but has the advantage of providing information about expression in the context of tissue architecture. Such information is important for identifying the source of message and likely origin of protein to provide context to candidate biomarkers. For example, *in situ* hybridization can provide insight as to whether a biomarker candidate is originating from a tumor or the adjacent tissue. RTQ-PCR increases the quantization sensitivity by three orders of magnitude when compared with other methods. Easy implementation and rapid results at minimal cost of precious samples makes this method a logical first step for confirmation of biomarker candidates in clinical samples.

Discovering protein biomarkers

Technological advances in miniaturization and the groundbreaking research done with nucleic acids have now been carried over into the study of proteins. Until recently most of the attention and technological advances had been made with DNA chips. The attention paid to developing similar technologies for the assessment of protein populations has resulted in the development of methods for differential proteomic analysis to more directly measure the changes in protein levels in biological samples. The wealth of information and knowledge generated using these methods has provided new insights into complex biological processes and has enormous potential for biomarker discovery and development [31–41].

ICAT

Isotope coded affinity tag (ICAT)-based protein profiling is a chromatography–mass spectrometry (MS)-based method wherein two samples are differentially labeled with tags that differ in their incorporation of a heavy label. The samples are typically treated in a series of steps; tissue preparation, covalent tagging of cyteinyl residues, and multidimensional chromatography followed by MS. Informatics is then used to deconvolute differentially expressed peptides indicated by quantitation differences in the MS signal and then subsequent identification of the parental protein molecule. The ICAT reagent has a thiol-specific reactive group adjacent to an alkyl linker, which contains either nine [^{12}C] or nine [^{13}C] atoms and a biotin moiety. Typically, a whole cell or tissue protein extract is divided into several chromatographic-exchange fractions, with each being subjected to avidin chromatography. Cysteine-containing tryptic peptides are then isolated, followed by LC/MS/MS

analysis to identify ICAT peptide pairs and data processing used to quantify the relative [^{12}C]/[^{13}C] ratios. The mass difference of 9 Da between the control and the experimental tryptic peptides obviates the need to analyze by MS the control and experimental samples separately [42–44].

Differential gel electrophoresis (DIGE)

Two-dimensional (2D) gel electrophoresis is a technique that has been around for many years. Traditionally, it has been used to analyze differences between two protein samples, particularly posttranslational modification of proteins (especially the addition of phosphate groups). The limitation with this method has been the difficulty in comparing the results from two different gels. Most laboratories were not equipped to standardize the procedure of running enough 2D gels to statistically validate the method. To solve this problem, researchers at Carnegie Mellon took a page out of the gene-analysis methods and devised a set of chemicals that could be used to differentiate two samples run on the same gel [45,46]. By fluorescently tagging two samples with two different dyes, running them on the same 2-D gel, post-run fluorescence image the gel into two images, and then superimpose the images. The amine-reactive dyes they synthesized were designed to ensure that proteins common to both samples have the same relative mobility regardless of the dye used to tag them. This circumvented the need to compare several 2-D gels. With one sample labeled with Cy3 and the other with Cy5 dyes, images can be processed post-electrophoresis and, using image processing, determine the relative abundance of each protein spot. A third dye (Cy2) is now available to label all proteins in both samples and thereby serves as a useful internal control.

Tissue microarrays

As is the case for *in situ* hybridization for confirmation of nucleic acid biomarker candidates derived from expression analysis, tissue microarrays are a means of confirming protein expression and also provide information about protein distribution. This method has become a popular means for triaging biomarker candidates and an early step for validating protein biomarkers. These arrays can consist of a single tissue-type from many individuals, or arrays of many tissues from the same or different individuals. By multiplexing the number of screens performed, the use of these arrays can decrease both the time and expense needed to analyze a protein biomarker. Tissue microarrays have been used to profile and create databases of markers (including DNA, RNA and protein) for tumors [47–48]. These methods have been applied to archival specimens more than 60 years old [49] and tissue microarray technology has been used, discussed, or described in more than 200 scientific publications. The integration of cDNA microarray, high-density tissue microarray, and linked clinical and pathology data is a powerful approach to molecular profiling of human cancer [50] and holds great promise for the identification and preclinical validation of biomarkers for patient stratification and selection. Tissue microarrays have several advantages compared with conventional proteomics and genomics approaches. The speed of molecular analyses is increased several-fold by multiplexing, precious tissues are conserved, and a large number of targets can be analyzed from consecutive tissue sections [48,51]. Tissue microarrays are convenient because they can be treated as a single histological slide during staining, immunohistochemistry, and also for *in situ* hybridization, thus tying message information to protein production in the context of tissue architecture.

Correlating expression and protein production

There has been some skepticism regarding the usefulness of measuring mRNA levels to infer what is happening at the protein level. With few exceptions (such as when an interfering RNA is produced as a consequence of increased expression), when an mRNA modulates, the rate of production of its encoded protein is correspondingly affected. Whether mRNA modulation results in an immediately measurable change of the encoded protein steady-state level will be dependent on the protein turnover rate and other factors. However, there are examples of discordance between mRNA levels and protein levels. Reticulocyte α- and β-globin proteins are present at equimolar concentration, yet their encoding mRNAs are not. In this case there is more α- than β-globin mRNA, but this is compensated for by different rates of translation initiation of the two mRNAs [44]. Therefore, the level of an mRNA, when compared with the level of another mRNA, might not be a true indication of the relative levels of their encoded proteins. Biomarker candidates derived from expression analysis must be triaged according to their tractability for protein assessment unless it is certain that their use is only as a nucleic acid biomarker. Because the time between early drug discovery and the clinic is long, and the final indications and clinical samples might be hard to anticipate, it is ideal to have the ability to assay a biomarker as a nucleic acid and a protein.

Biochemical profiling – metabolomics

The intracellular concentrations of metabolites can reveal phenotypes for proteins active in metabolic regulation. Quantification of the change of several metabolite concentrations relative to the change of one selected metabolite

can reveal the site of action in a metabolic network [52]. Metabolomics provides a snapshot of the small molecules (amino acids, organic acids, sugars) in cells, fluids and tissues at a given time. Mass spectrometers are typically used to analyze samples and the peaks compared with a standards library is used to identify and name the known biochemicals and catalogue the unknown ones.

Cautious optimism

Correctly identifying a disease and choosing the best therapeutic is often a matter of knowledge, intuition, creativity and frequently, trial and error. Biomarker discovery efforts might well become as integral to drug development as toxicology and pharmacology and will likely add to the costs of early drug development. The question remains as to what will be the return on this investment. The usefulness of the technologies to discover and validate biomarkers should be evaluated according to their ability to reduce the costs of biomarker discovery and the efficiency with which they can be translated for application in a clinical setting.

If biomarkers are going to be used to contribute to preclinical decisions, they will have to be precise, sensitive and flexible in their use. Biomarker technologies will have to be amenable to many sample inputs and have readouts that can be compared across multiple platforms. The cost of applying technologies in preclinical and early clinical development must be evaluated based on the potential savings of advancing the most viable compounds and the risk of prematurely terminating a strong clinical candidate.

The development of biomarkers is not dissimilar to the development of diagnostics. A useful biomarker must have a degree of clinical sensitivity and specificity to be of greatest value. In fact, the synergy of anti-HER-2/*neu* antibody therapy (HerCeptin) for the treatment of breast cancer coupled to the HercerpTest has become a canonical example of how a biomarker/diagnostic can guide patient selection to the degree that it becomes the approved means of patient selection [53].

The study and development of biomarkers should be considered an emerging field with the potential to become a valuable tool for the pharmaceutical industry to develop better drugs at lower cost. The increasing focus on biomarkers is a function of the interest that has been taken by regulatory agencies as well as the potential for biomarkers to affect the economics of drug development.

The practice of using biological readouts to evaluate the status of a patient is ancient. After all, the definitive diagnosis of *diabetes mellitus*, 'the pissing evil' of Medieval times, had since ancient times been made by tasting patient urine to detect glucose content [54]. The use of technology to enable the marriage of biomarkers and drug development promises to reduce the cost of drug development and, hopefully, increase the frequency of sweet success.

Acknowledgements

The authors wish to thank J. Chant, R. Kinders and E. Gunther for critical reading of the manuscript.

References

1 Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* 69, 89–95
2 Frank, R. and Hargreaves, R. (2003) Clinical biomarkers in drug discovery and development. *Nat. Rev. Drug Discov.* 2, 566–580
3 Evans, S.M. and Koch, C.J. (2003) Prognostic significance of tumor oxygenation in humans. *Cancer Lett.* 195, 1–16
4 Mildvan, D. *et al.* (1997) An approach to the validation of markers for use in AIDS clinical trials. *Clin. Infect. Dis.* 24, 764–774
5 Lesko, L. *et al.* (2003) Guidance for Industry: Pharmacogenomic Data Submissions. U.S. Department of Health and Human Services, Food and Drug Administration, CDER, CBER, CDRH. http://www.fda.gov/cder/guidance/5900dft.pdf
6 Caldwell, G.W. *et al.* (2001) The new pre-preclinical paradigm: compound optimization in early and late phase drug discovery. *Curr. Top. Med. Chem.* 1, 353–366
7 Ballet, F. (1997) Hepatotoxicity in drug development: detection, significance and solutions. *J. Hepatol.* 26(Suppl 2), 26–36
8 Olson, H. *et al.* (2000) Concordance of the toxicity of pharmaceuticals in humans and in animals. *Regul. Toxicol. Pharmacol.* 32, 56–67
9 Lindpaintner, K. (2002) The impact of pharmacogenetics and pharmacogenomics on drug discovery. *Nat. Rev. Drug Discov.* 1, 463–469
10 Lakkis, M.M. *et al.* (2002) Application of toxicogenomics to drug development. *Expert Rev. Mol. Diagn.* 2, 337–345
11 Ulrich, R. and Friend, S.H. (2002) Toxicogenomics and drug discovery: will new technologies help us produce better drugs? *Nat. Rev. Drug Discov.* 1, 84–88
12 Ellinger-Ziegelbauer, H. *et al.* (2003) Characteristic expression profiles induced by genotoxic carcinogens in rat liver. *Toxicol. Sci.* 77, 19–34
13 Lipinski, C.A. *et al.* (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46, 3–26
14 Gunther, E.C. *et al.* (2003) Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles *in vitro*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9608–9613
15 Ye, S.Q. *et al.* (2002) Gene expression profiling of human diseases by serial analysis of gene expression. *J. Biomed. Sci.* 9, 384–394
16 Strausberg, R.L. *et al.* (2002) An international database and integrated analysis tools for the study of cancer gene expression. *Pharmacogenomics J.* 2, 156–164
17 Porter, D.A. *et al.* (2001) A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res.* 61, 5697–5702
18 Carulli, J.P. *et al.* (1998) High throughput analysis of differential gene expression. *J. Cell. Biochem. Suppl.* 30–31, 286–296
19 Green, C.D. *et al.* (2001) Open systems: panoramic views of gene expression. *J. Immunol. Methods* 250, 67–79
20 Velculescu, V.E. *et al.* (1995) Serial analysis of gene expression. *Science* 270, 484–487
21 Boon, K. and Riggins, G.J. (2003) SAGE as a strategy to isolate cancer-related genes. *Methods Mol. Biol.* 222, 463–479
22 Mei, R. *et al.* (2003) Probe selection for high-density oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.* 100, 11237–11242
23 Fodor, S.P. *et al.* (1993) Multiplexed biochemical assays with biological chips. *Nature* 364, 555–556
24 Pease, A.C. *et al.* (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5022–5026

25 Lipshutz, R.J. *et al.* (1995) Using oligonucleotide probe arrays to access genetic diversity. *BioTechniques* 19, 442–447
26 Pollock, J.D. (2002) Gene expression profiling: methodological challenges, results, and prospects for addiction research. *Chem. Phys. Lipids* 121, 241–256
27 Gnatenko, D.V. *et al.* (2003) Transcript profiling of human platelets using microarray and serial analysis of gene expression. *Blood* 101, 2285–2293
28 Nimgaonkar, A. *et al.* (2003) Reproducibility of gene expression across generations of Affymetrix microarrays. *BMC Bioinformatics* 4, 27
29 Heid, C.A. *et al.* (1996) Real time quantitative PCR. *Genome Res.* 6, 986–994
30 Kang, J.J. *et al.* (2000) Transcript quantitation in total yeast cellular RNA using kinetic PCR. *Nucleic Acids Res.* 28, e2
31 Ge, H. *et al.* (2003) Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet.* 19, 551–560
32 Kiechle, F.L. and Holland-Staley, C.A. (2003) Genomics, transcriptomics, proteomics, and numbers. *Arch. Pathol. Lab. Med.* 127, 1089–1097
33 Voshol, H. *et al.* (2003) Proteomics in the discovery of new therapeutic targets for psychiatric disease. *Curr. Mol. Med.* 3, 447–458
34 He, Q.Y. and Chiu, J.F. (2003) Proteomics in biomarker discovery and drug development. *J. Cell. Biochem.* 89, 868–886
35 Petricoin, E.F. and Liotta, L.A. (2003) Clinical applications of proteomics. *J. Nutr.* 133(Suppl), 2476S–2484S
36 Phizicky, E. *et al.* (2003) Protein analysis on a proteomic scale. *Nature* 422, 208–215
37 Patterson, S.D. and Aebersold, R.H. (2003) Proteomics: the first decade and beyond. *Nat. Genet.* 33(Suppl), 311–323
38 Sellers, T.A. and Yates, J.R. (2003) Review of proteomics with applications to genetic epidemiology. *Genet. Epidemiol.* 24, 83–98
39 Zhu, H. *et al.* (2003) Proteomics. *Annu. Rev. Biochem.* 72, 783–812
40 Hunter, T.C. *et al.* (2002) The functional proteomics toolbox: methods and applications. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 782, 165–181
41 Michaud, G.A. and Snyder, M. (2002) Proteomic approaches for the global analysis of proteins. *BioTechniques* 33, 1308–1316
42 Han, D.K. *et al.* (2001) Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat. Biotechnol.* 19, 946–951
43 Smolka, M.B. *et al.* (2001) Optimization of the isotope-coded affinity tag-labeling procedure for quantitative proteome analysis. *Anal. Biochem.* 297, 25–31
44 Gygi, S.P. *et al.* (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17, 994–999
45 Unlu, M. *et al.* (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071–2077
46 Tonge, R. *et al.* (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 1, 377–396
47 Kononen, J. *et al.* (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med.* 4, 844–847
48 Kallioniemi, O.P. *et al.* (2001) Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum. Mol. Genet.* 10, 657–662
49 Camp, R.L. *et al.* (2000) Validation of tissue microarray technology in breast carcinoma. *Lab. Invest.* 80, 1943–1949
50 Dhanasekaran, S.M. *et al.* (2001) Delineation of prognostic biomarkers in prostate cancer. *Nature* 412, 822–826
51 Fejzo, M.S. and Slamon, D.J. (2001) Frozen tumor tissue microarray technology for analysis of tumor RNA, DNA, and proteins. *Am. J. Pathol.* 159, 1645–1650
52 Raamsdonk, L.M. *et al.* (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat. Biotechnol.* 19, 45–50
53 Birner, P. *et al.* (2001) Evaluation of the United States Food and Drug Administration approved scoring and test system of HER-2 protein expression in breast cancer. *Clin. Cancer Res.* 7, 1669–1675
54 Harvey, R. (1998) The judgment of urines. *CMAJ* 159, 1482–1484