Combinatorial biomarkers: from early toxicology assays to patient population profiling

Ronald Koop

Advanced biomarkers for improved prediction and monitoring of disease and toxicology mechanisms are needed to control the high clinical failure rates among new compounds. Along with other uses, biomarkers are currently used in the industry to screen for toxic side effects of drug candidates and to identify appropriate patient populations. The use of combinatorial biomarkers is a first step towards an effective systems biology approach to drug development. In combination with high content screening, and in particular a novel, high content *in situ* proteomic technology, combinatorial biomarkers will strongly support the knowledge-based decision-making process by providing crucial information on functional biology.

The major challenge for the drug development industry is to reduce the high attrition rates of drugs entering clinical trials. Ever-increasing demands on drug safety and efficacy by the regulatory agencies and the public in general are generating growing pressure on the industry to move towards speciality drugs and to detect any possible safety complications as early as possible. A better understanding of disease and toxicology mechanisms is urgently needed to control the high clinical failure rate among new compounds, which at present is primarily a result of safety issues and lack of efficacy. Current industry estimates give a new medical compound entering Phase I testing only an 8% chance of reaching the market. The cost of bringing a new drug to market continues to escalate, with current estimates as high as US$0.8 billion to US$1.7 billion [1].

Over the past two decades, molecular biology has unraveled many detailed disease mechanisms but the translation of this knowledge into profitable drug development has been painfully slow and burdened with many failures. Drugs, of course, must function in patients; it is clear, however, that whole organisms are exceedingly more complicated than our current molecular models have led us to believe. This is not to say that the accumulated insight into disease mechanisms has not been of enormous importance, it is just that a more comprehensive understanding of biological networks is needed for rational drug design to really succeed. Ultimately, a systems biology approach is indispensable to our understanding of complex organisms [2,3]. In the long term, systems biology will almost certainly revolutionize the drug discovery and development process [4–6]. However, the question remains: what can be done now and in the immediate future to improve the efficacy and safety of novel therapeutics? Combinatorial biomarkers (Box 1) provide a potential answer to this problem. They are identified by systematic approaches, such as gene expression profiling, and represent a more holistic view of the organism, without the requirement for a complete understanding and systematic modeling.

Toxic effects are a main reason for compound failure (www.fda.gov/oc/initiatives/criticalpath/whitepaper.html) and there is a significant need for their

Ronald Koop
MelTec,
Leipziger Strasse 44,
D-39120 Magdeburg,
Germany
e-mail:
ronald.koop@meltec.de

1359-6446/05/$ – see front matter PII: S1359-6446(05)03440-9

BOX 1

What defines a combinatorial biomarker?

A biomarker is a measure of a biological response to a system change, for example, disease or environmental change, and can happen at the molecular, cellular or whole organism level. Biomarkers can be used for the study of specific disease processes.

The screening of single biomarkers has been in practice for several years and a popular example would be the determination of prostate specific antigen (PSA) for the early detection of prostate cancer. Most of the single biomarker measurements only have limited predictive values (e.g. ~20% for PSA) and it would only be logical to increase the predictive value by including additional biomarkers.

Theoretically, a combination of a pair of markers would generate a combinatorial biomarker, but in practice the identification of combinatorial markers usually involves gene expression profiling or other highly multivariable assays. Most combinatorial biomarkers consist of about five or more markers. Dependent on the particular assay used, each individual marker has to display a change of approximately twofold or more to be considered significant.

A combinatorial biomarker should be seen as a specific pattern that confers much more information than simply several individual measurements. The identification of these patterns requires the sophisticated bioinformatics analysis tools that are necessary for all multivariable assays, like gene expression or proteomic biochips.

Combinatorial biomarkers enable better specificity and/or sensitivity than single markers. Variants of these complex marker patterns can define subpopulations of patients or the severity of toxic effects, for example.

detection as early as possible in development. Several toxicological biomarkers and detection systems are available to determine toxicity; a systematic approach based on the best available knowledge of the biological system is needed to capitalize on their individual advantages of combining speed and confidence.

Companies are striving for safer, more-efficient drugs by targeting disease-specific entities, especially in oncology [7]. This creates enormous difficulties in the case of strongly diverse diseases like cancer because only a subset of patients might express the responding phenotype [8]. This has resulted in a strong need for appropriate biomarkers at the clinical end of the drug development process for patient population profiling towards individualized medicine [9]. The functional complexity of disease mechanisms calls for a combinatorial approach to designing screens that can accurately identify specific, yet comprehensive, biomarkers.

Early *in vitro* toxicology evaluation

Drug safety problems are still the main cause of failure in the clinic and predictive assays to assess these difficulties as early as possible during pharmaceutical development continue to be in high demand. When it comes to predictive toxicology, two things have to be considered, the cost of drug development and the power of predictivity for safety problems throughout human treatment.

Cost increases dramatically during the course of drug development and the ability to make decisions regarding termination of clinical development of a non-viable drug candidate as early as possible has a large financial impact for the pharmaceutical company.

Nevertheless, to ensure product safety, regulatory agencies require classical animal toxicology studies to show that the product is safe enough for early human testing and clinical studies to demonstrate safety for commercial distribution. Even so, these rigorous toxicology studies have not been able to avoid many extremely costly drug failures as a result of safety problems late in clinical development or even during commercial distribution. There is clearly room for improvement of toxicology evaluation methods and procedures also at these late stages in pharmaceutical development.

Gene expression profiling

The potential impact of gene expression, protein and metabolite profiling in toxicology, often referred to as toxicogenomics, is enormous and will enable a much better understanding of toxicology mechanisms. In a recent review, the authors argue that toxicogenomics will evolve into systems toxicology and eventually describe all toxicological interactions that occur within a living system [10]. However, it is clear that the development of useable systems toxicology will take time and cannot solve the drug safety challenges the industry is facing today. The following discussion will focus on gene expression profiling as the core of toxicogenomics but it should be noted that proteomics, and especially metabolite profiling (usually called metabolomics or metabonomics), is vigorously studied to date with respect to toxicology [11–15].

In the area of hepatotoxicity, toxicogenomics can no longer be regarded as 'new' technology and the application of gene expression markers to early stage preclinical safety assessment is probably already impacting the lead optimization process in ongoing drug development [16,17]. This is supported by several published studies on the changes in gene expression profiles on treatments of rats, cultured primary human hepatocytes or HepG2 cells with known genotoxic hepatotoxicants [18–22], or even with nongenotoxic hepatotoxicants, which currently can only be identified through long-term animal studies [23]. Published research in areas outside of hepatotoxicity is scarce but a couple of studies on renal toxicity have been described in the literature [24,25].

All these studies almost certainly identify excellent combinatorial biomarkers based on differential gene expression that work exceptionally well in the experimental setting and can be used to classify novel agents [26]. However, the real value of these novel combinatorial biomarkers will probably only be realized over time when they have contributed to lower the attrition rate of novel therapeutics entering clinical trials. In the meantime, and specifically to achieve this ambitious goal, data collection

and management systems have to be developed and employed. The experimental work required for a toxicogenomics study is extensive and the amount of generated data are vast. To create corresponding databases with sufficient data for confident statistic analysis, three main collaborative research consortia have been formed with scientists from industry, government and academia, as well as from regulatory agencies. The International Life Sciences Institute Health and Environmental Services Institute (HESI) Genomics Committee established a collaboration with the EMBL-European Bioinformatics Institute to develop a database [27,28], the Toxicogenomics Research Consortium of the National Institute of Environmental Health Services National Center for Toxicogenomics, which mainly analyzes environmental stress responses and the Consortium for Metabonomics Technology is producing a metabonomics database [29]. Also noteworthy in this respect are the efforts to combine gene expression data with histopathological information [30,31]. The basis for such a combined effort is to have comparable microarray results for different platforms and laboratories in the identification of response profiles. With the recent developments in microarray standardization, this seems to be the case, although differences in the response of individual genes of the patterns were observed [24,32–34] and some discrepancies still remain [35]. The preliminary conclusions of the HESI Genotoxicity Working Group so far are slightly disappointing. They conclude that microarray technologies are not currently amenable for use as a HTS tool for genotoxicity [27]. The technology appears to be less sensitive than standard genotoxic measures, which could be caused by relatively high variability and the expensive assays are complicated to analyze.

However, the same group also concludes that transcription profiling can be used to differentiate compounds by their mode-of-action and later the possibilities of combining proteomic profiling, as a further step in gene expression profiling, with 'classical' cell-based high-content analysis will be discussed.

In silico *approaches*

Despite their growing importance, *in silico* methods for the prediction of metabolism and toxicity as tools to assist library design and lead optimization will be discussed only briefly in this article. *In silico* ADME-Toxicity (ADMET) models have been around for slightly over a decade with initially unsatisfactory results but they have evolved into a helpful, but still not too trusted, tool that certainly still requires improvements [36]. By integrating experimental and computational technologies with respect to ADMET evaluation in drug discovery, it has been possible to increase the value of *in silico* methods [37]. This raises the important possibility that commercial enterprises could exploit this field by combining *in silico* ADMET with a gene-expression profiling approach. Companies that focus on *in silico* ADMET, such as Bio-Rad, Accelrys or Tripos, already use models validated with experimental data in standard toxicology assays. Others, like UK-based Inpharmatica, have by now started to integrate microarray annotation utilities in their suite of programs and the incorporation of a toxicogenomics database is only a logical, however ambitious, next step. By contrast, there are companies, like Gene Logic or ACLARA, that already have some of the largest toxicogenomics/metabonomics databases and the analytical capabilities for predicting potential toxicity. The integration of chemistry and genomics to profile drug candidates is also used by Iconix Pharmaceuticals with their extensive chemogenomics database, which contains gene expression and molecular pharmacology information. The incorporation of these technologies in a systems biology approach, such as that being currently developed by GeneGo, for example, will strongly improve '*in silico*' ADMET assessment early in the drug development process [38].

Cell- and tissue-based high-content analysis

Although there is certainly some movement towards *in silico* studies, ADMET usually involves animals, or at least cells, because the decisive test for new compounds is still the clinical trial. When it comes to cells or tissues, imaging technologies enable an expanded toxicology toolkit because several factors, such as the mitotic index, cell viability, micronuclei, apoptosis and cell-cycle progression, can be monitored. Imaging technologies have been used in scientific research and possibly early target identification studies for a long time but only made their appearance in the drug development process about eight years ago, with the advent of high-content screening (HCS) [39]. Using HCS, multiple interactions or independent targets are analyzed simultaneously in cells, using fluorescent reagents and advanced imaging equipment. The technology has obvious potential for ADMET analysis and is used widely in that respect for genotoxicity studies by screening for micronucleus formation [40]. The sophistication of the method is demonstrated in the kinetic multiparametric cytotoxicity assay for live cells offered by Cellomics. Nuclear morphology and size, calcium homeostasis, the mitochondrial transmembrane potential and membrane permeability are assayed simultaneously to provide cell-based biomarkers for toxicology analysis [41].

To combine high content analysis with the combinatorial potential of toxicogenomics, MelTec, a biotechnology company based in Germany, has developed a technology called MELK that provides combinatorial, cell-based biomarkers. MELK is a robotic whole-cell imaging technology that integrates cell biology and biomathematical tools to visualize simultaneously dozens of proteins in a structurally intact cell or tissue. The technology is largely incompatible with paraffin-embedded, formalin-fixed tissue because of difficulties with antigen retrieval procedures and only cryo-fixed tissue can be used at present. It consists of a cycling process based on immunofluorescence microscopy, whereby cellular proteins are labeled

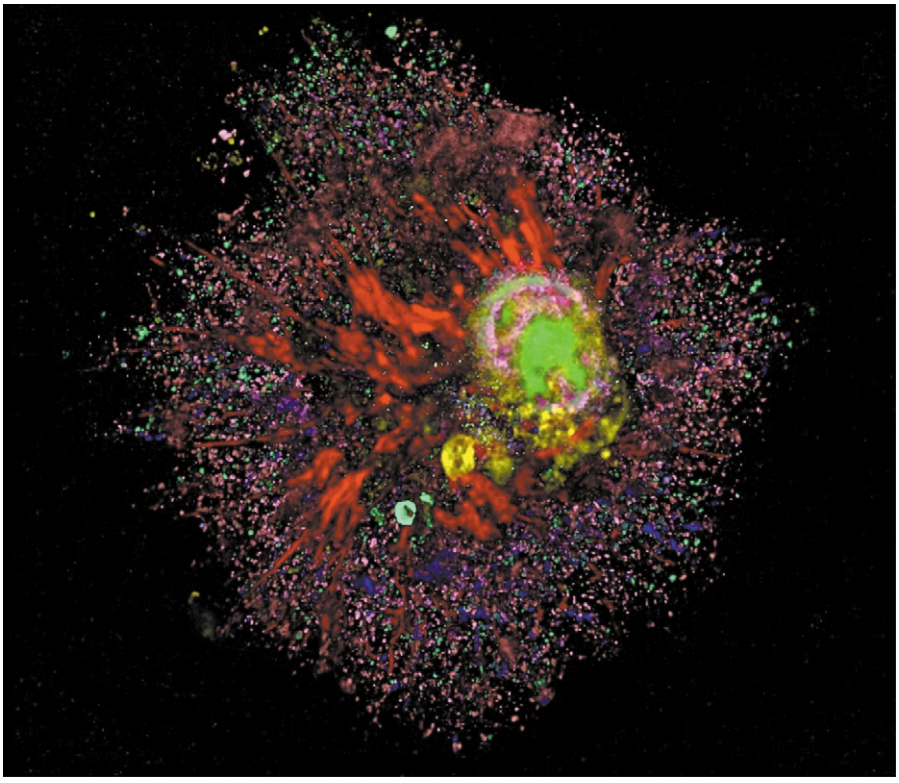

FIGURE 1

A multiplexed, 3D image of 20 proteins in a primary hepatocyte. Different colors represent different proteins. It should be noted that it is difficult to display 20 proteins simultaneously in a single image capture. Visual image analysis with a sophisticated 3D software tool (Imaris, Bitplane AG) reveals an astounding complexity of cellular structures and endosomal vesicles.

with fluorescent tags, usually dyed antibodies. After the dye is bleached (destroyed) by light, the next round of MELK is started by adding another fluorescent dye-labeled antibody, directed against the second protein and so the cycle continues. The visualized information generated by MELK is then processed through the company's data analysis software, enabling the identification of protein networks that have a crucial role in biological processes. Because of the difficulty associated with quantifying protein concentration using signal intensities, the data analysis is primarily based on changes in protein patterns that are more or less independent of protein concentrations. By simultaneously elucidating disease mechanisms and screening for the effects of compounds on these and toxicology-related pathways in a tissue context, a combination of toxicity analysis with pathway monitoring is possible. This technology is based on immunofluorescence so the value of any assay is dependent on the proper selection of antibodies and, therefore, somewhat reliant on existing knowledge. Although currently not useful for high-throughput applications, MELK technology nonetheless addresses crucial decision-making steps involving selected targets or leads.

Novel methodologies to quantitative microscopy have been reported recently [42,43]. These approaches use sophisticated computation methods to define features of single- or double-labeled immunofluorescent images that can be compared statistically between different experiments. By contrast, MELK enables access to this complexity at the level of protein networks (Figure 1) by performing proteomics characterization of single cells for a large number of proteins simultaneously. The outcome of such analyses is expressed in a 'partial Toponome map' that provides a 3D description of a subset of protein networks, depending on the antibody library used, that enables the construction of functional protein linkage maps. The toponome is the entirety of all protein networks in a biological sample and encompasses the total functional code of a cell. Whereas the computational methods use readily available microscopic instruments for analysis, MELK is a novel microscopic tool. It will be interesting to see in the future how accurate the predictions of complex protein co-localizations by computational methods will prove to be in comparison with the actual measurements by multiplexed immunofluorescence.

By screening for changes in the distribution of all possible combinations of proteins (50 proteins give rise to 2^{50} or $\approx 10^{15}$ different possible combinatorial protein patterns) on drug treatment or existence of disease, MelTec can identify specific combinatorial biomarkers for toxicity analysis, lead compound selection and clinical monitoring.

Biomarkers for patient stratification and clinical trial diagnosis

It appears to be an obvious statement that a drug should only be given to a patient suffering from the disease for which the drug is targeted. However, this requires an exact diagnosis of the illness, which is exceedingly difficult in certain diseases – particularly highly heterogeneous diseases like cancer. It is now generally understood that the development of cancer is a multistep process requiring multiple genetic mutations, resulting in a high variability between individual tumors [44]. Furthermore, the advent of molecular targeting in drug discovery has opened new opportunities for progress against cancer with highly specific and less harmful treatments, in contrast to the broadly acting cytotoxic agents. Probably, these targeted therapies will prove beneficial only in specific patient subsets as defined by their molecular signatures, therefore, biomarkers are urgently needed to support the classification

of patients. Protein biomarkers are likely to have an increasing role in this procedure, as well as in several other areas in the drug development process. Hanke *et al.* [45] discuss this in an excellent review, along with the important role of protein biomarkers in identifying patients at high risk for prevention therapy.

The screening of patients for genes or genetic mutations to determine the appropriateness of a specific treatment is already a reality. Examples like the gene encoding Her2 and Herceptin, or the gene encoding epidermal growth factor receptor (EGFR) and Erbitux, show that potent biomarkers are needed to maximize the potential of many pharmaceuticals, as well as to minimize the limitations of these drugs. In these cases, the biomarker is a single gene or protein and also the treatment target. These examples reflect the early stage of molecular biomarker utilization in the clinic and leave ample opportunity for further development. The advantage of combinatorial biomarkers in clinical trial monitoring arises from their broad application because they can be used even if the mechanism-of-action of a drug is less clear. They can also reveal details about the mechanism and monitor toxic side effects at the same time. It is desirable to monitor multiple markers and pathways from small blood or tissue samples in their natural biological context because the availability and amount of patient samples is often limited.

Genetic screens and specified targets

The ErbB2 (HER2/neu) protein is overexpressed in ~25–30% of breast cancer tumors, among others, and is linked with a relatively poorer clinical outcome in a variety of epithelial malignancies [46]. Genentech developed a monoclonal antibody, Herceptin, that inhibits ErbB2 and was one of the first approved targeted cancer therapeutics. The response rate in breast cancer patients overexpressing ErbB2 was ~50% when used in combination with chemotherapy [47]. It can easily be seen that the response rate in all breast cancer patients, regardless of their ErbB2 status, would be only 15%. Therefore, it was important that at the time the FDA approved Herceptin for the treatment of metastatic breast cancer, the FDA's Center for Devices and Radiological Health approved a diagnostic kit for Her2 expression. Currently, there are two diagnostic tests available, HercepTest and PathVysion, which are used for the identification of tumors that show the best promise for susceptibility to Herceptin. There is clearly room for improvement in the response rate of cancer patients depending on better diagnostic tools. Studies with cell lines have demonstrated that not all ErbB-expressing cell lines respond to inhibition of ErbB receptors [48]. A recent retrospective study on tissue microarrays from breast cancer patients that were treated with Herceptin and chemotherapy demonstrated that, apart from ErbB2 expression, the status of EGFR and the ErbB ligands, neu differentiating factor (heregulin) and transforming growth factor α, also affect therapy response [49]. This is a clear demonstration for the value of combinatorial biomarkers in the stratification of cancer patients.

The Erb family of proteins appears to be attractive for targeted cancer therapy. The colorectal cancer drug Erbitux (ImClone System), and the non-small cell lung cancer drug Iressa (AstraZeneca) are two FDA-approved agents targeting the EGFR; numerous other drugs are currently in clinical trials in various therapeutic areas. Biomarkers for susceptible tumors are necessary for effective targeting with these drugs. A companion diagnostic test for EGFR, called the 'EGFR pharm DX Kit', was approved for use with Erbitux in early 2004. Iressa initially baffled clinicians with mixed results: it only seemed to work in ~10% of the patients; however, in that 10% of patients, it resulted in an often dramatic reduction of tumor size. The FDA approved the drug for use in the USA and, later, two independent research groups discovered that particular somatic mutations in the gene encoding EGFR in lung tumor tissue corresponded with the significant clinical responses seen in earlier trials [50–52]. This information immediately encouraged the development of companion diagnostic tests and quickly resulted in independent laboratories offering a genetic test for patients, screening for EGFR mutations in DNA isolated from tumors. However, despite encouraging results in some patients, the company eventually pulled Iressa off the market in late 2004 after it failed to show significant effects on prolonged survival rates in additional clinical trials.

These examples show the importance of biomarker use to stratify patient populations to increase efficacy for molecular-targeted drugs. With numerous speciality drugs on the horizon, determining the optimal drug–patient combination is a challenging task for biomarker discovery.

In addition, the identification of patients who might be susceptible to dangerous side effects will result in overall improved drug safety. For instance, there are now clinical tests available to screen for mutations in the gene encoding thiopurine S-methyl-transferase (TPMT), which affects drug metabolism of the immunosuppressants mercaptopurine and azathiopurine. The routine use of TPMT genotyping is still limited [9]. Clearly, there is still a long way to go until we reach a level where medicine is truly individualized. Most likely, this will involve the development of comprehensive combinatorial biomarkers with the help of '-omics' technologies and systems biology.

Combinatorial biomarkers or '-omic signatures'

Advanced developments in biotechnology enable researchers and clinicians to get a more holistic view on disease processes with a combination of genomics, proteomics and metabolomics, providing new combinatorial markers for the susceptibility of individual patients to specific therapies. These technologies enable the efficient monitoring of a large number of key cellular pathways simultaneously, to enable improved cancer screening, patient monitoring and choice of therapy [53]. The advantage of combinatorial

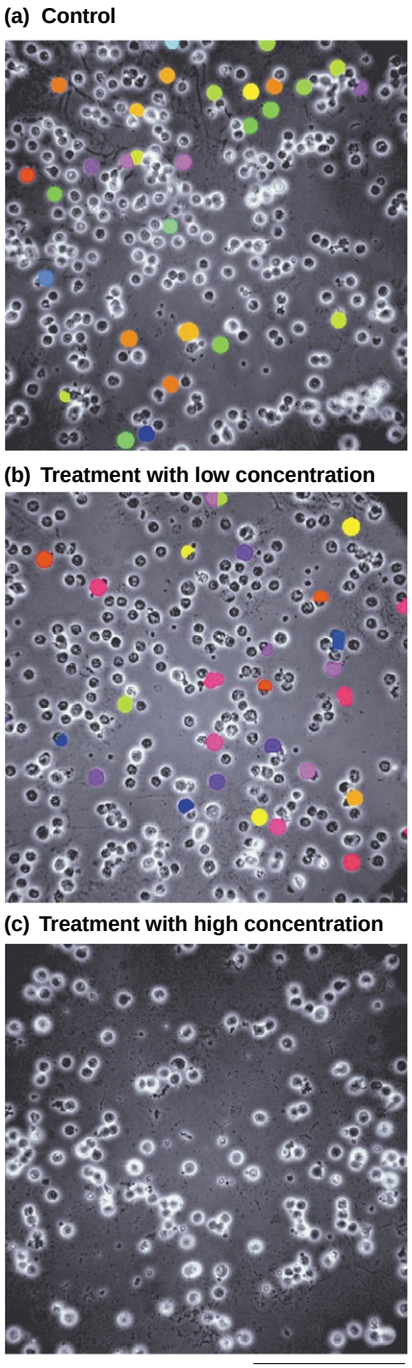

FIGURE 2

A combinatorial biomarker for toxicity in peripheral blood lymphocytes. Peripheral blood lymphocytes were treated in vitro with different concentrations of a chemotherapeutic drug and MELK runs were performed with a library that included 30 antibodies against leukocyte surface antigens. Changes in the distribution of cell subpopulations in control **(a)** and samples treated with high **(b)** or low **(c)** concentrations of cisplatin were observed. A subpopulation of lymphocytes disappeared after treatment. Cells containing the distinguishing protein combination are colored. The differences in colors reflect the difference in distribution of other proteins.

biomarkers in clinical trial monitoring is a result of their broad application because they can be used even if the mechanism-of-action of a drug is relatively unclear. In addition, they can be used to assist treatment decisions by providing markers from different stages of disease progression. They can also reveal details about the mechanism and can monitor toxic side effects at the same time.

Application of the new '-omic' technologies is broad and genetic screening for single nucleotide polymorphisms, for example, has been shown to have some promise in the prediction of restenosis [54], whereas proteomics techniques have been used for the diagnosis of, among others, ovarian [55] and prostate [56] cancers, as well as gastric adenocarcinoma [57].

Another novel technology that has not been mentioned here is the screening of DNA methylation markers, and their potential for diagnostics and therapy management is discussed in a review with respect to gastrointestinal cancers [58].

One drawback to classical proteomic technologies is that they require tissue homogenization, thereby losing sensitivity and topological information. It is possible to partially overcome this disadvantage by using laser capture microdissection [59]. Although time consuming, rapid robotic laser dissecting machines are now available that make this approach less cumbersome and have led to promising results, especially in combination with protein arrays, where hundreds of proteins have been analyzed from only a few hundred cells [60,61]. However, even using this technique, several-to-many cells must be pooled, and the tissue context is still lost. In this respect, immunohistochemistry (IHC) and high content analysis offer a major advantage because the specific location of the target within the tissue can be determined. This can be crucial because, for example, tumors also contain stroma, fibroblasts, immune cells, blood vessels, and so on. Obviously, this heterogeneity creates enormous challenges for bioinformatic analysis and, although companies like Definiens offer automated cell detection systems for tissue samples, the data-mining process is labor-intensive and requires experienced users. The recent development of tissue microarrays progresses IHC in the direction of systems biology and, by using analytical methods developed for gene expression profiling to tissue microarray immunostaining data, prognostics of breast carcinoma can be improved [62].

With the ability to stain and monitor more than 100 proteins in the same tissue section, the MELK technology platform brings IHC closer to a proteomics technology and merges the combinatorial power of proteomics with the topological advantages and challenges of IHC, thereby creating 'toponomics'. Expert knowledge and protein analysis by classical proteomic technologies combined with protein network databases facilitate the formation of suitable antibody libraries for specific assays. Such antibody libraries enable the separation of groups of tissue samples from patients by statistical data mining. For the analysis of heterogeneous cell populations, it is

also possible to recognize individual cells automatically and perform cell-based data mining (Figure 2). MELK can monitor multiple markers and pathways in small blood or tissue samples in their natural biological context, which might make it a useful option for the application of '-omic' technologies in patient stratification and monitoring, although its value remains to be established in comparison with more established methods, such as HCS, gene expression profiling and classical proteomics.

Conclusion

Biomarkers, ranging from histopathological analysis of liver toxicology to the screening of cancer patients for mutations in their EGFR, are widely used in the pharmaceutical industry and in clinical medicine. Generally speaking, diagnostics is nothing more than the use of biomarkers. The intention of this review is to show that, to keep pace with the growing need, the development of combinatorial '-omic' biomarkers is paramount. Drug safety problems continue to be the leading cause of drug failure in the clinic, therefore, predictive biomarker assays to assess these difficulties as early as possible during pharmaceutical development continue to be in high demand. With the advent of molecular targeting in drug discovery, determining the optimal drug–patient combination becomes a central issue for biomarker discovery.

References

1 Gilbert, J. *et al.* (2003) Rebuilding Big Pharma's Business Model. *In vivo: The Business & Medicine Report* 21, 73
2 Ideker, T. *et al.* (2001) A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* 2, 343–372
3 Nurse, P. (2003) Understanding cells. *Nature* 424, 883
4 Hood, L. and Perlmutter, R.M. (2004) The impact of systems approaches on biological problems in drug discovery. *Nat. Biotechnol.* 22, 1215–1217
5 Mack, G.S. (2004) Can complexity be commercialized? *Nat. Biotechnol.* 22, 1223–1229
6 Butcher, E.C. *et al.* (2004) Systems biology in drug discovery. *Nat. Biotechnol.* 22, 1253–1259
7 Strausberg, R.L. *et al.* (2004) Oncogenomics and the development of new cancer therapies. *Nature* 429, 469–474
8 Stratton, M.R. and Futreal, P.A. (2004) Understanding the target. *Nature* 430, 30
9 Evans, W.E. and Relling, M.V. (2004) Moving towards individualized medicine with pharmacogenomics. *Nature* 429, 464–468
10 Waters, M.D. and Fostel, J.M. (2004) Toxicogenomics and systems toxicology: aims and prospects. *Nat. Rev. Genet.* 5, 936–948
11 Merrick, B.A. and Tomer, K.B. (2003) Toxicoproteomics: a parallel approach to identifying biomarkers. *Environ. Health Perspect.* 111, A578–A579
12 Petricoin, E.F. *et al.* (2004) Toxicoproteomics: serum proteomic pattern diagnostics for early detection of drug induced cardiac toxicities and cardioprotection. *Toxicol. Pathol.* 32 (Suppl. 1), 122–130
13 Grifin, J.L. and Bollard, M.E. (2004) Metabonomics: its potential as a tool in toxicology for safety assessment and data integration. *Curr. Drug Metab.* 5, 389–398
14 Schmidt, C.W. (2004) Metabolomics: what's happening downstream of DNA. *Environ. Health Perspect.* 112, A410–A415
15 Weckwerth, W. (2003) Metabolomics in systems biology. *Annu. Rev. Plant Biol.* 54, 669–689
16 Yang, Y. *et al.* (2004) Toxicogenomics in drug development: from preclinical studies to clinical trials. *Chem. Biol. Interact.* 150, 71–85
17 de Longueville, F. *et al.* (2004) DNA microarrays as a tool in toxicogenomics. *Comb. Chem. High Throughput Screen.* 7, 207–211
18 Harris, A.J. *et al.* (2004) Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutat. Res.* 549, 79–99
19 Huang, Q. *et al.* (2004) Gene expression profiling reveals multiple toxicity endpoints induced by hepatotoxicants. *Mutat. Res.* 549, 147–167
20 Heinloth, A.N. *et al.* (2004) Gene expression profiling of rat livers reveals indicators of potential adverse effects. *Toxicol. Sci.* 80, 193–202
21 Steiner, G. *et al.* (2004) Discriminating different classes of toxicants by transcript profiling. *Environ. Health Perspect.* 112, 1236–1248
22 Ulrich, R.G. *et al.* (2004) Overview of an interlaboratory collaboration on evaluating the effects of model hepatotoxicants on hepatic gene expression. *Environ. Health Perspect.* 112, 423–427
23 Kramer, J.A. *et al.* (2004) Acute molecular markers of rodent hepatic carcinogenesis identified by transcription profiling. *Chem. Res. Toxicol.* 17, 463–470
24 Thompson, K.L. *et al.* (2004) Identification of platform-independent gene expression markers of cisplatin nephrotoxicity. *Environ. Health Perspect.* 112, 488–494
25 Taulan, M. *et al.* (2004) Renal toxicogenomic response to chronic uranyl nitrate insult in mice. *Environ. Health Perspect.* 112, 1628–1635
26 Waring, J. *et al.* (2002) Identifying toxic mechanisms using DNA microarrays: evidence that an experimental inhibitor of cell adhesion molecule expression signals through the aryl hydrocarbon nuclear receptor. *Toxicology* 181–182, 537–550
27 Newton, R.K. *et al.* (2004) The utility of DNA microarrays for characterizing genotoxicity. *Environ. Health Perspect.* 112, 420–422
28 Mattes, W.B. *et al.* (2004) Database development in toxicogenomics: issues and efforts. *Environ. Health Perspect.* 112, 495–505
29 Lindon, J.C. *et al.* (2003) Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. *Toxicol. Appl. Pharmacol.* 187, 137–146
30 Kriete, A. *et al.* (2003) Combined histomorphometric and gene-expression profiling applied to toxicology. *Genome Biol.* 4, R32
31 Lee, W-S. *et al.* (2004) The intelligent data management system for toxicogenomics. *J. Vet. Med. Sci.* 66, 1335–1338
32 Baker, V.A. *et al.* (2004) Clofibrate-induced gene expression changes in rat liver: a cross-laboratory analysis using membrane cDNA arrays. *Environ. Health Perspect.* 112, 428–438
33 Chu, T.M. *et al.* (2004) Cross-site comparison of gene expression data reveals high similarity. *Environ. Health Perspect.* 112, 449–455
34 Waring, J.F. *et al.* (2004) Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. *Environ. Health Perspect.* 112, 439–448
35 Goodsaid, F.M. *et al.* (2004) Quantitative PCR deconstruction of discrepancies between results reported by different hybridization platforms. *Environ. Health Perspect.* 112, 456–459
36 Dearden, J.C. (2003) *In silico* prediction of drug toxicity. *J. Comput. Aided Mol. Des.* 17, 119–127
37 Yu, H. and Adedoyin, A. (2003) ADME-Tox in drug discovery: integration of experimental and computational technologies. *Drug Discov. Today* 8, 852–861
38 Bugrim, A. *et al.* (2004) Early prediction of drug metabolism and toxicity: systems biology approach and modeling. *Drug Discov. Today* 9, 127–135
39 Giuliano, K.A. *et al.* (1997) High-content screening: a new approach to easing key bottlenecks in the drug discovery process. *J. Biomol. Screen.* 2, 249–259
40 Fenech, M. (2000) The *in vitro* micronucleus technique. *Mutat. Res.* 455, 81–95
41 Abraham, V.C. *et al.* (2004) High content screening applied to large-scale cell biology. *Trends Biotechnol.* 22, 15–22
42 Huang, K. and Murphy, R.F. (2004) From quantitative microscopy to automated image understanding. *J. Biomed. Opt.* 9, 893–912
43 Perlman, Z.E. *et al.* (2004) Multidimensional drug profiling by automated microscopy. *Science* 306, 1194–1198
44 O'Shaughnessy, J. *et al.* (2002) Treatment and prevention of intraepithelial neoplasia: an important target for accelerated new agent development. *Clin. Cancer Res.* 8, 314–346
45 Hanke, J.H. *et al.* (2004) Protein biomarkers and drug design for cancer treatments. *Eur. J. Cancer Prev.* 13, 297–305
46 Thor, A.D. *et al.* (2000) Activation (tyrosine phosphorylation) of ErbB-2 (HER-2/neu): a study of incidence and correlation with outcome in breast cancer. *J. Clin. Oncol.* 18, 3230–3239
47 Hortobagyi, G.N. (2001) Overview of treatment results with trastuzumab (Herceptin) in metastatic breast cancer. *Semin. Oncol.* 28 (Suppl. 18), 43–47
48 Motoyama, A.B. *et al.* (2002) The efficacy of ErbB receptor-targeted anticancer therapeutics is influenced by the availability of epidermal growth

factor-related peptides. *Cancer Res.* 62, 3151–3158
49 Smith, B.L. *et al.* (2004) The efficacy of Herceptin therapies is influenced by the expression of other erbB receptors, their ligands and the activation of downstream signaling proteins. *Br. J. Cancer* 91, 1190–1194
50 Minna, J.D. *et al.* (2004) A bull's eye for targeted lung cancer therapy. *Science* 304, 1458–1461
51 Paez, J.G. *et al.* (2004) EGFR mutations in lung cancer: correlation with clinical response to gefinitib therapy. *Science* 304, 1497–1500
52 Lynch, T.J. *et al.* (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350, 2129–2139
53 MacGregor, J.T. (2004) Biomarkers of cancer risk and therapeutic benefit: new technologies, new opportunities, and some challenges. *Toxicol. Pathol.* 32 (Suppl. 1), 99–105
54 Monraats, P.S. *et al.* (2004) Genetic predictive factors in restenosis. *Pathol. Biol. (Paris)* 52, 186–195
55 Posadas, E.M. *et al.* (2004) Proteomics and ovarian cancer: implications for diagnosis and treatment: a critical review of the recent literature. *Curr. Opin. Oncol.* 16, 478–484
56 Petricoin, E.R. *et al.* (2002) Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* 94, 1576–1578
57 He, Q.Y. *et al.* (2004) Diverse proteomic alterations in gastric adenocarcinoma. *Proteomics* 4, 3276–3287
58 Lofton-Day, C. and Lesche, R. (2003) DNA methylation markers in patients with gastrointestinal cancers. *Dig. Dis.* 21, 299–308
59 Player, A. *et al.* (2004) Laser capture microdissection, microarrays and the precise definition of a cancer cell. *Expert Rev. Mol. Diagn.* 4, 831–840
60 Liotta, L.A. *et al.* (2003) Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* 3, 317–325
61 Wulfkuhle, J.D. *et al.* (2003) Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase microarrays. *Proteomics* 3, 2085–2090
62 Makretsov, N.A. *et al.* (2004) Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma. *Clin. Cancer Res.* 10, 6143–6151