BIOMARKERS, VOL. 8, NO. 2 (MARCH–APRIL 2003), 79–92

Taylor & Francis
healthsciences

REVIEW
Genomics and the search for novel biomarkers in toxicology

JONATHAN D. TUGWOOD[1]*, LAURA E. HOLLINS[2] and MARK J. COCKERILL[2]

[1] Molecular Toxicology Group, Safety Assessment Department, AstraZeneca Pharmaceuticals, Alderley Park, Macclesfield, UK
[2] University of Manchester, Manchester, UK

Received 19 August 2002, revised form accepted 25 November 2002

The advent of 'genomics' technology, in particular transcript profiling, has already had a measurable impact on the drug discovery process in the areas of target identification and validation. This review is concerned with the potential application of this technology to toxicology and drug safety assessment, with particular emphasis on biomarker discovery and characterization. An advantage (or possibly a drawback!) of transcript profiling is that candidate biomarkers of toxicity can be speedily identified, with the caveat that a significant amount of subsequent experimental and bioinformatic effort needs to be expended in order to evaluate and validate them. Attention is also drawn to the critical need for robust experimental design with studies of this type and to issues associated with the analysis of large data sets. In summary, while genomics technology undoubtedly offers much that can assist drug safety assessment, its potential has yet to be realized fully in this area. However, a large amount of resource continues to be applied to 'toxicogenomics'. Tangible benefits, in terms of new biomarkers of toxicity and reduced numbers of adverse drug effects, remain realistic objectives.

Introduction

The identification, evaluation and implementation of biomarkers is crucial to the process of pharmaceutical discovery and development in a number of respects, not least of which is drug safety assessment. Existing industry-wide rates of 'attrition' of candidate drugs due to adverse toxicity in humans are unacceptably high, and there is room for a lot of improvement in our ability to detect pharmacologically related drug toxicities, both preclinically and clinically. Robust surrogate biomarkers for the major drug-induced toxicities would have a significant impact on both these areas.

Genomics technology, including the ability to quantify simultaneously many thousands of gene transcripts in a biological sample, potentially offers exciting possibilities for biomarker discovery and evaluation. This applies to all classes of biomarkers: type 0 (those that relate to pre-existing conditions or risk factors), type 1 (markers of an early pharmacological or toxic response), and type 2 (markers of disease progression). In particular, 'toxicogenomics' offers not just the possibility of determining which molecular pathways are perturbed by toxic concentrations of compounds, but also a means of exploiting this information, either for the development of novel screens or for the development of new biomarkers.

* Corresponding author: Jonathan Tugwood, Molecular Toxicology Group, Safety Assessment Department, AstraZeneca Pharmaceuticals, Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK. Tel: (+44) 1625 512868; fax: (+44) 1625 513779; e-mail: jonathan.tugwood@astrazeneca.com

Biomarkers ISSN 1354-750X print/ISSN 1366-5804 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 1354750031000070103

Genomics technology

For the purposes of this review, genomics is taken to mean large-scale detection and quantification of gene transcripts, or 'transcript profiling'. Methods for transcript profiling fall broadly into two categories: 'open' technologies, with which potentially any transcript may be analysed, and 'closed' technologies, in which the analysis is restricted to a specific number of transcripts. Examples of these are given in Table 1 (see also Farr and Dunn 1999, Pennie *et al*. 2000, Steiner and Anderson 2000, Waring and Ulrich 2000, for reviews).

The great advantage offered by all these technologies is the ability to be able to quantify hundreds or thousands of transcripts simultaneously from the same sample. If the assertion that 'no toxicity develops in the absence of changes in gene expression at some stage' is true, then there are obvious potential benefits from applying gene profiling to toxicology. These include:

i. discovery of novel toxic mechanisms
ii. development of predictive toxicity databases
iii. detection of the onset of toxicity in advance of pathological or phenotypic changes
iv. evaluation of *in vitro* models for toxicity
v. identification of preclinical and clinical biomarkers for toxicity
vi. development of screens for individuals susceptible to particular adverse drug effects.

These potential benefits have prompted a large amount of experimental work in both academic and industrial environments. The majority of published reports are 'proof of concept' experiments. While not contributing much to an overall understanding of toxicology, these studies have shown that well-known toxic compounds can bring about robust time- and concentration-dependent changes in gene expression in various biological systems (e.g. Frueh *et al*. 2001, Reilly *et al*. 2001, Ruepp *et al*. 2002, Yamazaki *et al*. 2002). Where these and other similar studies have also been useful is in highlighting the shortcomings of the 'toxicogenomics' approach, and focussing the attention of researchers on key experimental and data handling issues. These include:

i. definition and selection of the toxic endpoint
ii. selection of a relevant experimental system
iii. robust experimental design
iv. appropriate data processing and analysis
v. evaluation/'validation' of putative differentially expressed genes.

These issues are particularly pertinent with respect to the use of genomics technology for the identification of biomarkers of toxicity.

Definition and selection of toxic endpoints

While the number of organ and tissue toxicities that can occur in a test animal may appear large, they are underpinned by a relatively small number of responses at

Table 1. Technology platforms for expression profiling.

Technology	Principle	Example references or sources	Some advantages	Some disadvantages
Open technologies				
Differential display	Labelled cDNA samples from different cell populations are PCR amplified, and co-electrophoresed. cDNAs with different abundances are identified and isolated.	Liang *et al*. (1993)	Potentially any gene can be detected. Can be rapidly performed.	Downstream analysis needs to be performed to identify genes. May only detect transcripts that show high degree of differential expression, and are biased towards abundant transcripts (differential display and subtractive hybridization).
Subtractive hybridization	cDNAs from two different cell populations are subtracted from one another, usually by a PCR-mediated process. The subtraction process enriches for differentially expressed sequences.	Diatchenko *et al*. (1996)		
Serial analysis of gene expression (SAGE)	Concatenation and manual sequencing of short diagnostic sequence 'tags' from a cDNA population.	Velculescu *et al*. (1995)		
GeneCalling™	High-resolution electrophoresis followed by sequencing.	Simpson *et al*. (2000); www.curagen.com		
Closed technologies				
Nylon membrane cDNA arrays	cDNA sequences are immobilized on a nylon support, and hybridized with labelled total cDNA preparations from cells or tissues.	www.clontech.co.uk	More consistent data, as are analysing same gene sets each time. More quantitative than open techniques, and more easily amenable to downstream data processing and bioinformatic analysis.	Restricted to genes that are present on the array. Relatively inflexible – can be difficult to add more genes.
Glass cDNA arrays	cDNA sequences are immobilized on a glass slide support, and hybridized with labelled total cDNA preparations from cells or tissues. Both 'one colour' and 'two colour' systems available.	www.clontech.co.uk; www.apbiotech.com		
Oligonucleotide arrays	Complementary oligonucleotides are designed to specific genes, and immobilized on glass or synthesized *in situ*. These are hybridized with labelled cDNA preparations.	www.affymetrix.com; www.sigma-genosys.-com		

the cellular level, including necrosis, apoptosis, and stimulation or inhibition of cell proliferation (Farr and Dunn 1999). Given that these are fundamental biological processes regulated by gene activity, in principle genomics technology can be used to understand the basis of this regulation and, furthermore, to understand how this regulation is perturbed by toxic compounds. Genes whose regulation is consistently disrupted in certain toxic responses are good candidates for genetic biomarkers.

In practice, however, the characteristics of good biomarkers (Timbrell 1998) do not readily apply to genetic biomarkers of toxicity, in particular the desirability of compound or class specificity and non-invasiveness. Nonetheless, a good example of the robust association of genetic biomarkers with a toxic endpoint is observed with the peroxisome proliferator class of compounds, which are hepatotoxic in rodents but apparently not in humans (see Roberts 1999 for review). It is known that these compounds act via the ligand-activated transcription factor peroxisome proliferator activated receptor-α (PPARα) (Issemann and Green 1990). It is known that PPARα is expressed at significantly higher levels in rat and mouse liver than in human liver, and this may, at least in part, be responsible for the species differences in toxicity (Vanden Heuvel 1999). Studies on the regulation of gene expression by PPARα activators, including work with gene arrays, have repeatedly identified a number of transcript changes consistent with the known mechanism of action, that is dysregulation of peroxisomal β-oxidation of fatty acids and hepatomegaly (Baker *et al*. 2001, Thomas *et al*. 2001, Yamazaki *et al*. 2002). Importantly, many of these gene changes are also detected using primary rodent hepatocyte cultures, and the toxicity endpoints of induction of peroxisome proliferation and S-phase, and the suppression of apoptosis, can also be demonstrated in these *in vitro* systems (Hasmall *et al*. 2000). While this potentially offers a toxicity screen based on primary cell culture and gene expression changes, the important factor is the rodent specificity of these toxic effects. It is still important to be able to predict adverse *rodent* effects using rapid *in vitro* screens, to reduce the burden on *in vivo* preclinical testing. In this instance, however, it is clear that a human-based system needs to be employed, as rodents are not a good surrogate for the prediction of human toxicity. This is particularly true as new, more potent compounds are developed. For example, a recently described potent PPARα agonist upregulates *CYP4A* mRNA in human hepatocyte cultures, whereas the less potent fenofibric acid does not (Lawrence *et al*. 2001). Whether or not this can be considered as a genetic biomarker for potential human hepatotoxicity is unclear, as no other PPARα agonists have been reported to have this effect in human hepatocyte cultures and no clinical information relating to this compound is available. Although also without precedent in human hepatocyte cultures, other genetic biomarkers of peroxisome proliferator-induced toxicity may be useful in this context. In particular, biomarkers of cell proliferation such as Ki-67 and proliferating cell nuclear antigen (PCNA) may be used as early indicators of a potential compound-induced pro-proliferative effect. *In vivo* hepatomegaly is a good indicator of incipient hepatocarcinogenesis in rodents with these compounds (see Ashby *et al*. 1994 for review), and upregulation of Ki-67 and PCNA in human hepatocytes, while not necessarily indicating a human carcinogenic risk, would give cause for concern. This is an example of where a good understanding of animal

toxicity has allowed the development of genetic biomarkers, which have the potential to be applied to humans to allow assessment of novel drugs.

Toxic endpoints in extrahepatic tissues are rather less tractable at present, partly as a consequence of a lack of understanding of mechanisms and/or experimental inaccessibility, for example a lack of a good *in vitro* model. Toxicities of the central nervous system, heart and eye fall into this category, although in other areas, notably bone marrow and testis, good progress has been made in the development of *ex vivo* and *in vitro* models. For example, *in vitro* bone marrow cultures have been used for some time in attempts to predict *in vivo* bone marrow toxicity (for example, see Schoeters *et al*. 1995). The development of immortalized progenitor cells offers potentially greater utility for toxicity testing and biomarker development. For example, FDCP-Mix cells were obtained by viral immortalization of long-term bone marrow cultures, and these possess the capability of differentiating into various cell lineages depending on the growth factor environment (Spooncer *et al*. 1984, 1986). These are beginning to be exploited for toxicity evaluation of chemicals, for example with chlorinated pesticides (Henschler *et al*. 2001), and work is currently underway to attempt to obtain genomic biomarkers of bone marrow toxicity using this model.

Selection of experimental system and experimental design

Given that the overall aim is to identify genetic biomarkers that are predictive of *in vivo* animal and human toxicity, the most relevant data is obtained from *in vivo* experimental systems. Clearly this is not possible for the identification of human biomarkers, and raises certain practical difficulties for animal biomarkers. For example, it is desirable to detect gene changes that occur as an early response to a toxic insult, since these are more likely to be directly mechanistically linked and to predate any pathological change. This would require a complex time course and dose–response study to be set up, with a need for large numbers of animals and posing logistic problems with dosing and sampling. Also, such a study would be subject to considerable inter-animal variation, both due to differences in exposure to the dosed compound and to the biological variation that occurs even between animals of inbred strains. It is also common that toxic events are restricted to subpopulations of cells within an organ or tissue type. Therefore, tissue samples may consist of a majority of non-responding cells, which could dilute and obscure the key regulatory changes in the cell population of interest. Techniques such as laser capture microdissection have been used successfully to enrich for cell populations of interest in array studies (Leethanakul *et al*. 2000), but this is labour-intensive, particularly when large numbers of samples are involved. Nonetheless, data exists to support the distinction of different classes of hepatotoxins based on *in vivo* transcript profiling and the correlation of specific gene changes with histological and clinical chemistry measurements (Waring *et al*. 2001, Hamadeh *et al*. 2002). Interestingly, the Waring *et al*. (2001) study employed a single dose/single time point design, which, as discussed above, is suboptimal. The rationale was to use a dose of each of the 15 test compounds that produced liver pathology changes after 7 days, and then to take samples after 3 days' exposure. This intermediate time point was selected in the hope of capturing early,

mechanism-related gene changes while avoiding complex later changes that may be related to pro-inflammatory, necrotic or fibrotic events. As well as simplifying experimental design and reducing cost, Waring *et al.* (2001) argued that in order for this type of approach to be useful for screening large numbers of compounds, it needs to be sufficiently robust to provide useful data at a single 'snapshot' time point. This simplified approach was justified in the sense that mechanism-related gene changes were identified for the well-characterized hepatotoxins included in the study, and related compounds were shown to be 'clustered' together. It could be argued that this is hardly surprising, and that the definitive test of a predictive gene expression database is the ability to detect the toxicity of 'new' compounds, that is those with an aetiology that is not represented in the database. This will require not only an extensive database, but also robust predictive models based on the data. A number of companies are developing such databases and predictive modelling tools (see below). The Waring *et al.* (2001) and Hamadeh *et al.* (2002) studies use liver *in vivo* as a model system to attempt to 'prove the principle' that toxicity endpoints can be associated with specific gene expression profiles. It could be argued that liver is structurally more 'simple' than other toxicity target organs, and therefore more amenable to such analysis. It will be a greater challenge to demonstrate similar correlations for other tissues, where toxic mechanisms are also less well understood at the gene expression level.

Some of the difficulties with *in vivo* studies can be addressed by the use of primary cell cultures and cell lines rather than intact animals. However, this immediately raises the issue of the relevance of any *in vitro* findings to the toxic response *in vivo*. It is documented that, in hepatocytes at least, the isolation process upregulates 'immediate-early' transcription factors such as c-fos and c-jun (Rana *et al*. 1994), which as potent regulatory molecules themselves have profound effects on 'downstream' gene expression. Also in hepatocytes, a number of investigations have noted a dedifferentiation process with time in culture, with a concomitant loss of liver-specific functions. Liver-derived cell lines have been used successfully in some situations to recapitulate *in vivo* toxicity (Bayly *et al*. 1993), although many lines are tumour-derived, karyotypically abnormal and often refractory to the effects of known hepatotoxins.

In selecting the appropriate experimental system for array experiments, whether *in vivo* or *in vitro*, the important criterion is to be able to demonstrate and measure the onset of the compound-related toxic effect(s) by means other than transcript profiling. This will allow the differential gene expression information to be placed in context with respect to the toxic endpoint, which is critical to determining which changes may be causally related.

There are a number of important issues to consider when designing array experiments in order to enable the experimenter to have confidence that the observed gene expression changes are genuine and not artefactual. A detailed discussion of these issues is not appropriate here, but in general the number of replicate arrays run per sample can have a significant effect on data reliability. For cDNA arrays, running three replicate arrays has been shown to improve substantially the detection of false negatives and false positives (Lee *et al*. 2000). Reliance on 'fold-change' measurements to quantify gene expression changes can

also cause difficulties, since the magnitude of a fold-change is highly dependent on transcript abundance. The use of appropriate statistical models, both to determine the number of replicates required and to ascribe 'significance' to observed expression changes (Wolfinger *et al*. 2001), is an improvement on ranking differentially expressed genes simply on the basis of fold-changes. Finally, our own experience and that of others indicates that there are several sources of variation in array experiments, some of which may be unforeseen. These include variation due to operator, date of RNA preparation, date of hybridization, and reagent batch. Tools such as principle component analysis can help identify such variation when it occurs, but the preferred solution is to reduce the effect of this by rational experimental design. In essence, this requires arranging the experimental work so that, for example, 'control' and 'drug-treated' samples are not processed on different days.

Identifying biomarkers – analytical issues with transcript profiles

The large volume of information generated by transcript profile studies presents the experimenter with analytical problems. Whereas a relatively simple approach using spreadsheets can perform a rudimentary analysis, for example sorting genes according to fold-change and performing analysis of variance, specialist tools are required for more complex evaluation. Proprietary analysis packages such as Spotfire (Spotfire, Inc.) and GeneSpring (Silicon Genetics) give the investigator a large number of options, including 'clustering' genes with similar or (dissimilar) expression patterns to identify genes that may be co-ordinately regulated (Eisen *et al*. 1998).

As mentioned previously, a desirable outcome of transcript profiling experiments is the creation of a database of gene profiles associated with toxic endpoints, both from the point of view of predicting toxicity and identifying new biomarkers. One such database currently being compiled is ToxExpress™ (Gene Logic Inc.; www.genelogic.com). This programme is focussing initially on liver toxicity, and a large number of hepatotoxins are being used in *in vivo* and *in vitro* transcript profiling studies. One of the goals is to build predictive models of liver toxicity, but another important anticipated outcome is the identification of genes or clusters of genes predictive of certain types of toxicity that can be employed as biomarkers in screening experiments. A similar approach has recently yielded a number of potential new markers of pancreatic cancer, as well as confirming some previously known gene changes associated with this disease (Iacobuzio-Donahue *et al*. 2002).

Confirmation and 'validation' of putative differentially expressed genes

A typical gene array experiment will generate several hundred potential gene changes, and data analysis of the type referred to above will help the experimenter prioritize which of these are associated with chemical concentration and/or time course of administration. Much has been written of the desirability of confirming these putative gene changes by some independent means, such as quantitative reverse transcription-polymerase chain reaction (RT-PCR). Indeed, peer-reviewed journals will not publish 'lists of genes' from gene profiling experiments without such corroboration.

'Validation' when used in this context implies more than simple confirmation of a change in transcript abundance, but a robust association of the gene product with the toxic endpoint. It is useful to provide a reminder here of what is being measured with transcript profiling, that is steady-state mRNA levels. As illustrated in Figure 1, an increase in steady-state mRNA abundance should not necessarily be taken to imply an increase in protein abundance, still less an increase in enzyme activity. This serves to exemplify one of the shortcomings of using gene profiling to generate biomarkers for toxicity: a gene identified as a putative biomarker on the basis of a transcript profile may not 'translate' into something useful as an indicator of a toxic endpoint. There are good examples of increases in mRNA abundance not being reflected by corresponding protein increases, and of protein/enzyme activity being increased by post-translation mechanisms while mRNA levels remain unchanged (Anderson and Seilhamer 1997). Also, the temporal relationship between gene transcription and the appearance of active protein is often complex.

The main test of the 'validity' of a genetic biomarker is whether alteration in gene activity, whether measured at the mRNA, protein or activity level, is consistently associated with the appearance of a toxic endpoint. Such a biomarker is especially useful if alteration can be detected irrespective of the toxin, although a biomarker that is toxin-specific or toxin-class-specific nonetheless has significant utility.

Use of suppression subtractive hybridization technology to identify genetic biomarkers of drug-induced liver toxicity in dogs

As an example of the use of genomics technology in biomarker identification, the following experiments were undertaken to identify gene regulatory changes associated with drug-induced hepatosteatosis in dogs. Steatosis and hepatosteatosis refer to the accumulation of fat in the liver and can be one of the consequences of alcohol abuse. Non-alcoholic steatosis, however, is associated with a number of factors, including mitochondrial dysfunction (Pessayre *et al*. 2002), oxidative stress (Robertson *et al*. 2001) and altered macrophage function (Diehl 2002).

A number of pharmaceutical agents, including amiodarone and perhexiline, can induce these types of changes, resulting in hepatic steatosis (Berson *et al*. 1998). The early molecular changes that underlie the development of hepatic steatosis are

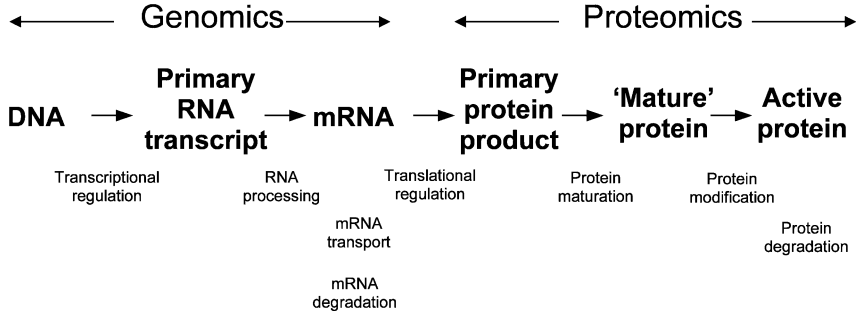

Figure 1. A simplified linear representation of the process of gene expression, from structural gene to active protein, indicating some of the complex, independently regulated processes involved.

unclear, and the aim of this study was to elucidate the molecular pathways that are perturbed by a steatotic agent and to identify biomarkers that may potentially be used in primary hepatocyte cultures to identify possible hepatosteatotic agents.

The candidate pharmaceutical #6021 has been shown to cause focal necrosis and microvesicular steatosis in the livers of treated dogs. Such changes were not observed in rodent studies, and it is likely that the apparent canine-specific effect is due to differences in exposure. For these experiments, we chose to use dog primary hepatocyte cultures, since a dog *in vivo* study was felt to be impractical for generating the early time-course material required in these experiments. Also, use of this *in vitro* system should overcome the difficulties and sources of variation associated with compound administration and inter-animal variation.

The lack of annotated gene information in dogs has up to now precluded the development of dog gene arrays for transcript profiling, so the technique of suppression subtractive hybridization (SSH) was employed for these studies. SSH is an improvement over conventional subtractive hybridization in that only two rounds of subtraction are required and the likelihood of isolating differentially expressed rare cDNAs is enhanced (Diatchenko *et al*. 1996).

Initially, the cytotoxicity of #6021 in dog primary hepatocyte cultures was determined in time course and dose–response studies using Alamar Blue fluorescence (Nakayama *et al*. 1997) (Figure 2) to select appropriate concentrations for the SSH experiments. Toxic and subtoxic concentrations of 20 μM and 2 μM #6021, respectively, were selected. After exposure of hepatocyte cultures to these concentrations for 24 h, the accumulation of lipid was demonstrated within the cells using Nile Red staining (McMillian *et al*. 2001) (data not shown). Separate cultures were exposed to these concentrations for 2 and 8 h; these time points were selected in order to maximize the chances of obtaining early

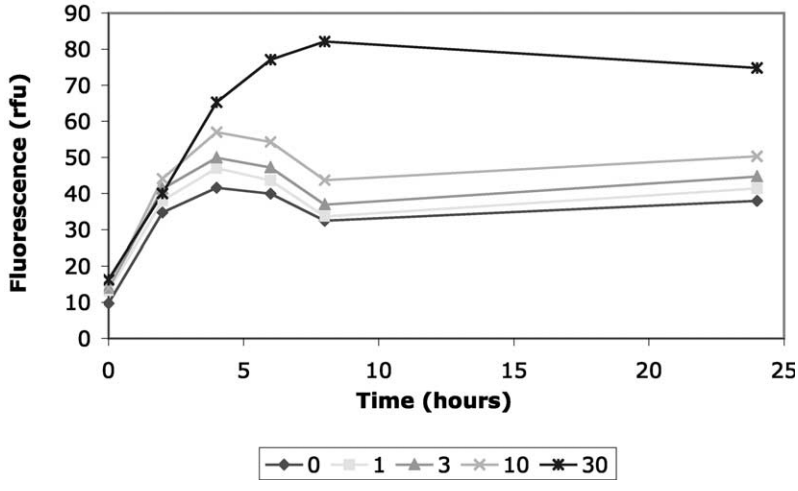

Figure 2. Cytotoxicity of #6021 in primary dog hepatocytes. The key below refers to the concentration of #6021 in μM. Cells were exposed to the indicated concentration of compound in the presence of Alamar Blue reagent (Nakayama *et al*. 1997), and the fluorescence in relative fluorescence units (rfu) measured at 590 nm after 0, 1, 2, 4, 8 and 24 h to obtain a quantitative indication of cellular toxicity.

mechanistically relevant biomarkers. cDNA was prepared from treated cultures, and the 20 μM #6021 samples were pooled and subtracted against cDNA from vehicle-treated cultures. Both 'forward' and 'reverse' subtractions were performed to allow isolation of both up- and downregulated transcripts. To determine whether the subtraction had been successful, the relative levels of the abundant glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA in subtracted and unsubtracted populations were determined by PCR (Figure 3). This indicated that less than 0.01% of the GAPDH cDNA remained following the SSH process. Putative differentially expressed cDNAs were cloned into plasmid vectors, and 200 clones (100 each putative up- and downregulated) were randomly selected for sequencing. Some of the sequenced clones that could be identified from public sequence databases are presented in Table 2. In addition, 33 clones were isolated that had no matches in the database, and these represent potential novel biomarkers that are currently being investigated further. The identified sequences included some genes that are known to be abundantly expressed, for example albumin and ribosomal proteins, indicating that the SSH process does not completely remove common sequences. However, some of the sequenced cDNAs were from genes known from the published literature to be associated with hepatosteatosis, for example genes for cytochrome P450 2E1 (CYP2E1), stearoyl coenzyme A (stearoyl-CoA) desaturase and apolipoprotein-AII (apo-AII). These were among those selected for evaluation using RT-PCR. For these experiments, RNA samples from cultures treated for 2, 8 and 24 h with either 2 or 20 μM #6021 were analysed (Table 3). Interestingly, the magnitude of the changes determined by RT-PCR was relatively small – it is doubtful whether these types of changes would have been detected using conventional subtractive hybridization. In most cases the RT-PCR results correlated qualitatively with the SSH data, for example apo-AII being downregulated, CYP2E1 and apolipoprotein B-100 (apoB-100) being upregulated. The exception

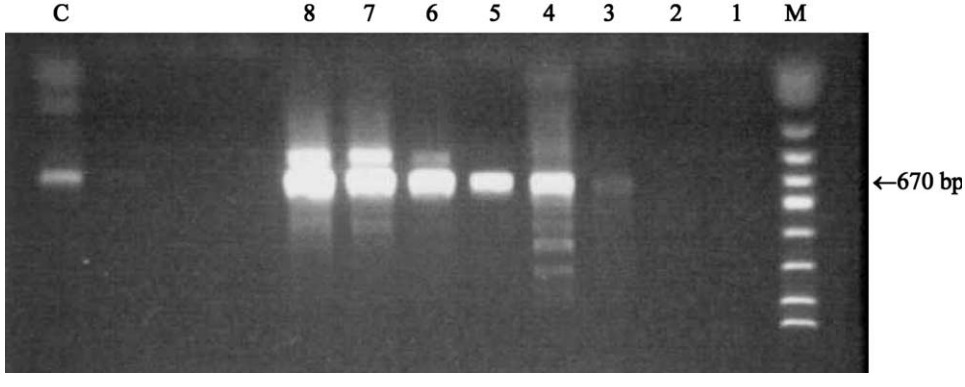

Figure 3. Verification of subtraction of common cDNA sequences. Equal amounts of subtracted (lanes 1–4) and unsubtracted (lanes 5–8) dog liver cDNAs were analysed for the presence of the common abundant sequence GAPDH using PCR. Lanes 1 and 5, 18 PCR cycles; lanes 2 and 6, 23 cycles; lanes 3 and 7, 28 cycles; lanes 4 and 8, 33 cycles. The intensity of the band in lane 4 is between that of lanes 5 and 6, indicating that approximately 13 more PCR cycles are required to obtain equal GAPDH amplification in subtracted versus unsubtracted cDNA. This suggests that GAPDH is greater than 99.99% subtracted using this procedure. M, DNA markers; C, control (GAPDH PCR product from dog liver cDNA).

Table 2. Putative differentially expressed clones identified by SSH.

Description	No. of sequences
UPREGULATED CLONES	
Translationally controlled tumour-associated protein 1 (TCTP-1)	26
Ribosomal S27a protein	18
Albumin	11
Vitronectin	6
Cytochrome P450 2E1 (CYP2E1)	5
Serum response factor (SRF) accessory protein 1A	5
Apolipoprotein B-100 (apoB-100)	5
p21-Waf1	4
Insulin-like growth factor binding protein-1 (IGF-BP1)	4
Ribosomal S25 protein	4
Nuclear factor of activated T-cells-1 (NFAT-1)	4
Gene 33 polypeptide	3
Bile salt export pump (BSEP)	2
Tumour necrosis factor (TNF) regulating transmembrane protein	2
α_1-Antitrypsin	2
Phospholipase D2	2
Growth arrest and DNA damage induced-45 (GADD45)	1
Manganese superoxide dismutase	1
Metastasis inhibitory protein	1
Serum and glucocorticoid-regulated kinase (sgk-2) Ser-Thr kinase	1
DOWNREGULATED CLONES	
Tissue inhibitor of metalloproteinase-3 (TIMP-3)	22
Uridine disphosphate glucuronosyltransferase (UDP-GT)	12
N-Cadherin	12
Calmodulin	11
Glutamate dehydrogenase	5
Stearoyl coenzyme A desaturase	5
Apolipoprotein-AII	5
Ribosomal protein L30	4
Thrombospondin	4
Activated C kinase	4
Rab14 guanosine triphosphatase (GTPase)	3
Albumin	2
F box protein 9 (FBX9)	2
Ribosomal protein S6	2
R-Cadherin	2
Heat shock protein 70 (HSP70)	1
Cathepsin Z	1
Glucocerebrosidase	1
Annexin A2	1
Cathepsin Y	1

Upregulated clones were isolated following 'forward' subtraction (#6021-treated minus untreated); downregulated clones were isolated following 'reverse' subtraction (untreated minus #6021-treated). The number of sequences indicates the number of times the clone was independently isolated using SSH. Clones shown in bold are abundant, presumably common sequences that have come through the subtraction process; clones underlined are those that have been identified in the literature as being associated with hepatosteatosis. The clones shown are those that gave robust identifications in the European Molecular Biology Laboratory (EMBL) sequence database. Other clones could not be identified and represent potentially novel differentially expressed sequences.

was stearoyl-CoA desaturase, which was slightly upregulated in the RT-PCR experiments whereas SSH suggested a downregulation by #6021.

In summary, in these experiments the SSH technique facilitated the identification of a number of potential genetic biomarkers, including some novel ones, which would have been difficult by other means given the paucity of gene sequence

Table 3. RT-PCR analysis of clones identified by SSH.

Clone	#6021 (μM)	Treatment duration 2 h	8 h	24 h
Apolipoprotein-AII	2	1.3	*0.5*	**2.8**
	20	*0.6*	**1.5**	**1.7**
Apoprotein B-100	2	**2**	0.9	**2**
	20	**1.5**	**2.3**	**1.6**
Cytochrome P450 2E1 (CYP2E1)	2	**1.5**	0.8	*0.5*
	20	1	**4.2**	*0.5*
Manganese superoxide dismutase	2	1.2	0.7	1.2
	20	1	**1.7**	**1.5**
Stearoyl coenzyme A desaturase	2	0.8	0.7	**2**
	20	1.3	1.4	1.3
Translationally controlled tumour-associated protein (TCTP)	2	1.1	0.7	**2**
	20	0.9	1.7	**1.6**
Transmembrane protein induced by tumour necrosis factor-α (TMPIT)	2	1	0.7	*0.6*
	20	0.9	**1.9**	**2.2**

The values given indicate fold-changes relative to vehicle controls at the same time point. Clones that were up- or downregulated by greater than 1.5-fold are indicated by bold and italic type, respectively.

information in the dog. A parallel series of experiments is currently in progress using rat primary hepatocytes and gene arrays, which should help with corroboration of the dog SSH results.

Conclusion

Although the concept of toxicogenomics has been around for a few years, the real impact of this technology has still to be felt in drug safety assessment in terms of facilitating toxicity input early in drug discovery programmes and in reducing the 'attrition rate' of candidate drugs failing due to adverse animal toxicity. This can be attributed in part to technology issues such as inter-experiment variation, the inter-animal variation that has always hindered interpretation of toxicity studies, and problems associated with handling and analysing large data sets. However, many within the field remain optimistic that the application of genomics technology to toxicology will provide useful benefits, not least in the delivery of robust mechanism-associated genetic biomarkers for toxic endpoints in preclinical and clinical situations. This view is supported by the huge investment that continues to be made in toxicogenomics, coupled with encouraging reports at conferences and in the literature. Although companies have hitherto been understandably reluctant to use transcript profile data to support drug regulatory packages, this may change when there is a general acceptance in the wider scientific community of the reproducibility and robustness of this type of data. International collaborative pro-

grammes such as the International Life Sciences Institute (ILSI) sponsored committee on the 'Application of Genomics Technology to Mechanism-Based Risk Assessment' should help bring this about. Encouragingly, there are signs that the regulatory authorities themselves are adopting an inclusive attitude to genomics technology.

References

Anderson, L. and Seilhamer, J. 1997, A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis*, **18**, 533–537.

Ashby, J., Brady, A., Elcombe, C. R., Elliott, B. M., Ishmael, J., Odum, J., Tugwood, J. D., Kettle, S. and Purchase, I. F. 1994, Mechanistically-based human hazard assessment of peroxisome proliferator-induced hepatocarcinogenesis. *Human and Experimental Toxicology*, **13**, S1–S117.

Baker, T. K., Carfagna, M. A., Gao, H., Dow, E. R., Li, Q. Q., Searfoss, G. H. and Ryan, T. P. 2001, Temporal gene expression analysis of monolayer cultured rat hepatocytes. *Chemical Research in Toxicology*, **14**, 1218–1231.

Bayly, A. C., French, N. J., Dive, C. and Roberts, R. A. 1993, Non-genotoxic hepatocarcinogenesis *in vitro*: the FAO hepatoma line responds to peroxisome proliferators and retains the ability to undergo apoptosis. *Journal of Cell Science*, **104**, 307–315.

Berson, A., De Beco, V., Letteron, P., Robin, M. A., Moreau, C., El Kahwaji, J., Verthier, N., Feldmann, G., Fromenty, B. and Pessayre, D. 1998, Steatohepatitis-inducing drugs cause mitochondrial dysfunction and lipid peroxidation in rat hepatocytes. *Gastroenterology*, **114**, 764–774.

Diatchenko, L., Lau, Y. F. C., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D. and Siebert, P. D. 1996, Suppression subtractive hybridization – a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences of the USA*, **93**, 6025–6030.

Diehl, A. M. 2002, Nonalcoholic steatosis and steatohepatitis IV. Nonalcoholic fatty liver disease abnormalities in macrophage function and cytokines. *American Journal of Physiology: Gastrointestinal and Liver Physiology*, **282**, G1–G5.

Eisen, M. B., Spellman, P. T., Brown, P. O. and Botstein, D. 1998, Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the USA*, **95**, 14863–14868.

Farr, S. and Dunn II, R. T. 1999, Concise review: gene expression applied to toxicology. *Toxicological Sciences*, **50**, 1–9.

Frueh, F. W., Hayashibara, K. C., Brown, P. O. and Whitlock, J. P. 2001, Use of cDNA microarrays to analyze dioxin-induced changes in human liver gene expression. *Toxicology Letters*, **122**, 189–203.

Hamadeh, H. K., Bushel, P. R., Jayadev, S., Martin, K., DiSorbo, O., Sieber, S., Bennett, L., Tennant, R., Stoll, R., Barrett, J. C., Blanchard, K., Paules, R. S. and Afshari, C. A. 2002, Gene expression analysis reveals chemical-specific profiles. *Toxicological Sciences*, **67**, 219–231.

Hasmall, S. C., James, N. H., Macdonald, N., Soames, A. R. and Roberts, R. A. 2000, Species differences in response to diethylhexylphthalate: suppression of apoptosis, induction of DNA synthesis and peroxisome proliferator activated receptor alpha-mediated gene expression. *Archives of Toxicology*, **74**, 85–91.

Henschler, R., Appel, K. E., Heyworth, C. M. and Glatt, H. 2001, Proliferation and differentiation of murine haemopoietic progenitor cells in stroma-free culture in the presence of metabolites of chlorinated pesticides. *Toxicology In Vitro*, **15**, 31–37.

Iacobuzio-Donahue, C. A., Maitra, A., Shen-Ong, G. L., van Heek, T., Ashfaq, R., Meyer, R., Walter, K., Berg, K., Hollingsworth, M. A., Cameron, J. L., Yeo, C. J., Kern, S. E., Goggins, M. and Hruban, R. H. 2002, Discovery of novel tumor markers of pancreatic cancer using global gene expression technology. *American Journal of Pathology*, **160**, 1239–1249.

Issemann, I. and Green, S. 1990, Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature*, **347**, 645–650.

Lawrence, J. W., Li, Y., Chen, S. Y., DeLuca, J. G., Berger, J. P., Umbenhauer, D. R., Moller, D. E. and Zhou, G. C. 2001, Differential gene regulation in human versus rodent hepatocytes by peroxisome proliferator-activated receptor (PPAR) alpha – PPAR alpha fails to induce peroxisome proliferation-associated genes in human cells independently of the level of receptor expression. *Journal of Biological Chemistry*, **276**, 31521–31527.

Lee, M. L. T., Kuo, F. C., Whitmore, G. A. and Sklar, J. 2000, Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proceedings of the National Academy of Sciences of the USA*, **97**, 9834–9839.

Leethanakul, C., Patel, V., Gillespie, J., Pallente, M., Ensley, J. F., Koontongkaew, S., Liotta, L. A., Emmert-Buck, M. and Gutkind, J. S. 2000, Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene*, **19**, 3220–3224.

Liang, P., Averboukh, L. and Pardee, A. B. 1993, Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Research*, **21**, 3269–3275.

McMillian, M. K., Grant, E. R., Zhong, Z., Parker, J. B., Li, L., Zivin, R. A., Burczynski, M. E. and Johnson, M. D. 2001, Nile Red binding to HepG2 cells: an improved assay for *in vitro* studies of hepatosteatosis. *In Vitro and Molecular Toxicology*, **14**, 177–190.

Nakayama, G. R., Caton, M. C., Nova, M. P. and Parandoosh, Z. 1997, Assessment of the Alamar Blue assay for cellular growth and viability *in vitro*. *Journal of Immunological Methods*, **204**, 205–208.

Pennie, W. D., Tugwood, J. D., Oliver, G. J. A. and Kimber, I. 2000, The principles and practice of toxicogenomics: applications and opportunities. *Toxicological Sciences*, **54**, 277–283.

Pessayre, D., Mansouri, A. and Fromenty, B. 2002, Nonalcoholic steatosis and steatohepatitis v. mitochondrial dysfunction in steatohepatitis. *American Journal of Physiology: Gastrointestinal and Liver Physiology*, **282**, G193–G199.

Rana, B., Mischoulon, D., Xie, Y., Bucher, N. L. and Farmer, S. R. 1994, Cell–extracellular matrix interactions can regulate the switch between growth and differentiation in rat hepatocytes: reciprocal expression of C/EBP alpha and immediate-early growth response transcription factors. *Molecular and Cellular Biology*, **14**, 5858–5869.

Reilly, T. P., Bourdi, M., Brady, J. N., Pise-Masison, C. A., Radonovich, M. F., George, J. W. and Pohl, L. R. 2001, Expression profiling of acetaminophen liver toxicity in mice using microarray technology. *Biochemical and Biophysical Research Communications*, **282**, 321–328.

Roberts, R. A. 1999, Peroxisome proliferators: mechanisms of adverse effects in rodents and molecular basis for species differences. *Archives of Toxicology*, **73**, 413–418.

Robertson, G., Leclercq, I. and Farrell, G. C. 2001, Nonalcoholic steatosis and steatohepatitis. II. Cytochrome P-450 enzymes and oxidative stress. *American Journal of Physiology: Gastrointestinal and Liver Physiology*, **281**, G1135–G1139.

Ruepp, S. U., Tonge, R. P., Shaw, J., Wallis, N. and Pognan, F. 2002, Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. *Toxicological Sciences*, **65**, 135–150.

Schoeters, G. E. R., Vander Plaetse, F., Leppens, H., Van Vlasselaer, P. and Van den Heuvel, R. 1995, Haemopoietic and osteogenic toxicity testing *in vitro* using murine bone marrow cultures. *Toxicology In Vitro*, **9**, 421–428.

Simpson, J. W., Ruiz-Martinez, M. C., Mulhern, G. T., Berka, J., Latimer, D. R., Ball, J. A., Rothberg, J. M. and Went, G. T. 2000, A transmission imaging spectrograph and microfabricated channel system for DNA analysis. *Electrophoresis*, **21**, 135–149.

Spooncer, E., Boettiger, D. and Dexter, T. M. 1984, Continuous *in vitro* generation of multipotential stem cell clones from src-infected cultures. *Nature*, **310**, 228–230.

Spooncer, E., Heyworth, C. M., Dunn, A. and Dexter, T. M. 1986, Self-renewal and differentiation of interleukin-3-dependent multipotent stem cells are modulated by stromal cells and serum factors. *Differentiation*, **31**, 111–118.

Steiner, S. and Anderson, N. L. 2000, Expression profiling in toxicology – potentials and limitations. *Toxicology Letters*, **112**, 467–471.

Thomas, R. S., Rank, D. R., Penn, S. G., Zastrow, G. M., Hayes, K. R., Pande, K., Glover, E., Silander, T., Craven, M. W., Reddy, J. K., Jovanovich, S. B. and Bradfield, C. A. 2001, Identification of toxicologically predictive gene sets using cDNA microarrays. *Molecular Pharmacology*, **60**, 1189–1194.

Timbrell, J. A. 1998, Biomarkers in toxicology. *Toxicology*, **129**, 1–12.

Vanden Heuvel, J. P. 1999, Peroxisome proliferator-activated receptors (PPARS) and carcinogenesis. *Toxicological Sciences*, **47**, 1–8.

Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. 1995, Serial analysis of gene expression. *Science*, **270**, 484–487.

Waring, J. F. and Ulrich, R. G. 2000, The impact of genomics-based technologies on drug safety evaluation. *Annual Review of Pharmacology and Toxicology*, **40**, 335–352.

Waring, J. F., Jolly, R. A., Ciurlionis, R., Lum, P. Y., Praestgaard, J. T., Morfitt, D. C., Buratto, B., Roberts, C., Schadt, E. and Ulrich, R. G. 2001, Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicology and Applied Pharmacology*, **175**, 28–42.

Wolfinger, R. D., Gibson, G., Wolfinger, E. D., Bennett, L., Hamadeh, H., Bushel, P., Afshari, C. and Paules, R. S. 2001, Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biology*, **8**, 625–637.

Yamazaki, K., Kuromitsu, J. and Tanaka, I. 2002, Microarray analysis of gene expression changes in mouse liver induced by peroxisome proliferator activated receptor alpha agonists. *Biochemical and Biophysical Research Communications*, **290**, 1114–1122.