Review

Practical Application of Toxicogenomics for Profiling Toxicant-Induced Biological Perturbations

Naoki Kiyosawa [1,]*, **Sunao Manabe** [2], **Takashi Yamoto** [1] **and Atsushi Sanbuissho** [1]

[1] Medicinal Safety Research Laboratories, Daiichi Sankyo Co., Ltd., 717 Horikoshi, Fukuroi, Shizuoka 437-0065, Japan; E-Mails: yamoto.takashi.vw@daiichisankyo.co.jp (T.Y.); sanbuissho.atsushi.tv@daiichisankyo.co.jp (A.S.)

[2] Global Project Management Department, Daiichi Sankyo Co., Ltd., 1-2-58, Hiromachi, Shinagawa, Tokyo 140-8710, Japan; E-Mail: manabe.sunao.u2@daiichisankyo.co.jp (S.M)

* Author to whom correspondence should be addressed;
E-Mail: kiyosawa.naoki.wr@daiichisankyo.co.jp; Tel.: +81-538-42-4356;
Fax: +81-538-42-4350.

Abstract: A systems-level understanding of molecular perturbations is crucial for evaluating chemical-induced toxicity risks appropriately, and for this purpose comprehensive gene expression analysis or toxicogenomics investigation is highly advantageous. The recent accumulation of toxicity-associated gene sets (toxicogenomic biomarkers), enrichment in public or commercial large-scale microarray database and availability of open-source software resources facilitate our utilization of the toxicogenomic data. However, toxicologists, who are usually not experts in computational sciences, tend to be overwhelmed by the gigantic amount of data. In this paper we present practical applications of toxicogenomics by utilizing biomarker gene sets and a simple scoring method by which overall gene set-level expression changes can be evaluated efficiently. Results from the gene set-level analysis are not only an easy interpretation of toxicological significance compared with individual gene-level profiling, but also are thought to be suitable for cross-platform or cross-institutional toxicogenomics data analysis. Enrichment in toxicogenomics databases, refinements of biomarker gene sets and scoring algorithms and the development of user-friendly integrative software will lead to better evaluation of toxicant-elicited biological perturbations.

1. Introduction

The term 'toxic agent' can be defined as any substance that causes harmful effects on living organisms, but in general, such hazardous effects are substantially dependent on the chemical's exposure level. For instance, dietary salt may cause nephrotoxicity if an extreme amount is ingested at any time; however, do we consider sodium chloride a toxic agent? This is not the case because ordinarily we only need a spoon of salt for cooking and that amount will not harm healthy human bodies. In addition, in the case of pharmaceutical drugs, any chemical can be toxic when overdosed, and what makes them beneficial or toxic depends on their dose levels. Various types of preclinical toxicity studies are required before starting a human clinical trial to collect information on the toxicological profile such as target organs of toxicities, dose-response profile, recovery, toxicokinetics, genotoxicity, teratogenicity and carcinogenicity, using a sufficient number of experimental animals. However, even after a long and costly preclinical toxicity evaluation, species difference in mechanism of action (MOA) between humans and animals sometimes brings about unpredictable toxicities in humans [1]. Furthermore, a chemical once regarded as non-toxic could cause exaggerated toxicity in case undesirable chemical-chemical interaction occurs through inhibition or induction of ABC transporters [2] or hepatic drug metabolizing enzymes [3]. To evaluate and manage the potential hazardous risks of the chemicals, it is desirable to gain insight into the chemical-induced MOA in the target organ of toxicity, by which we can appropriately evaluate whether the chemical should be regarded as toxic or not according to its overall risk/benefit property.

For living organisms, it is sometimes necessary to modulate biological homeostasis to overcome potential hazardous effects caused by chemical exposure. For example, administration of anticonvulsant phenobarbital to rats causes liver hypertrophy, which is associated with induction of hepatic drug metabolizing enzymes such as CYP2B and CYP3A through the activation of the nuclear receptor CAR [4]. Although an increase in liver weight may look like a deleterious perturbation of hepatic homeostasis, toxicologists usually regard it as a non-toxic but rather a desirable response or adaptive response for the body. This is because such hepatic enzyme induction facilitates efficient metabolism and disposition of the exposed chemicals. In a systems-biological point of view, such liver system is called 'robust', not 'homeostatic' against phenobarbital exposure [5], where biological systems are not static but exhibit dynamic molecular reconstructions to maintain their cellular functional integrity. From this perspective, drug-induced toxicity can be defined as a 'collapse of the biological robustness by drug exposure'. The recent advancement in Systems Biology is supported by dramatic advances in functional genomics techniques, especially the microarray technique by which expression levels of tens of thousands of genes can be measured simultaneously. Application of the microarray technique to toxicology research is called toxicogenomics (TGx), and is now widely utilized by pharmaceutical scientists in drug development [6]. The major problem in utilizing the TGx technique lies in its huge data size, as well as the complexity of the systems-level molecular interactions. We cannot avoid performing multivariate analyses to handle huge data sets, but

toxicologists usually struggle to implement complicated statistical analysis. To overcome such difficulties, an easy, simple and practical analytical flow is desired. In this paper, we present practical methods of TGx data analysis for profiling chemical-elicited molecular perturbations using an open source analytical software, which will lead to better and easier utilization of TGx data to understand the MOA of toxicities.

2. Advancement of Toxicogenomics

### 2.1. Toxicogenomic Biomarker Gene Sets

In 2001, it was reported that the hepatic gene expression profiles in rats following treatment with various chemicals showed clear chemical-specific patterns when measured with the microarray technique [7,8]. Such chemical-specific changes in the transcriptome profile leads to changes in the proteome profile, the metabolome profile and eventually the tissue-level phenotypes. Thus, it is natural that the transcriptome profile would contain a significant degree of information for biological conditions at the moment, which may lead to a profound understanding of chemical-induced molecular perturbations. However, such chemical-specific gene expression data contain mixed molecular events that reflect complicated interactions among biological pathways such as xenobiotic metabolism, stress response, energy metabolism, protein synthesis/degradation, mRNA transcription/degradation, DNA repair/replication, cell proliferation/cell death control, *etc*. Microarray analysis measures tens of thousands of gene expression levels simultaneously, and is usually too complicated to appropriately interpret the significance of the gene expression changes at a time. Instead, it would be more practical to focus on the data for certain gene sets whose expression levels are closely associated with certain biological functions like glycolysis and cell proliferation. Such gene sets can be prepared from public information such as PubMed literature search, Gene Ontology [9], KEGG [10] and GenMAPP [11] biological pathway information. In addition, a number of gene sets have been reported whose expression levels are closely associated with certain toxicological endpoints, or TGx biomarker gene sets [12] such as cell injury [13], carcinogenicity [14,15], phospholipidosis [16,17] and glutathione depletion [18,19]. These TGx biomarkers can then be utilized for evaluation, diagnosis or prediction of toxicity based on their expression changes. As shown in Figure 1, it is much more informative and easy to interpret the microarray data by focusing on certain gene sets rather than observing whole data sets (*i.e.*, >30,000 gene probe sets in the case of Affymetrix GeneChip system). However, it is still too complicated when we need to handle a number of biomarker gene sets at a time. Because pharmaceutical toxicologists are usually not experts in handling vast amounts of data sets, it is crucial to develop user-friendly analytical software whose output is clear enough to interpret the toxicological significance.

2.2. Public and Commercial Microarray Database

A high quality and large-scale reference microarray database is desired for an appropriate interpretation of TGx data. A number of public databases are currently available, such as Gene Expression Ominibus (GEO) [20], ArrayExpress [21], Chemical Effects in Biological Systems (CEBS) [22], Comparative Toxicogenomics Database (CTD) [23] or EDGE [24]. In addition, the

Toxicogenomics Project in Japan (http://wwwtgp.nibio.go.jp/index.html) and the InnoMed PredTox Consortium (http://www.innomed-predtox.com/) developed large-scale toxicogenomic databases, both of which contain microarray datasets for well-studied toxicants, as well as proprietary drugs using both *in vivo* and *in vitro* systems.

Figure 1. Expression profiling for toxicogenomic biomarker gene sets. Gene sets whose expression levels are closely associated with cell proliferation, glutathione metabolism and inflammatory responses are presented. The heat map represents gene expression changes, where up-regulation, no change and down-regulation are colored in red, white and blue, respectively.
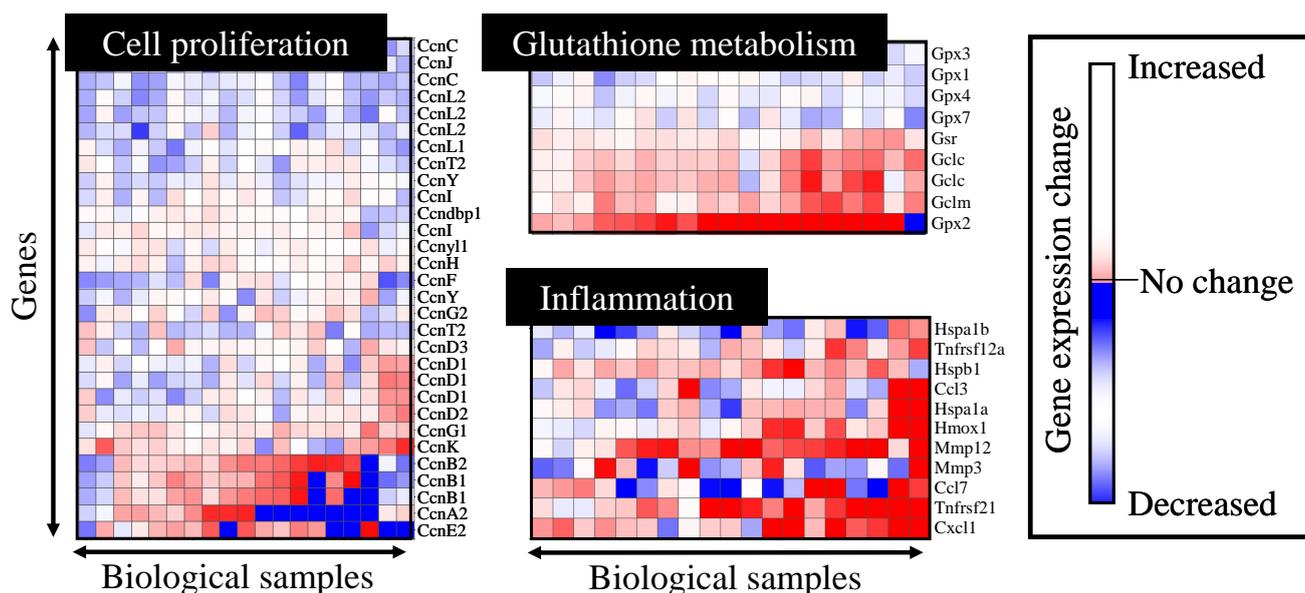
2.3. Open Source Software

Development of open source software is another driving factor of the recent advances in functional genomics research [25], although in many cases they require users to possess a certain degree of computational skills. In this paper, Bioconductor [26] (http://www.bioconductor.org/), which is implemented on the statistical software R [27] (http://cran.r-project.org/), GraphViz (http://www.graphviz.org/) and Cytoscape [28] (http://www.cytoscape.org/) were actually utilized and the results are presented.

3. Practical Application of TGx Database and Biomarkers

### 3.1. Scoring the Gene Set-Level Expression Changes

As stated before, the process of data analysis and interpretation of the results becomes increasingly complex when we handle large-scale microarray data sets and multiple TGx biomarker gene sets simultaneously. To facilitate our understanding of the toxicological significance based on the TGx data, a simple scoring strategy was introduced to evaluate gene set-level expression changes. Figure 2 shows the general concept on how microarray data and multiple TGx biomarker gene sets can be processed to

generate a simple score for each TGx biomarker gene set, by which toxicologists can identify which biological endpoints were affected by chemical exposure, and thereby can evaluate the toxicological significance efficiently.

In the Toxicogenomics Project in Japan, a large-scale TGx database called the Toxicogenomics Project-Genomics Assisted Toxicity Evaluation system (TG-GATEs) was developed that consisted of Affymetrix GeneChip data for approximately 150 prototypical toxicants on rat liver, kidney, hepatocytes and human hepatocytes [29]. In the TG-GATEs system, an expression ratio-based scoring method called 'TGP1 score' was utilized to facilitate understanding of the toxicological significance based on the TGx data set [30].

The TGP1 score is calculated by multiplying two elements: one element represents an index for the "overall direction of the expression change per probe set", and the other represents the "overall magnitude of the expression change per gene" of the *Biomarker X* gene set. The sign of the first index will be either positive or negative when the overall expression changes of the genes in *Biomarker X* were up- or down-regulated, respectively, and will be expected to approach zero when the direction of expression changes is divergent. The second index is always positive and will be higher when the expression change levels of the genes show a higher value. Collectively, the TGP1 score will be higher when the genes included in *Biomarker X* show uniform up-regulation with higher expression changing levels.

Figure 2. Scoring multiple toxicological endpoints using toxicogenomics data. Multiple toxicological endpoint-associated gene sets or TGx biomarkers need to be prepared in advance, and the overall expression changing levels for each gene set are calculated by certain algorithms such as the D-score, by which affected levels for each biological pathway can be evaluated intuitively.
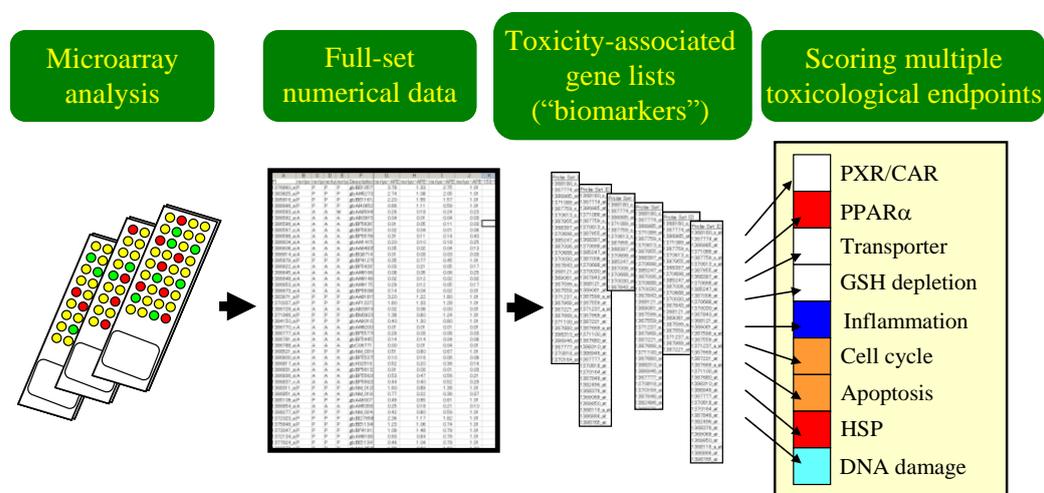

3.2. Differentially-Regulated Gene Score (D-score)

The TGP1 score was found to be useful for efficient comprehension of large-scale microarray data sets. However, the TGP1 score calculation treats both low and high quality gene expression data equally, and therefore the calculated score would be flawed if low quality data was involved in the calculation. To overcome this shortcoming, we introduced a new scoring method called *Differentially-*

expressed gene score (D-score) [31], where the data quality as well as the expression changing level for each gene is considered in the score calculation. Thus, the calculated score is much more reliable compared with the TGP1 score.

Figure 3 represents D-scores for microarray data on rat livers treated with one of the four hepatotoxicants: acetaminophen (APAP), phenobarbital (PB), clofibrate (CFB) or acetamidofluorene (AAF) using 10 biomarker gene sets for score calculation. Stimulation of glutathione depletion and inflammation by APAP, induction of Cyp2b and Cyp3a genes, induction of PPARα by CFB and induction of Cyp1a genes are evident by D-score analysis results (Figure 3(A)), where the dose-dependent stimulation of these endpoints or genes can also be evaluated (Figure 3(B)). These results demonstrate that TGx biomarker gene sets and the D-score calculation method dramatically facilitate the interpretation of the TGx data.

> **Figure 3.** Detection of affected toxicological endpoints by D-score. (**A**) Rats were treated with prototypical hepatotoxicants acetaminophen (APAP), phenobarbital (PB), clofibrate (CFB) or acetamidofluorene (AAF), and the hepatic microarray data were obtained at 3, 6, 9 and 24 h after treatment. The D-score highlights the activated toxicological endpoints elicited by the chemicals: glutathione depletion and inflammation by APAP, Cyp2b/Cyp3a induction by PB, PPARα activation by CFB and Cyp1 induction by AAF; (**B**) All the D-scores except for that of Cyp2b exhibited a clear dose-response. Data are reprinted from [31] with permission from Elsevier.
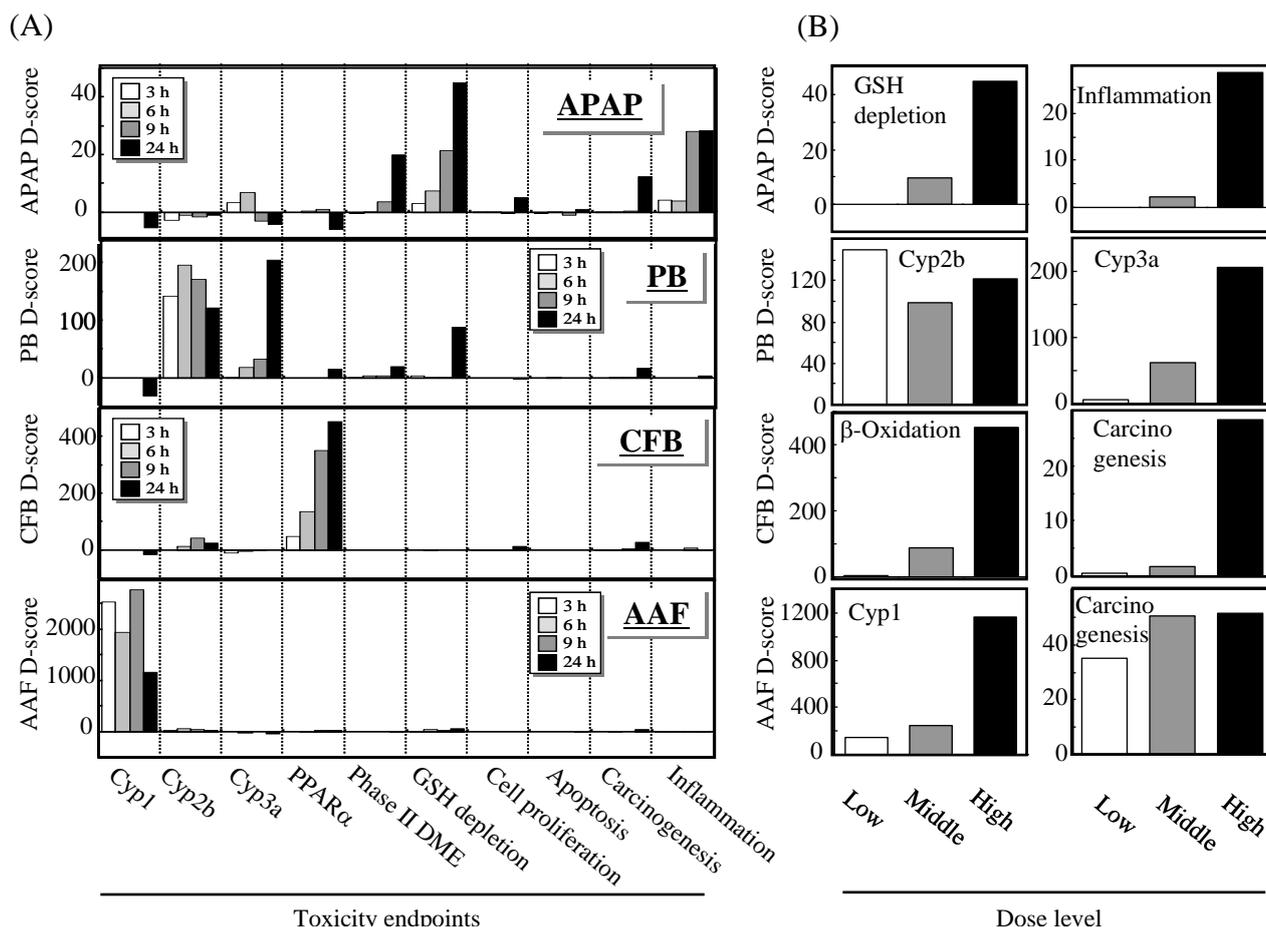
3.3. Inference of Gene Set-Level Network Structure Using a TGx Database

Previous studies reported inconsistency of interlaboratory/inter-platform microarray results [32,33], while others have reported good concordance among laboratories [34–36] or inconclusive results [37,38]. Because a gene set-level or biological pathway-level analysis has been reported to be more robust and comparable for microarray data sets obtained from different studies [39,40], we hypothesized that the biological pathway-level interactions could be better evaluated using D-scores for multiple gene sets as compared with generic molecular network analysis methods, which were conducted with individual gene-by-gene-level analysis [41]. We inferred the gene set-level network structure using a large-scale TGx database, TG-GATEs, and a total of 58 gene sets [42] with a Gaussian graphical model (GGM) algorithm to calculate partial correlation coefficients among the gene sets [43]. In addition to gene expression data, we also included changing levels of phenotypic data such as organ weight, blood chemistry and hematology parameters for calculation.

The inferred network presented in Figure 4 was found to contain a number of toxicologically-relevant gene set—gene set and gene set-phenotype relationships, such as blood glucose level and hepatic glycolysis-associated gene sets, or blood aminotransferase enzyme activity and inflammation-associated gene sets [42]. These results demonstrate that the retrospective network inference using a GGM algorithm successfully highlighted toxicologically significant gene set- and phenotype-level relationships from a large-scale TGx database. Furthermore, microarray data set obtained outside TG-GATEs was found to be well compatible with the network structure inferred based on TG-GATEs [42], suggesting that the gene set-level network structure was robust enough to be applicable for external microarray data sets.

4. Case Study: Bromobenzene-Induced Molecular Perturbation

In this section, a case study is presented for evaluating hepatic molecular perturbations elicited by 300 mg/kg bromobenzene (BBz) treatment in rats. D-score was calculated for a total of 58 gene sets and the calculated score was presented as either a radar chart (Figure 5), heat map (Figure 6) or network structure (Figure 7).

4.1. Radar Chart Presentation

Figure 5 shows the time course of absolute D-score values presented in radar charts, where a total of 58 gene sets, a detailed gene list which has been reported previously [42], were used for score calculation. Significant biological perturbations became evident at 6 h after BBz treatment, where *DNA damage-* and *Glutathione depletion*-associated gene sets exhibited high D-scores. BBz is reported to cause hepatic glutathione depletion through the generation of reactive metabolites [44], which is concordant with the high D-score for the glutathione depletion-associated genes. At 12 h, D-scores for *Oxidative stress-* and *Inflammation*-associated gene sets exhibited high scores, suggesting that oxidative stress, supposedly associated with glutathione depletion, induced liver injury, which was followed by an inflammatory response. *Gluthathione homeostasis*-associated genes, which include glutathione synthesis-related genes, were also activated at 12 h, which would contribute to feedback up-regulation of glutathione synthesis against acute glutathione depletion. The D-scores for

Inflammation- and *DNA damage*-associated gene sets became considerably high at 24 h after BBz treatment, suggesting the level of liver injury progressed in spite of activation of *Glutathione homeostasis*-associated genes.

Figure 4. Gene set- and phenotype-level network analysis. A large-scale TGx database, TG-GATE, was used for extracting statistically significant relationships among gene sets and phenotypes by utilizing a GGM algorithm. The network consists of D-scores for 58 gene sets, as well as changing levels of phenotype data such as organ weight, blood chemistry and hematology. Purple and green represent positive and negative partial correlation coefficients, respectively, and the width of the lines represents strength of the correlation measured with partial correlation coefficient. The network was drawn with open-source software Cytoscape.
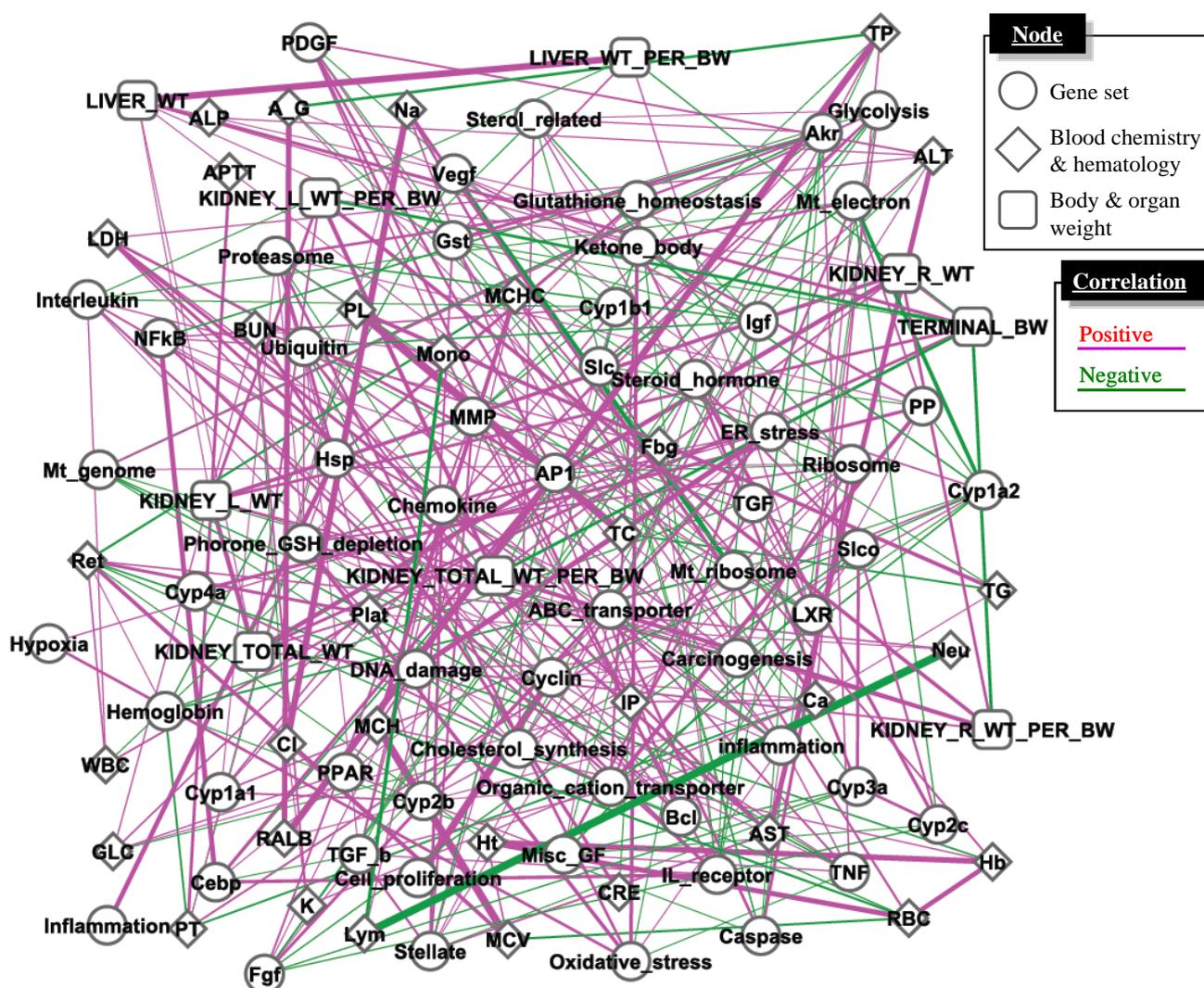
Figure 5. Time course of D-score: Radar chart presentation. D-scores were calculated for rat livers harvested at 2, 6, 12 and 24 h after bromobenzene treatment, and are presented in a radar chart. The red line indicates the D-score for each gene set, and the blue circle indicates a D-score = 20 for each gene set. Detailed information for the gene sets used can be obtained in a previous report [42].
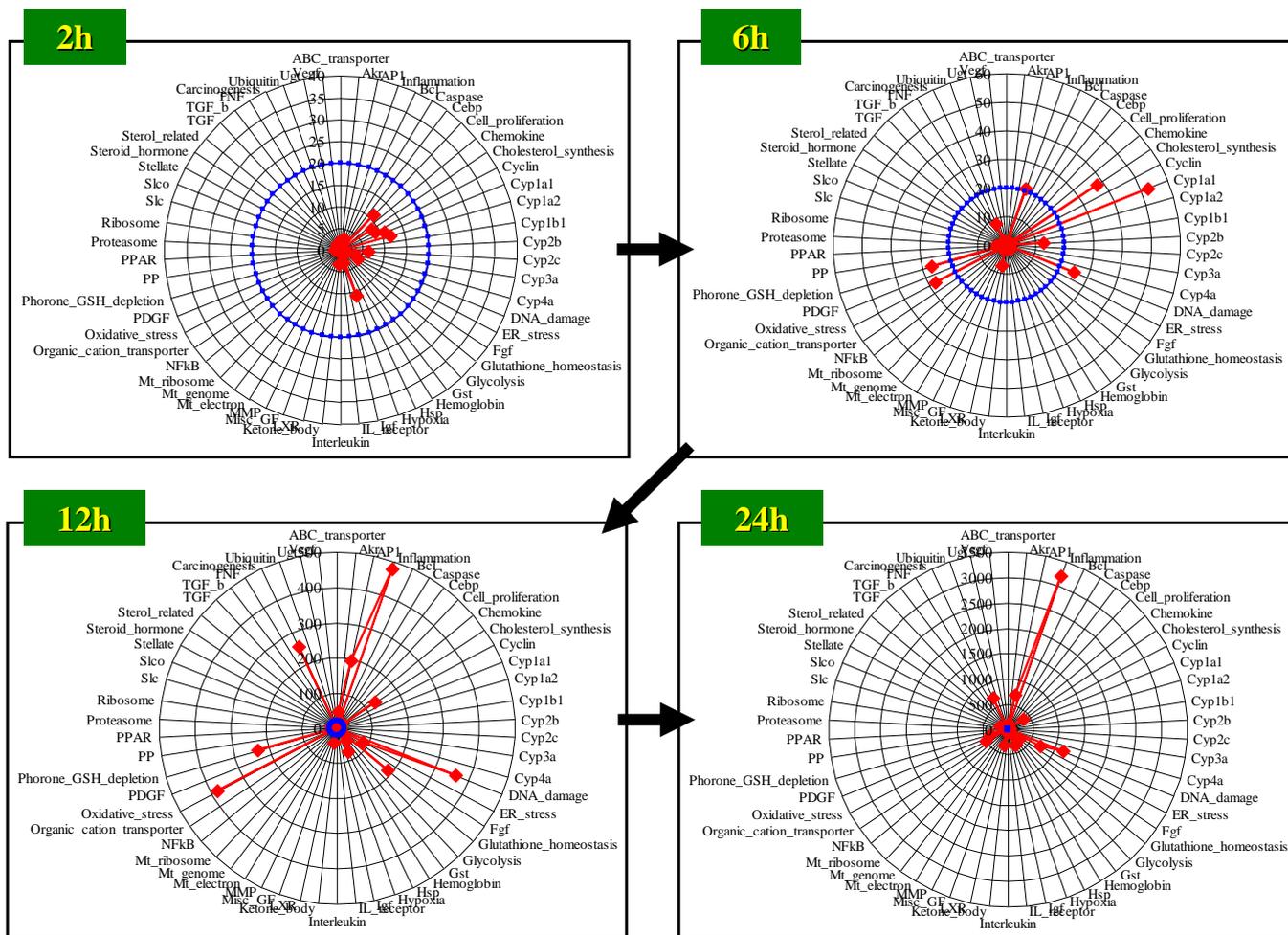
4.2. Heat Map Presentation

Time course transition of D-score profiles is presented in Figure 6, which clearly demonstrates the molecular mechanisms of BBz-elicited toxicity: Glutathione depletion and oxidative stress are the first triggers after BBz treatment, which caused cell death suggested by up-regulation of DNA damage-associated genes. This was followed by inflammation, tissue repair, as well as activation of antioxidant factors such as up-regulation of glutathione synthesis-related genes, glutathione *S*-transferase (Gst) genes or aldo-keto reductase (Akr) genes.

Figure 6. Time course of D-score: Heat map presentation. Red and blue indicate high and low D-scores, or up- and down-regulation for each gene set, respectively. The heat map indicates that the first trigger invoked by bromobenzene exposure was glutathione depletion and associated oxidative stress responses, followed by cell death, inflammation and up-regulation of antioxidant factors, as well as down-regulation of energy metabolism and drug metabolizing enzymes.
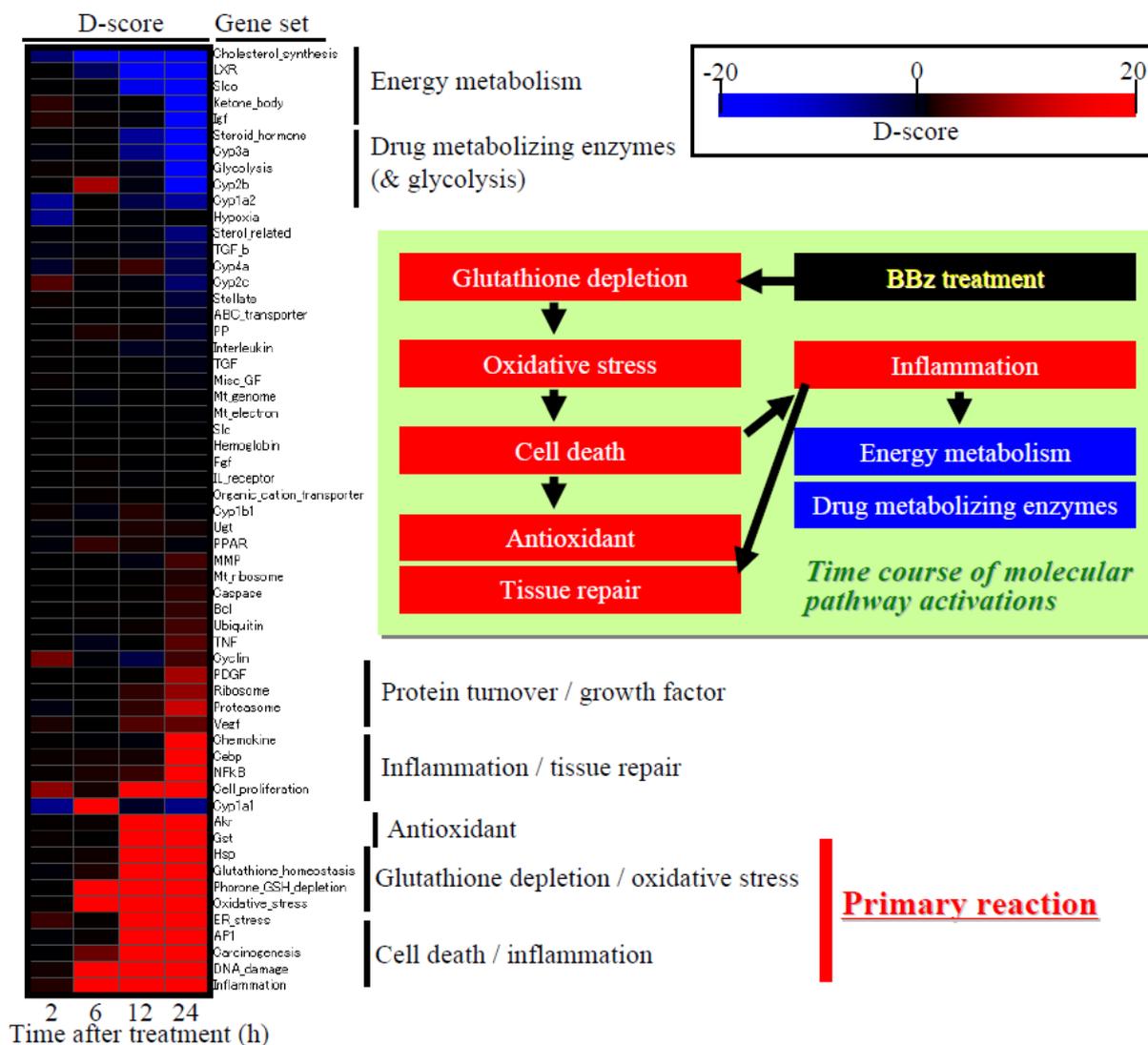
4.3. Supervised Network Structure Presentation

The BBz-elicited biological responses can also be visualized with the pre-defined biological network structure. Figure 7 represents a D-score profile at 24 h after BBz treatment, where gene sets associated with glutathione depletion, oxidative stress and inflammation were up-regulated (colored in red). On the other hand, gene sets associated with energy metabolism (*i.e.*, cholesterol synthesis and glycolysis) were down-regulated (colored in blue). Thus, the systems-level biological responses can be intuitively characterized with the supervised gene set-level biological network. However, the network structure should be kept up-to-date in a timely manner when novel toxicological knowledge has been obtained.

Figure 7. Network structure presentation of D-scores. Biological and toxicological relationships among gene sets were visualized as a supervised network using GraphViz software. The D-scores calculated for microarray data on rat livers at 24 h after bromobenzene treatment are presented in a network structure using GraphViz software, where red and blue indicate high and low D-scores, respectively. The network demonstrates oxidative stress and DNA damage were induced by bromobenzene treatment, while sterol metabolism was down-regulated.
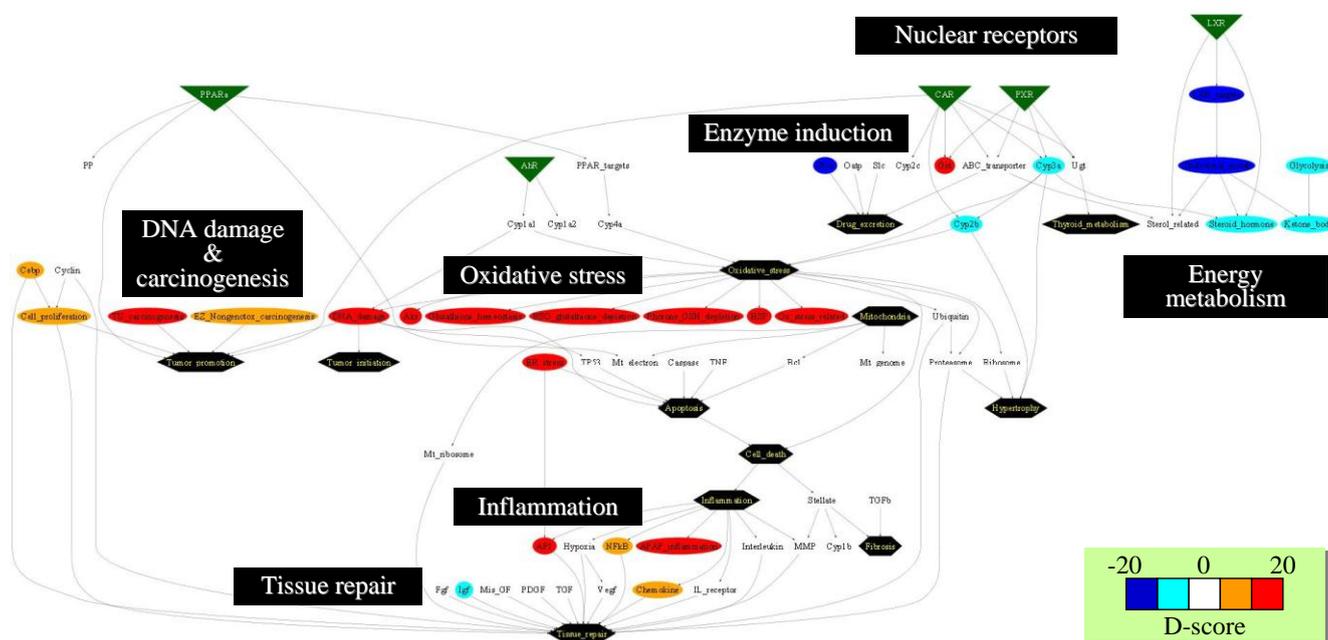

5. Conclusion

Systems-level understanding of molecular perturbations is crucial for evaluating chemical-induced toxicity risks. Microarray data provides comprehensive gene expression responses against chemical exposure, and therefore the TGx approach is highly advantageous for understanding systems-level biological perturbations. To solve the difficulty in handling huge amounts of TGx data sets, preparation of TGx biomarker gene sets and implementation of gene set-level data analysis are effective. The general flow of TGx data analysis is shown in Figure 8. The first step is to identify biological pathways that were affected by the chemical exposure, for which a radar chart presentation will be useful (Figure 8(A)). Detailed expression analysis of the individual genes will be needed to focus on the affected biological pathways (Figure 8(B)). When available, large-scale TGx reference database will be utilized for comparative analysis to appropriately evaluate the toxicological significance (Figures 8(C) and 8(D)). To comprehend the systems-level molecular dynamics, relationships among pathways should be taken into consideration (Figure 8(F)). Such analytical flow can be automated if appropriate computational skills are available. Refining toxicogenomic biomarker gene sets, scoring algorithm and development of user-friendly integrative software will substantially help the utilization of the TGx data set to evaluate biological response by which hazardous effects of exposed chemicals could be appropriately managed.

Figure 8. Analytical flow for toxicity evaluation using TGx data. (**A**) Radar chart for D-scores using 58 gene sets; (**B**) Box plot for comparative analysis using a reference TGx database; (**C**) Heat map for individual genes; (**D**) TGx reference database and TGx biomarker knowledgebase; (**E**) Unsupervised gene set-level network inference to extract toxicological relationships among pathways; (**F**) Supervised gene set-level network.
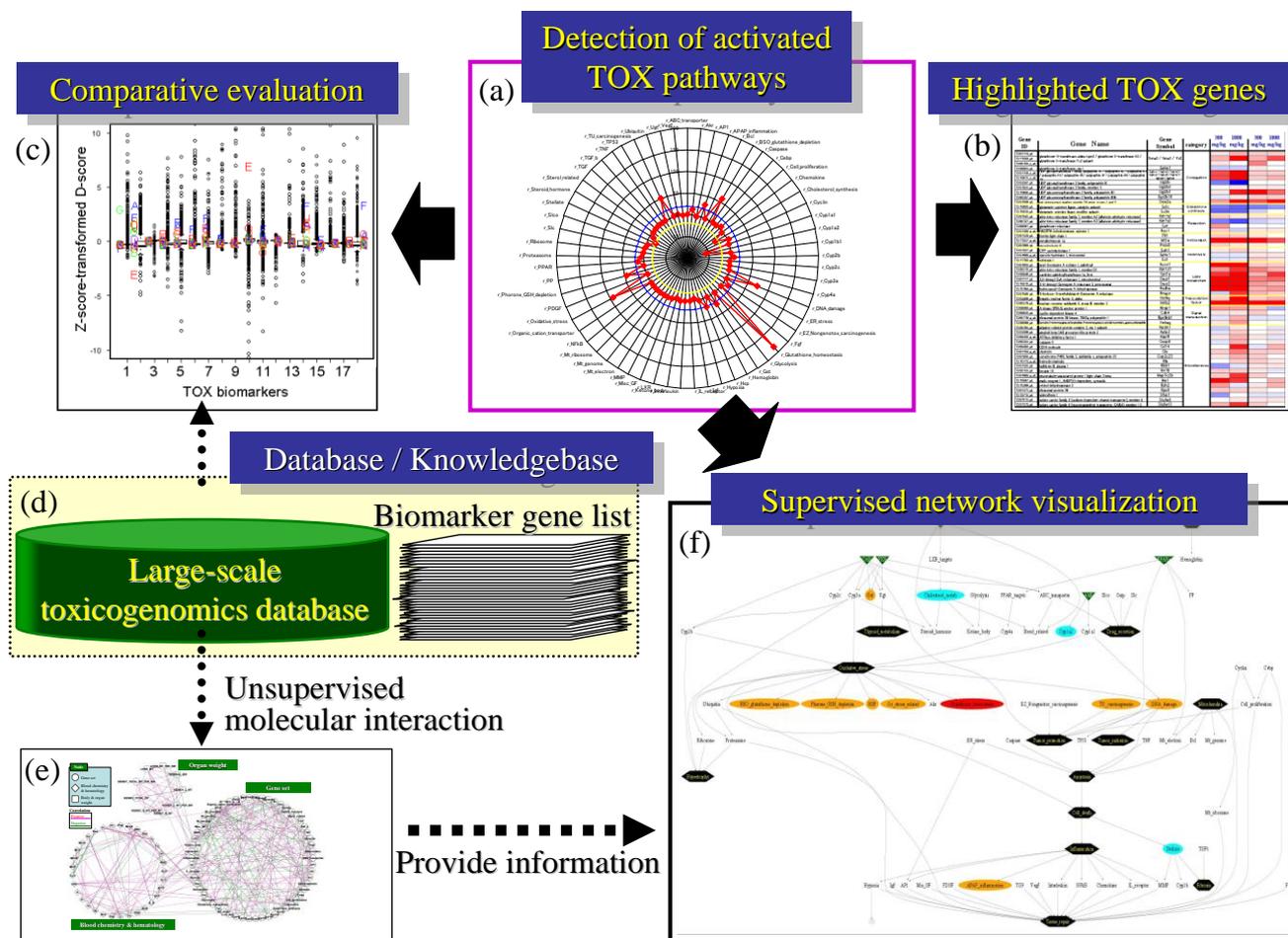
References

1. Kaplowitz, N. Idiosyncratic drug hepatotoxicity. *Nat. Rev. Drug Discov.* **2005**, *4*, 489–499.
2. Marchetti, S.; Mazzanti, R.; Beijnen, J.H.; Schellens, J.H. Concise review: Clinical relevance of drug drug and herb drug interactions mediated by the ABC transporter ABCB1 (MDR1, P-glycoprotein). *Oncologist* **2007**, *12*, 927–941.
3. Haddad, A.; Davis, M.; Lagman, R. The pharmacological importance of cytochrome CYP3A4 in the palliation of symptoms: Review and recommendations for avoiding adverse drug interactions. *Support. Care Cancer* **2007**, *15*, 251–257.

4. Kodama, S.; Negishi, M. Phenobarbital confers its diverse effects by activating the orphan nuclear receptor car. *Drug Metab. Rev.* **2006**, *38*, 75–87.

5. Kitano, H. Towards a theory of biological robustness. *Mol. Syst. Biol.* **2007**, *3*, 137.

6. Boverhof, D.R.; Zacharewski, T.R. Toxicogenomics in risk assessment: applications and needs. *Toxicol. Sci.* **2006**, *89*, 352–360.

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

8. Waring, J.F.; Ciurlionis, R.; Jolly, R.A.; Heindel, M.; Ulrich, R.G. Microarray analysis of hepatotoxins *in vitro* reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol. Lett.* **2001**, *120*, 359–368.

9. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; Harris, M.A.; Hill, D.P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J.C.; Richardson, J.E.; Ringwald, M.; Rubin, G.M.; Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29.

10. Goto, S.; Bono, H.; Ogata, H.; Fujibuchi, W.; Nishioka, T.; Sato, K.; Kanehisa, M. Organizing and computing metabolic pathway data in terms of binary relations. *Pac. Symp. Biocomput.* **1997**, 175–186.

11. Dahlquist, K.D.; Salomonis, N.; Vranizan, K.; Lawlor, S.C.; Conklin, B.R. GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. *Nat. Genet.* **2002**, *31*, 19–20.

12. Kiyosawa, N.; Ando, Y.; Manabe, S.; Yamoto, T. Toxicogenomic biomarkers for liver toxicity. *J. Toxicol. Pathol.* **2009**, *22*, 35–52.

13. Kier, L.D.; Neft, R.; Tang, L.; Suizu, R.; Cook, T.; Onsurez, K.; Tiegler, K.; Sakai, Y.; Ortiz, M.; Nolan, T.; Sankar, U.; Li, A.P. Applications of microarrays with toxicologically relevant genes (tox genes) for the evaluation of chemical toxicants in Sprague Dawley rats *in vivo* and human hepatocytes *in vitro*. *Mutat. Res.* **2004**, *549*, 101–113.

14. Ellinger-Ziegelbauer, H.; Gmuender, H.; Bandenburg, A.; Ahr, H.J. Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term *in vivo* studies. *Mutat. Res.* **2008**, *637*, 23–39.

15. Uehara, T.; Hirode, M.; Ono, A.; Kiyosawa, N.; Omura, K.; Shimizu, T.; Mizukawa, Y.; Miyagishima, T.; Nagao, T.; Urushidani, T. A toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats. *Toxicology* **2008**, *250*, 15–26.

16. Sawada, H.; Takami, K.; Asahi, S.A. Toxicogenomic approach to drug-induced phospholipidosis: Analysis of its induction mechanism and establishment of a novel *in vitro* screening system. *Toxicol. Sci.* **2005**, *83*, 282–292.

17. Hirode, M.; Ono, A.; Miyagishima, T.; Nagao, T.; Ohno, Y.; Urushidani, T. Gene expression profiling in rat liver treated with compounds inducing phospholipidosis. *Toxicol. Appl. Pharmacol.* **2008**, *229*, 290–299.

18. Kiyosawa, N.; Ito, K.; Sakuma, K.; Niino, N.; Kanbori, M.; Yamoto, T.; Manabe, S.; Matsunuma, N. Evaluation of glutathione deficiency in rat livers by microarray analysis. *Biochem. Pharmacol.* **2004**, *68*, 1465–1475.

19. Kiyosawa, N.; Uehara, T.; Gao, W.; Omura, K.; Hirode, M.; Shimizu, T.; Mizukawa, Y.; Ono, A.; Miyagishima, T.; Nagao, T.; Urushidani, T. Identification of glutathione depletion-responsive genes using phorone-treated rat liver. *J. Toxicol. Sci.* **2007**, *32*, 469–486.

20. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic. Acids Res.* **2002**, *30*, 207–210.

21. Brazma, A.; Parkinson, H.; Sarkans, U.; Shojatalab, M.; Vilo, J.; Abeygunawardena, N.; Holloway, E.; Kapushesky, M.; Kemmeren, P.; Lara, G.G.; Oezcimen, A.; Rocca-Serra, P.; Sansone, S.A. Array Express—A public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.* **2003**, *31*, 68–71.

22. Waters, M.; Stasiewicz, S.; Merrick, B.A.; Tomer, K.; Bushel, P.; Paules, R.; Stegman, N.; Nehls, G.; Yost, K.J.; Johnson, C.H.; Gustafson, S.F.; Xirasagar, S.; Xiao, N.; Huang, C.C.; Boyer, P.; Chan, D.D.; Pan, Q.; Gong, H.; Taylor, J.; Choi, D.; Rashid, A.; Ahmed, A.; Howle, R.; Selkirk, J.; Tennant, R.; Fostel, J. CEBS—Chemical Effects in Biological Systems: A public data repository integrating study design and toxicity data with microarray and proteomics data. *Nucleic Acids Res.* **2008**, *36*, D892–D900.

23. Mattingly, C.J.; Rosenstein, M.C.; Davis, A.P.; Colby, G.T.; Forrest, J.N., Jr.; Boyer, J.L. The comparative toxicogenomics database: A cross-species resource for building chemical-gene interaction networks. *Toxicol. Sci.* **2006**, *92*, 587–595.

24. Hayes, K.R.; Vollrath, A.L.; Zastrow, G.M.; McMillan, B.J.; Craven, M.; Jovanovich, S.; Rank, D.R.; Penn, S.; Walisser, J.A.; Reddy, J.K.; Thomas, R.S.; Bradfield, C.A. EDGE: A centralized resource for the comparison, analysis, and distribution of toxicogenomic information. *Mol. Pharmacol.* **2005**, *67*, 1360–1368.

25. Gehlenborg, N.; O'Donoghue, S.I.; Baliga, N.S.; Goesmann, A.; Hibbs, M.A.; Kitano, H.; Kohlbacher, O.; Neuweger, H.; Schneider, R.; Tenenbaum, D.; Gavin, A.C. Visualization of omics data for systems biology. *Nat. Methods* **2010**, *7*, S56–S68.

26. Gentleman, R.C.; Carey, V.J.; Bates, D.M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A.J.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J.Y.; Zhang, J. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **2004**, *5*, R80.

27. *R: A Language and Environment for Statistical Computing*; R Development Core Team: Vienna, Australia, 2008.

28. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504.

29. Urushidani, T. Prediction of hepatotoxicity based on the toxicogenomics database. In *Hepatotoxicity from Genomics to in Vitro and in Vivo Models*; Sahu, S., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2008; pp. 507–529.

30. Kiyosawa, N.; Shiwaku, K.; Hirode, M.; Omura, K.; Uehara, T.; Shimizu, T.; Mizukawa, Y.; Miyagishima, T.; Ono, A.; Nagao, T.; Urushidani, T. Utilization of a one-dimensional score for surveying chemical-induced changes in expression levels of multiple biomarker gene sets using a large-scale toxicogenomics database. *J. Toxicol. Sci.* **2006**, *31*, 433–448.

31. Kiyosawa, N.; Ando, Y.; Watanabe, K.; Niino, N.; Manabe, S.; Yamoto, T. Scoring multiple toxicological endpoints using a toxicogenomic database. *Toxicol. Lett.* **2009**, *188*, 91–97.

32. Mah, N.; Thelin, A.; Lu, T.; Nikolaus, S.; Kuhbacher, T.; Gurbuz, Y.; Eickhoff, H.; Kloppel, G.; Lehrach, H.; Mellgard, B.; Costello, C.M.; Schreiber, S. A comparison of oligonucleotide and cDNA-based microarray systems. *Physiol. Genomics* **2004**, *16*, 361–370.

33. Severgnini, M.; Bicciato, S.; Mangano, E.; Scarlatti, F.; Mezzelani, A.; Mattioli, M.; Ghidoni, R.; Peano, C.; Bonnal, R.; Viti, F.; Milanesi, L.; De Bellis, G.; Battaglia, C. Strategies for comparing gene expression profiles from different microarray platforms: application to a case-control experiment. *Anal. Biochem.* **2006**, *353*, 43–56.

34. Waring, J.F.; Ulrich, R.G.; Flint, N.; Morfitt, D.; Kalkuhl, A.; Staedtler, F.; Lawton, M.; Beekman, J.M.; Suter, L. Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. *Environ. Health Perspect.* **2004**, *112*, 439–448.

35. Chu, T.M.; Deng, S.; Wolfinger, R.; Paules, R.S.; Hamadeh, H.K. Cross-site comparison of gene expression data reveals high similarity. *Environ. Health Perspect.* **2004**, *112*, 449–455.

36. Fielden, M.R.; Nie, A.; McMillian, M.; Elangbam, C.S.; Trela, B.A.; Yang, Y.; Dunn, R.T., II.; Dragan, Y.; Fransson-Stehen, R.; Bogdanffy, M.; Adams, S.P.; Foster, W.R.; Chen, S.J.; Rossi, P.; Kasper, P.; Jacobson-Kram, D.; Tatsuoka, K.S.; Wier, P.J.; Gollub, J.; Halbert, D.N.; Roter, A.; Young, J.K.; Sina, J.F.; Marlowe, J.; Martus, H.J.; Aubrecht, J.; Olaharski, A.J.; Roome, N.; Nioi, P.; Pardo, I.; Snyder, R.; Perry, R.; Lord, P.; Mattes, W.; Car, B.D. Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicol. Sci.* **2008**, *103*, 28–34.

37. Beyer, R.P.; Fry, R.C.; Lasarev, M.R.; McConnachie, L.A.; Meira, L.B.; Palmer, V.S.; Powell, C.L.; Ross, P.K.; Bammler, T.K.; Bradford, B.U.; Cranson, A.B.; Cunningham, M.L.; Fannin, R.D.; Higgins, G.M.; Hurban, P.; Kayton, R.J.; Kerr, K.F.; Kosyk, O.; Lobenhofer, E.K.; Sieber, S.O.; Vliet, P.A.; Weis, B.K.; Wolfinger, R.; Woods, C.G.; Freedman, J.H.; Linney, E.; Kaufmann, W.K.; Kavanagh, T.J.; Paules, R.S.; Rusyn, I.; Samson, L.D.; Spencer, P.S.; Suk, W.; Tennant, R.J.; Zarbl, H. Multicenter study of acetaminophen hepatotoxicity reveals the importance of biological endpoints in genomic analyses. *Toxicol. Sci.* **2007**, *99*, 326–337.

38. Baker, V.A.; Harries, H.M.; Waring, J.F.; Duggan, C.M.; Ni, H.A.; Jolly, R.A.; Yoon, L.W.; de Souza, A.T.; Schmid, J.E.; Brown, R.H.; Ulrich, R.G.; Rockett, J.C. Clofibrate-induced gene expression changes in rat liver: A cross-laboratory analysis using membrane cDNA arrays. *Environ. Health Perspect.* **2004**, *112*, 428–438.

39. Ulrich, R.G.; Rockett, J.C.; Gibson, G.G.; Pettit, S.D. Overview of an interlaboratory collaboration on evaluating the effects of model hepatotoxicants on hepatic gene expression. *Environ. Health Perspect.* **2004**, *112*, 423–427.

40. Manoli, T.; Gretz, N.; Grone, H.J.; Kenzelmann, M.; Eils, R.; Brors, B. Group testing for pathway analysis improves comparability of different microarray datasets. *Bioinformatics* **2006**, *22*, 2500–2506.

41. Ma'ayan, A. Insights into the organization of biochemical regulatory networks using graph theory analyses. *J. Biol. Chem.* **2009**, *284*, 5451–5455.

42. Kiyosawa, N.; Manabe, S.; Sanbuissho, A.; Yamoto, T. Gene set-level network analysis using a toxicogenomics database. *Genomics* **2010**, *96*, 39–49.

43. Crzegorczyk, M.; Husmeier, D.; Werhli, A.V. Reverse engineering gene regulatory networks with various machine learning methods. In *Analysis of Microarray Data—A Network Approach*; Emmert-Streib, F., Dehmer, M., Eds.; Wiley-VCH: Hoboken, NJ, USA, 2008.

44. Lau, S.S.; Monks, T.J. The contribution of bromobenzene to our current understanding of chemically-induced toxicities. *Life Sci.* **1988**, *42*, 1259–1269.