



Published in final edited form as:

Expert Opin Drug Metab Toxicol. 2008 November ; 4(11): 1379–1389. doi:10.1517/17425255.4.11.1379.

Systems biology and functional genomics approaches for the identification of cellular responses to drug toxicity

Alison Hege Harrill, B.S. [Pre-doctoral fellow]¹ and Ivan Rusyn, M.D., Ph.D. [Associate Professor]^{1,2,*}

¹University of North Carolina at Chapel Hill, Curriculum in Toxicology, 0031 Michael Hooker Research Center, CB 7431, Chapel Hill, NC, 27713, USA., Phone: +1 919 843 2596, Fax: +1 919 966 7911

²University of North Carolina at Chapel Hill, Department of Environmental Sciences and Engineering, 0031 Michael Hooker Research Center, CB 7431, Chapel Hill, NC, 27713, USA. Phone: +1 919 843 2596, Fax: +1 919 966 7911

Abstract

Extensive growth in the field of molecular biology in recent decades has led to the development of new and powerful experimental and computational tools that enable the analysis of complex biological responses to chemical exposure on both a functional and structural genetic level. The ability to profile global responses on a transcriptional level has become a valuable resource in the science of toxicology and attempts are currently being made to further understand toxicity mechanisms by incorporating metabolomics and proteomics approaches. In addition, recent progress in understanding the extent of the genetic diversity within and between species allows us to take a fresh look at research on genetic polymorphisms which may influence an individual's susceptibility to toxicity. While new technologies have the potential to make a sizeable impact on our understanding of the mechanisms of toxicity, considerable challenges remain to be addressed, especially with regards to the regulatory acceptance and successful integration of omics data. This review highlights recent advancements in the application of functional and structural genomics techniques to chemical hazard identification and characterization, and to the understanding of the inter-individual differences in susceptibility to adverse drug reactions.

Keywords

toxicogenomics; toxicogenetics; susceptibility; systems biology; high-throughput toxicity screening

1. Introduction

Advances in genomics in recent decades have led to exponential growth in the understanding of genome structure and function in both humans and other organisms. Due to the wealth of resources available, including databases of sequence variation and transcriptional changes following chemical exposure, both the industry and the federal government have begun to invest more resources into using genomic tools to improve drug development and drug and chemical safety evaluation. Systems toxicology, or the integration of traditional toxicology approaches with the development and implementation of toxicogenomics, proteomics, and

*To whom correspondence should be addressed..

Conflict of Interest

The authors declare that they have no competing interests.

metabolomics, offers the promise of the development of novel reliable toxicity biomarkers and more accurate predictions of adverse health effects in humans.

The need for development of sensitive biomarkers to predict drug toxicity was recently reinforced by the Food and Drug Administration (FDA) Critical Path Initiative which emphasizes the development of diagnostic tests to improve clinical trials and post-market surveillance¹. In 2005, the FDA released a final guidance document on genomic data submissions, *Guidance for Industry: Pharmacogenomic Data Submissions*², and announced the creation of a new FDA program entitled the Interdisciplinary Pharmacogenomic Review Group³. The importance of genomic information in FDA regulatory decisions is reflected in the increasing voluntary inclusion of these types of data in FDA submissions³. To establish reliable protocols for data analysis and quality control in genomics studies, several consortia have been established which included the academia, industry, FDA, Environmental Protection Agency, the National Institute of Standards and Technology, and the National Institutes of Health⁴⁻⁶. In the near future, a strong understanding of genomic data and its potential uses will be critical to the drug development process and clinical practice. The literature reviewed here focuses on the current progress in the application of genomic and genetic data to drug safety and highlights recent successes and current knowledge gaps in the following areas: 1) the discovery of toxicity biomarkers and identification of compounds that are predicted to cause toxicity, 2) the identification of susceptible individuals, and 3) integration of various omics datasets to improve understanding of toxicity mechanisms within a biological system.

2. Functional Genomics Strategies

2.1 Functional genomics allows for global gene transcript profiling

The major challenge of toxicology today is the ability to extrapolate risk from experimental systems to human populations. Rodents are most frequently used for *in vivo* toxicity testing, yet there often are major differences in clearance, metabolic activity of enzymes and other key factors between species. While it is difficult to directly translate rodent findings to human population, human testing is most often not an option and the answer to this problem is to develop biomarkers of toxicity that are sub-clinical and that can be assayed from a non-invasive tissue. The promise of toxicogenomics is to identify gene expression patterns or signatures that will indicate an adverse health event at low doses. However, to do this, it needs to be determined which changes and pathways are best associated with impending pathology vs. benign adaptive changes that are responsive to the chemical but are not associated with toxicity.

At the forefront of these emerging technologies is the use of functional genomics or analysis of molecular perturbations as measured by transcriptomics assayed by microarray technology. Gene expression analysis using current microarray platforms that encompass whole genomes allows for a comprehensive picture of cellular responses⁷, especially when examined in concert with proteomic and metabolomic data. While early microarray experiments were often criticized as “fishing expeditions”, current testing strategies enable the generation of testable hypothesis that aid in understanding the mechanism of toxicity⁸. The ultimate value of using transcript profiling data within the context of systems toxicology is that responses that are predictive of cellular injury can be sampled across doses, time points, and species, facilitating risk extrapolation to human populations.

2.2 Improving safety assessment through genomic profiling

Microarray profiling has become a major omics tool for the characterization of drug toxicities by allowing for large-scale determination of gene expression changes associated with a defined pathology. A key step in toxicogenomics is the ability to link a chemical-elicited phenotype with gene expression changes, termed “phenotypic anchoring”⁹. Experimentally determined

gene expression signatures can serve as a guide to determining biomarkers that are indicative of toxicological responses that may be as-yet sub-clinical, with no observable morphological changes¹⁰. To confirm the utility of the approach, Heinloth et al.¹¹ demonstrated that patterns of gene expression perturbations observed at sub-toxic doses of acetaminophen in rats may indicate subtle cellular injury that was not detectable by histopathology or clinical chemistry methods within the liver. At toxic doses, expression changes in the same subset of genes associated with mitochondrial dysfunction and oxidative stress were more exaggerated and changes were detected in additional genes associated with these processes. These data indicate that gene expression profiling has the potential to identify subtle markers of cellular injury that precipitate overt organ toxicity.

2.3 Identification of biomarkers of toxicity

Identification of sensitive biomarkers that will assist in monitoring drug therapy for evidence of toxicity or therapeutic outcome, and (in acute poisoning cases) to predict exposure levels, is a critical gap where omics technologies hold promise. Genomic biomarkers of toxicity have recently been identified for a wide variety of toxicants including nephrotoxic agents¹², testicular toxicants¹³, and for keratinocyte proliferation in papilloma murine skin model¹⁴, to name only a few. The potential of using this technology to identify safety biomarkers is great and may help to create better diagnostic tools for traditionally difficult toxicodynamic monitoring, such as in patients receiving immunosuppressive therapy¹⁵.

The majority of recent investigations have used microarrays to study toxicity in target organ tissue or in cultured cells. While these experiments often yield important insight into the mechanism of toxicity, they provide limited information for monitoring drug safety in patient populations through non-invasive means. To address this limitation, Bushel et al.¹⁶ investigated the utility of using gene expression signatures in peripheral blood as an early indicator of pathological changes in the liver following administration of varying doses of acetaminophen. In this study, a prediction algorithm using classifiers and a pattern-based method that was weighted toward non-injurious exposure levels was used to discriminate sub-toxic and toxic exposure doses. Characterization of acetaminophen-induced liver injury using gene expression profiles derived from blood was shown to better predict acute chemical exposure levels than clinical chemistry, hematology, or histopathology analysis, indicating that transcript profiles derived from blood may be a good marker for specific organ toxicity.

2.4 High-throughput chemical toxicity testing is facilitated by microarrays

Microarrays have also enabled high-throughput screening of chemicals for potential toxicity. Increased sensitivity to detect molecular signatures or “fingerprints” that are associated with a particular class of chemicals or toxic response in both *in vitro*^{17, 18} and *in vivo*^{19, 20} systems, coupled with data-rich output, offer promise in screening for potential adverse health effects. In addition, approaches have been developed that do not rely upon *a priori* knowledge of expression signatures using well-characterized reference compounds, instead providing an *ab initio* estimation of compound-induced liver injury based upon microarray transcript profiling²¹. Furthermore, gene expression profiling has been suggested to have better predictive power than standard toxicological endpoints in a two-year rodent bioassay to detect carcinogenic potential of compounds²².

To validate the technology as a tool that may be useful in the classification of compound toxicity, an *in vivo* approach has often been utilized, in which defined organ pathology was present. It is assumed that toxicants that elicit similar pathology or disease will elicit a common pattern of gene expression changes. Recent studies aimed at prediction of toxicity based upon gene expression signatures have shown some success at distinguishing between non-genotoxic and genotoxic carcinogenic hepatotoxicants²³, and renal and hepatic toxicants²⁴, and have

enabled classification of hepatotoxicants that are macrophage activators, peroxisome proliferators, and oxidative stressors²⁵.

Because an *in vivo* approach would be impractical for use in high-throughput toxicity screening, *in vitro* culture systems offer an attractive alternative. For example, it was demonstrated that gene expression signatures can distinguish between toxic and non-toxic chemicals in HepG2 cells²⁶. However, until recently there was little data that suggested that toxicity signatures determined in an *in vitro* system would predict *in vivo* responses. The Toxicogenomics Project (TGP), a consortia composed of the Japanese National Institute of Health Sciences, the National Institute of Biomedical Innovation, and 15 pharmaceutical companies, was developed in 2002 to address the problem of connecting model systems with prediction of toxicity in humans²⁷. From the data collected by the TGP on gene expression profiles collected from rat liver and rat and human hepatocytes treated with 150 chemicals, it became clear that a major challenge to overcome was the identification of species-specific and common responses. The possibility of an informational bridge connecting transcript responses between rat and human hepatocytes and rat liver *in vivo* was investigated after administration of coumarin²⁷. In this experiment, gene expression profiles were compared first between *in vivo* rat liver samples and coumarin-treated rat hepatocytes to determine those genes that were induced by treatment and that were significant in both models and that served as an *in vivo-in vitro* bridge. Orthologous bridging genes were then measured in coumarin-treated human hepatocytes and a signature of 14 up-regulated and 11 down-regulated mRNA transcripts was identified. While more data is needed to connect species and model systems with risk assessment to humans, the approach represents an important step in enabling high-throughput chemical toxicity screening.

The ability of traditional cell cultures of primary or cultured cells to predict toxicity can also suffer from an inability to accurately model multi-cellular processes and the myriad effects of cell-cell communication. A precision-cut liver slice model was suggested as a promising model system²⁸. Liver slices offer the advantage of containing all relevant cell types within their original spatial arrangements in which cellular communication is maintained²⁹. The presence of all relevant cell types is particularly important because liver non-parenchymal have been linked to the mediation of liver toxicity for diverse chemicals such as alcohol^{30, 31}, cyclohexamide³², thioacetamide³³, and acetaminophen³⁴. Prototype experiments using rat liver slices demonstrated good prediction for toxicity and pathology based on microarray data for liver slices treated with acetaminophen, lipopolysaccharide, and carbon tetrachloride and these data correlated well with *in vivo* profiles²⁸; however experiments using human liver slices are needed to enhance risk estimates.

2.5 Pathway analysis aids in determination of the mode of action

A major knowledge gap in connecting gene expression profiles to classes of toxicants that elicit similar phenotypes is a lack of complete information on the complexity of cellular molecular pathways. Regulation of responses can be relatively straightforward (i.e., at the transcriptional level), but it can also be quite complex and controlled by multiple genes, proteins, and metabolites. Therefore, accurate identification of gene-gene interactions and regulation is essential for determining well-defined pathways that could serve as potential targets in therapeutic development or intervention. The importance of understanding subtle pathway changes and the need for applying such data to risk assessment have been underlined in the recent National Research Council's report entitled "Toxicity Testing in the 21st Century"³⁵. Transcriptional pathway-level analysis may prove essential for determining molecular perturbations that are associated with sub-clinical pathologies because alterations in gene expression are likely to be a more sensitive endpoint than the clinically observed toxicity.

There are several public databases available which aid in the annotation and interpretation of gene expression data in terms of cellular process, functions, and pathways. These databases

include the Gene Ontology (GO; www.geneontology.org), Gene Map Annotator and Pathway Profiler (www.genmapp.org), the Science Signaling Connections Map (stke.sciencemag.org/cm/), BioCarta (www.biocarta.com/genes/index.asp), Reactome (www.genomeknowledge.org), and KEGG (www.genome.jp/kegg/pathway.html). Most often, these tools are used to associate a set of differentially expressed genes from a microarray experiment with a particular pathway with the goal of identifying key modes of action. Several groups have developed statistical methods for associating cellular pathways and annotations with gene expression changes such as GOMiner³⁶, Significance Analysis of Functional Categories³⁷, and Onto-Tools³⁸. While the methods employed by these statistical techniques vary, each is aimed at assigning biological meaning to gene expression data. A broad understanding of the pathways affected by a variety of xenobiotics is essential for the future of drug safety assessment. A complete understanding of the molecular perturbations that precipitate a toxic pathology *in vivo* will enable a shift from high dose to low dose testing and may enable a better prediction of pathology from *in vitro* data when the same molecular pathways are affected.

3. Structural Genomics

3.1 Identifying sensitive individuals

Toxicogenetics is a discipline that evaluates genetic sequence variations that may impact individual's susceptibility to toxicity. In the pharmaceutical arena, these efforts are largely aimed at identification of susceptible individuals within the population that enables personalized drug treatment and improves drug safety. It is well accepted that genetic variants affect responses to drugs^{39, 40}. One of the first studies in this field compared plasma drug half-lives in identical and fraternal twin pairs and showed that greater differences existed between fraternal twins⁴¹. Recently, several more monogenic toxicogenetic traits have been reported⁴². Such research has yielded some success into the identification of genetic alleles that predict drug responses, for example it has been demonstrated that *apolipoprotein E (APOE)-4/4* allele carriers are the worst responders to conventional Alzheimer's disease treatments⁴³.

Successes in this field have led to important regulatory action by the FDA that has allowed a number of drugs to remain on the market due to the availability of genetic testing. Current genetic tests include assigning the dose of 6-mercaptopurine based on the genotype of thiopurine *S*-methyltransferase (*TPMT*)^{44, 45}, and the dose of warfarin based on the genotypes of vitamin K epoxide reductase complex (*VKORC1*)⁴⁶ and *CYP2C9*⁴⁷. In the case of the anti-coagulant drug warfarin, the combination of *VKORC1* haplotyping and *CYP2C9* genotyping explained an estimated 25% and 6–10% of the variance in warfarin dose, respectively⁴⁸. In fact, warfarin represents one of the first drugs for which toxicogenetic information (by *VKORC1* haplotyping) better explained the dose variance than pharmacokinetic pharmacogenomic data (through *CYP2C9* genotyping)⁴⁹. The data on warfarin, therefore, represents a shift from monogenic toxicogenetic testing to a polygenic model, which is expected to be encountered with increasing frequency as research in this field expands to include a greater variety of pharmaceuticals.

The wealth of information on genetic polymorphisms now available through the Human Genome Project has led to a dramatic increase in studies that seek to connect genetic variants with toxicity and pharmacologic phenotypes. Because the base pair sequence variation among individuals averages to be about 1 in 500–1000 base pairs⁵⁰, it is reasonable to expect that a significant number of genes will contain polymorphisms that contribute to disease and that many will play a role in adverse drug responses. Most of the current research focuses on the association between phenotype/disease and single nucleotide polymorphisms (SNPs). SNPs are an attractive choice for biomarkers of adverse responses because, unlike other factors which

contribute to a toxicity phenotype such as age, co-morbidity, and environment, an individual's genetic code remains stable throughout their lifetime. Genetic testing offers the potential of replacing empirical dose adjustment for many drugs that is based upon therapeutic assessment of pharmacologic or toxic effect after initial dosing. In addition, predictive genetic tests could also be of value in the drug development process by rescuing drugs that failed Phase III clinical trials due to toxicity within a subset of participants⁵¹. A key example of this is the genetic testing available to patients with HIV who are prescribed the drug abacavir, in which screening for major histocompatibility complex, class I, B (HLA-B)*5701 reduces the risk of hypersensitivity reactions⁵².

3.2 Toxicogenetic models and approaches

A number of publications have reported a significant association between SNPs and disease phenotypes in many areas of research; however few have been validated and have often been followed by reports that refute the original conclusion⁵³. Reasons for discrepancies between genotype-and-phenotype are numerous, but are often due to a low sample size in the study, population stratification of alleles, and heterogeneity of phenotypic classification^{54, 55}. To better facilitate a clinical translation of toxicogenetic data, efficient and validated strategies are needed. Clinical data has been used to determine the genetic basis of complex traits such as Parkinson disease⁵⁶ and susceptibility to HIV infection⁵⁷.

Classical approaches in this area have focused on mapping quantitative trait loci (QTL) within the genome that influence a specific phenotype. While there are many approaches to identify QTL, all involve a population of individuals with a measurable phenotype, a database of genotypic variation present within that population, and statistical measures that serve to link the magnitude of the phenotype with a specific genotype or polymorphism⁵⁸. Classical approaches have often sought to utilize the genotypic and phenotypic diversity present in F2 or backcross mouse populations; however, this approach is limited by the necessity to genotype all individuals within a population. Due to the relatively low number of recombination events in F2 and backcross populations, identification of precise QTL locations is often more difficult. Recombinant inbred lines offer the advantage of fixed genomes, but these lines can be expensive to acquire and maintain.

A promising new alternative is a method known as genomic association mapping in classical inbred mouse lines, which takes advantage of the genetic variation that arose naturally across inbred mouse lines over decades of crosses and inbreeding by scientists and fanciers. In this approach, large SNP datasets for several dozen strains and a database of over 8 million SNPs for 15 inbred mouse strains have recently become publicly available⁵⁹. Computational methods for this approach have been described^{58, 60}.

Applications of genomic association algorithms to a toxicogenetic approach have been performed *in vivo* within inbred mouse strains to determine additional genetic factors that affect the metabolism of warfarin⁶¹. In that study, the metabolism of warfarin, specifically the generation of 7-hydroxywarfarin, was shown to vary across inbred strains and this phenotype was then computationally associated with the mouse genomic region that encodes for *Cyp2c* enzymes. Experimental validation narrowed the list of potential genomic candidates to show that *Cyp2c29* polymorphisms altered hepatic protein expression of this enzyme. In a subsequent publication⁶², this same group demonstrated the utility of the approach for use in an *in vitro* drug biotransformation system in which genomic association was performed for downstream metabolism of testosterone and irinotecan across 15 mouse lines. The results of these studies showed that genetic variation within the *Cyp2b9* and *Ugt1a* loci influenced the metabolism of α -hydroxytestosterone⁶¹ and irinotecan glucuronidation⁶², respectively; these results were then confirmed experimentally via recombinant enzymes. These results suggest that genomic association using inbred mouse strains has the potential to aid in identification of toxicogenetic

alleles, however further studies that demonstrate a clear translation to human populations is required.

The genetic variation within panels of inbred mouse strains, while greater than that in RI lines, can be somewhat limited due to the fact that the majority of classical inbred strains are derived from *Mus musculus domesticus*. To address this concern, the Collaborative Cross mouse strains were recently developed and were designed to incorporate large genetic variation⁶³. To develop this resource, a controlled breeding program was designed to randomize genetic elements among the progeny derived from eight parental strains. A recent study demonstrated that the genetic variation present in the Collaborative Cross represents the optimal polymorphism architecture for the study of systems biology when compared to RI lines or to panels of classical inbred strains, and was demonstrated to be more reflective of the genetic variation expected in natural populations⁶⁴. Further characterization of this resource may improve the translational ability of toxicogenetic studies in rodents.

It has been suggested that genome-wide association studies that focus on only the top several “most significant” SNPs may have some limitations that could be overcome by a pathway-based approach. These limitations are that genes that contribute in a smaller, but still significant, amount to disease risk may be overlooked, and that variants that confer a large effect may not be included if hundreds of thousands of markers have been tested and the sample size is relatively small. Wang et al. have proposed a method in which the power to detect causal mechanisms of disease may be more robust because multiple contributing factors are considered together, as opposed to focusing on a few SNPs with the highest association score⁶⁵. This approach, which combines genome-wide association with the gene-set enrichment algorithm⁶⁶, identified a number of pathways that may be associated with Parkinson disease and age-related macular degeneration.

3.3 Expression QTL – bridging functional and structural genomics

Toxicogenomic and toxicogenetic approaches can be utilized in concert to study the regulation of gene expression that affects the differential toxicity susceptibility between individuals. When combined, these datasets have the potential to provide a mechanistic foundation for determining key pathways where adjuvant therapies could be developed for the prevention of toxicity in genetically sensitive individuals. While this field is still developing, a few proof-of-concept studies were recently published that have integrated QTL analysis with gene expression data; these include a determination of genes that were implicated in insulin resistance⁶⁷ and in allergic asthma⁶⁸. Notably, this approach has been adapted to the toxicology field by using recombinant BxD inbred lines derived from C57BL/6J and DBA/2J mice to uncover genes that regulate acute alcohol withdrawal⁶⁹ and that regulate lipopolysaccharide (LPS)-induced airway disease⁷⁰.

In a combined expression QTL (eQTL) approach, the gene expression across thousands of genes serves as the phenotype for genome-wide association studies. eQTL data generate a comprehensive picture of the regulation of gene expression on a global scale and of the differences that exist between sensitive and resistant individuals. A recent paper evaluated basal liver gene expression in the genetic context of BxD recombinant inbred and parental mouse strains to identify polymorphic local and distant QTL that control expression of a wide variety of genes in the liver⁶⁹. When comparing this dataset to similar data generated from nervous system tissue⁷¹, tissue-specific differences in the regulation of mRNA transcript expression were identified.

4. Data integration between -omics technologies

The need to better describe biological systems has led to the study of systems toxicology. Systems toxicology comprises the integration of genetics, toxicogenomics and conventional toxicity endpoints into a systems biology approach. These integrative approaches often use mouse models for an improved characterization of toxicity pathways, the discovery of new molecular and cellular indicators of exposure and outcome, better dose-response assessment and for more accurate inter-individual/cross-species extrapolations (Figure 1).

While microarray-based approaches in toxicity studies generate a wealth of data on gene expression and pathways that are affected by treatment or that are associated with a particular phenotype, these data are most often descriptive and may not reflect changes at the protein level. A key limitation of an analysis of toxicogenomic data alone is that these data often do not take into account confounding factors such as the pharmacokinetics of the test chemical⁷² and environmental factors, such the gut microflora population⁷³. Genomic studies, while more comprehensive than high-throughput metabolite or protein analysis, fail to characterize the full complement of cellular proteins which are subject to post-translational modifications and additional regulation⁷⁴. To address this concern and to develop additional toxicity biomarkers, metabolomics and proteomic approaches are gaining popularity, however metabolomics data is more frequently utilized due to its relative ease of use and interpretation.

Methods of data integration for orthogonal omics datasets fall broadly into two categories: 1) biology-driven strategies and 2) data-driven strategies⁷⁵. Biological interpretations tend to be more hypothesis-driven and take advantage of a contextual understanding of what is currently known about the toxicity mechanism. An example of a biology-driven strategy is a study that integrated gene expression data with metabolomics to identify sensitive biomarkers of human neuroendocrine (NE) cancers⁷⁶. In this approach, microarray and metabolite data was collected from NE *in vivo* mouse tumors, mouse NE tumor cell cultures, and primary human NE tumors and the authors identified from the gene transcript data key metabolism pathways that were altered and validated by mass spectrometry.

Data-driven strategies are comprised of unsupervised methods such as principal components analysis, self-organizing maps, and regression-based partial-least square analysis. While biology-driven interpretations have been more commonly reported, data-driven techniques have been utilized to study diverse areas, such as hepatocellular metabolism⁷⁷, the unique effects of anti-inflammatory compounds on macrophage cells⁷⁸, and the identification of regulatory metabolites and gene transcription in glucosinolate and anthocyanin biosynthesis⁷⁹. Another comprehensive example is a 'pharmaco-metabolomic' approach in which pre-dose metabolite profiles of urine and chemometrics were used to predict toxicity responses to acetaminophen in an *in vivo* rat model⁸⁰. Genomic, proteomic and metabolomic approaches, when analyzed alone, are limited in their ability to describe the great degree of nonlinearity and complexity that is present within biological systems.

Databases and approaches are currently being developed that can utilize the synergy between these data types with the goal of placing toxicogenomics data into a larger biological perspective. These include the Comparative Toxicology Database (ctd.mdibl.org), the Distributed Structure-Searchable Toxicity database (www.epa.gov/nct/dsstox/index.html), the Critical Path Institute (www.c-path.org), and systems biology databases developed by Genelogic, Inc. (www.genelogic.com), Iconix Pharmaceuticals (www.iconixbiosciences.com) and Ceetox (www.ceetox.com). As more data is added to publicly- and privately-funded databases, the data can be re-used and integrated to better inform hypothesis-driven research⁸¹.

5. Expert Opinion

Without a doubt, toxicogenomics offers great promise for risk assessment and omics technologies have the potential to transform chemical safety testing and to improve clinical intervention. A number of proof-of-principle studies have shown the utility of toxicogenomic data to predict adverse health events and injurious phenotypes. Even though toxicogenomics is unlikely to replace classical toxicological testing paradigms, it enriches the field by allowing researchers to discover biomarkers and subtle changes that occur sub-clinically or that predispose toward an injurious event. The reality is that omics technologies will become increasingly integral to toxicity studies; therefore, it is important that toxicologists have sufficient training and education to be able to design and interpret data from toxicogenomic and toxicogenetic studies.

The realization of the full potential of omics data integration is still several years away due to our currently limited understanding of the complex biology behind many toxicity types in a variety of species. There has been much progress in developing experimental tools to collect data, and some progress in data analysis and interpretation through pathway-based approaches and identification of complex interactions between gene expression, SNPs, other cellular macromolecules, and phenotypes. Focus should be placed on finding the optimal approaches for the integration and interpretation of multiple data types together in order to fully characterize molecular perturbations following exposure to a xenobiotic.

Another arena which requires significant attention from the toxicological community is in improving the use of omic data in high-throughput screening in human and rodent cell lines. Current methods which employ animal-based test methods tend to be expensive, are low-throughput, and often use doses that are much higher than potential human exposures. Toxicogenomic strategies that examine rodent *in vivo* with *in vitro* and human data are critical for bridging the uncertainties between species. While the value of *in vivo* data can not be completely replaced by *in vitro* screening tools, we anticipate that genomics tools will improve the value of *in vitro* screening for risk assessment. One current limitation that must be overcome is that very detailed gene annotation is only available for a few mammalian species, such as man and mouse, but is relatively lacking in others, such as in rat. The availability of complete gene information is critical for understanding and extrapolating toxicity risks across model systems and species and, ultimately, to humans.

Furthermore, regulatory agencies, the pharmaceutical industry, and academics must establish clear guidelines for the integration of toxicogenomic and genetic data into drug safety assessment. Evidence of clinical or human *in vitro* translation of toxicogenomic is essential in order to properly interpret the data and to be able to fully distinguish pathways that are associated with toxic effect vs. a physiological or pharmacological response. In examining datasets of sensitive “omics” data, it is necessary to distinguish between changes that are adaptive in nature and those changes that are within a no observable adverse effect level (NOAEL). Determination of chemical NOAELs can only be accomplished by dose-response modeling of a variety of compounds in which a benchmark dose can be correlated with both transcript changes and *in vivo* pathology; some progress has been made into developing computational tools that enable genomic benchmark dose analysis⁸². The question of determining the NOAEL becomes more complex as omics technologies become more refined, with greater throughput and sensitivity, and as data from multiple platforms becomes integrated.

The next wave of toxicogenomic and genetic research is the individualization of toxicity prediction and the ability to tailor drug therapy for specific patients. This area will be greatly facilitated by the development of easy and inexpensive whole-genome genotyping on an

individual basis. In order to apply structural genomics data to pharmaceutical development and to clinical safety testing, carefully designed rodent models are needed that can serve as appropriate population-based models. While there is promising data that shows that genome-wide association can predict human toxicity responses to xenobiotics, unequivocal data that demonstrates that this method can identify human “susceptibility” genes is still lacking. Current mouse models used for toxicity testing are not optimized for modeling the genetic diversity present within human populations. Because genome association using inbred or complex recombinant-inbred lines is an emerging technique, we expect that the animal models will be refined to facilitate personalized assessment of efficacy and toxicity. We anticipate that controlled mouse genetic breeding programs, such as the Collaborative Cross⁶³ that enable a superior precision in genetic mapping than current resources⁸³, will revolutionize the ability to detect the genetic control of toxicity in individuals. Translational research bridging rodent and human toxicity will be the key to the success of this field and great emphasis should be given to collaboration across laboratories and disciplines. This effort will require team research within centers and consortia in order to be fully realized.

Reference List

1. Gutman S, Hackett J. Search for shortcuts on the critical path to market: US FDA perspectives from the diagnostic side. *Pharmacogenomics* 2006 Dec;7(8):1223–1227. [PubMed: 17184209]
2. U.S. Food and Drug Administration. Guidance for Industry: Pharmacogenomic Data Submissions. 2005.
3. Frueh FW. Impact of microarray data quality on genomic data submissions to the FDA. *Nat Biotechnol* 2006 Sep;24(9):1105–1107. [PubMed: 16964222]
4. Bammler T, Beyer RP, Bhattacharya S, et al. Standardizing global gene expression analysis between laboratories and across platforms. *Nat Methods* 2005;2(5):351–356. [PubMed: 15846362]
5. Shi L, Reid LH, Jones WD, et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006 Sep;24(9):1151–1161. [PubMed: 16964229]
6. Beyer RP, Fry RC, Lasarev MR, et al. Multicenter study of acetaminophen hepatotoxicity reveals the importance of biological endpoints in genomic analyses. *Toxicol Sci* 2007 Sep;99(1):326–337. [PubMed: 17562736]
7. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 2004 Mar;36(3):299–303. [PubMed: 14981516]
8. Arcellana-Panlilio M, Robbins SM. Cutting-edge technology. I. Global gene expression profiling using DNA microarrays. *Am J Physiol Gastrointest Liver Physiol* 2002 Mar;282(3):G397–G402. [PubMed: 11841989]
9. Paules R. Phenotypic anchoring: Linking cause and effect. *Environ Health Perspect* 2003;111(6):A338–A339. [PubMed: 12760838]
10. Powell CL, Kosyk O, Ross PK, et al. Phenotypic anchoring of acetaminophen-induced oxidative stress with gene expression profiles in rat liver. *Toxicol Sci* 2006 Sep;93(1):213–222. [PubMed: 16751229]
11. Heinloth AN, Irwin RD, Boorman GA, et al. Gene expression profiling of rat livers reveals indicators of potential adverse effects. *Toxicol Sci* 2004 Jul;80(1):193–202. [PubMed: 15084756]*One of the first toxicogenomics papers that demonstrated the sensitivity of microarray data analysis for detecting organ toxicity.
12. Wang EJ, Snyder RD, Fielden MR, et al. Validation of putative genomic biomarkers of nephrotoxicity in rats. *Toxicology* 2008 Apr 18;246(2–3):91–100. [PubMed: 18289764]
13. Khor TO, Ibrahim S, Kong AN. Toxicogenomics in drug discovery and drug development: potential applications and future challenges. *Pharm Res* 2006 Aug;23(8):1659–1664. [PubMed: 16858654]
14. Ridd K, Zhang SD, Edwards RE, et al. Association of gene expression with sequential proliferation, differentiation and tumor formation in murine skin. *Carcinogenesis* 2006 Aug;27(8):1556–1566. [PubMed: 16537558]

15. Christians U, Schmitz V, Schoning W, et al. Toxicodynamic therapeutic drug monitoring of immunosuppressants: promises, reality, and challenges. *Ther Drug Monit* 2008 Apr;30(2):151–158. [PubMed: 18367974]
16. Bushel PR, Heinloth AN, Li J, et al. Blood gene expression signatures predict exposure levels. *Proc Natl Acad Sci U S A* 2007 Nov 13;104(46):18211–18216. [PubMed: 17984051]**This study is an example of translational toxicogenomics research which shows that blood mRNA is a sensitive predictor of chemical exposure levels that cause liver toxicity.
17. Burczynski ME, McMillian M, Ciervo J, et al. Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. *Toxicol Sci* 2000 Dec;58(2):399–415. [PubMed: 11099651]
18. Waring JF, Ulrich RG, Flint N, et al. Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. *Environ Health Perspect* 2004 Mar;112(4):439–448. [PubMed: 15033593]
19. Hamadeh HK, Bushel PR, Jayadev S, et al. Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 2002 Jun;67(2):219–231. [PubMed: 12011481]
20. Inadera H, Tachibana S, Takasaki I, et al. Expression profile of liver genes in response to hepatotoxicants identified using a SAGE-based customized DNA microarray system. *Toxicol Lett* 2008 Feb 28;177(1):20–30. [PubMed: 18243596]
21. Dai X, He YD, Dai H, et al. Development of an approach for ab initio estimation of compound-induced liver injury based on global gene transcriptional profiles. *Genome Inform* 2006;17(2):77–88. [PubMed: 17503381]
22. Fielden MR, Brennan R, Gollub J. A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicol Sci* 2007 Sep;99(1):90–100. [PubMed: 17557906]*Study suggesting that gene expression profiling is a sensitive predictor of carcinogenesis in rodent two-year bioassay.
23. Eun JW, Ryu SY, Noh JH, et al. Discriminating the molecular basis of hepatotoxicity using the large-scale characteristic molecular signatures of toxicants by expression profiling analysis. *Toxicology* 2008 Jul 30;249(2–3):176–183. [PubMed: 18579281]
24. Bartosiewicz MJ, Jenkins D, Penn S, et al. Unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants. *J Pharmacol Exp Ther* 2001 Jun;297(3):895–905. [PubMed: 11356909]
25. McMillian M, Nie AY, Parker JB, et al. Inverse gene expression patterns for macrophage activating hepatotoxicants and peroxisome proliferators in rat liver. *Biochem Pharmacol* 2004 Jun 1;67(11):2141–2165. [PubMed: 15135310]
26. Harries HM, Fletcher ST, Duggan CM, Baker VA. The use of genomics technology to investigate gene expression changes in cultured human liver cells. *Toxicol In Vitro* 2001 Aug;15(4–5):399–405. [PubMed: 11566570]
27. Uehara T, Kiyosawa N, Shimizu T, et al. Species-specific differences in coumarin-induced hepatotoxicity as an example toxicogenomics-based approach to assessing risk of toxicity to humans. *Hum Exp Toxicol* 2008 Jan;27(1):23–35. [PubMed: 18480146]**One of the first explorations into creating an *in vivo-in vitro* bridge for the validation of toxicogenomic biomarkers.
28. Elferink MG, Olinga P, Draaisma AL, et al. Microarray analysis in rat liver slices correctly predicts *in vivo* hepatotoxicity. *Toxicol Appl Pharmacol* 2008 Jun 15;229(3):300–309. [PubMed: 18346771] *Study that suggests that a rat liver slice model can be used successfully to model *in vivo* toxicity responses.
29. Lerche-Langrand C, Toutain HJ. Precision-cut liver slices: characteristics and use for *in vitro* pharmacotoxicology. *Toxicology* 2000 Nov 16;153(1–3):221–253. [PubMed: 11090959]
30. Adachi Y, Bradford BU, Gao W, et al. Inactivation of Kupffer cells prevents early alcohol-induced liver injury. *Hepatology* 1994 Aug;20(2):453–460. [PubMed: 8045507]
31. Wheeler MD, Kono H, Yin M, et al. The role of Kupffer cell oxidant production in early ethanol-induced liver disease. *Free Radic Biol Med* 2001 Dec 15;31(12):1544–1549. [PubMed: 11744328]
32. Kumagai K, Kiyosawa N, Ito K, et al. Influence of Kupffer cell inactivation on cycloheximide-induced hepatic injury. *Toxicology* 2007 Nov 30;241(3):106–118. [PubMed: 17900782]

33. Andres D, Sanchez-Reus I, Bautista M, Cascales M. Depletion of Kupffer cell function by gadolinium chloride attenuates thioacetamide-induced hepatotoxicity. Expression of metallothionein and HSP70. *Biochem Pharmacol* 2003 Sep 15;66(6):917–926. [PubMed: 12963478]
34. Ito Y, Bethea NW, Abril ER, McCuskey RS. Early hepatic microvascular injury in response to acetaminophen toxicity. *Microcirculation* 2003 Oct;10(5):391–400. [PubMed: 14557822]
35. National Research Council Toxicity Testing in the 21st Century: A Vision and a Strategy 2007 Washington, DC National Academies Press**A comprehensive volume with broad recommendations for improving toxicity testing to better inform risk assessment.
36. Zeeberg BR, Feng W, Wang G, et al. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* 2003;4(4):R28. [PubMed: 12702209]
37. Barry WT, Nobel AB, Wright FA. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 2005 May 1;21(9):1943–1949. [PubMed: 15647293]
38. Khatri P, Bhavsar P, Bawa G, Draghici S. Onto-Tools: an ensemble of web-accessible, ontology-based tools for the functional design and interpretation of high-throughput gene expression experiments. *Nucleic Acids Res* 2004 Jul 1;32(Web Server issue):W449–W456. [PubMed: 15215428]
39. Burns JJ IV. Pharmacogenetics and drug toxicity. Variation of drug metabolism in animals and the prediction of drug action in man. *Ann N Y Acad Sci* 1968 Jul 31;151(2):959–967. [PubMed: 4189972]
40. Kalow W. Contribution of hereditary factors to the response to drugs. *Fed Proc* 1965 Nov;24(6):1259–1265. [PubMed: 5853146]*One of the original papers that demonstrated that genetics can affect drug responses.
41. Vesell ES, Page JG. Genetic control of dicumarol levels in man. *J Clin Invest* 1968 Dec;47(12):2657–2663. [PubMed: 4177776]
42. Nebert DW, Zhang G, Vesell ES. From human genetics and genomics to pharmacogenetics and pharmacogenomics: past lessons, future directions. *Drug Metab Rev* 2008;40(2):187–224. [PubMed: 18464043]
43. Cacabelos R. Pharmacogenetic basis for therapeutic optimization in Alzheimer's disease. *Mol Diagn Ther* 2007;11(6):385–405. [PubMed: 18078356]
44. Yates CR, Krynetski EY, Loennechen T, et al. Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. *Ann Intern Med* 1997 Apr 15;126(8):608–614. [PubMed: 9103127]
45. Relling MV, Hancock ML, Rivera GK, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J Natl Cancer Inst* 1999 Dec 1;91(23):2001–2008. [PubMed: 10580024]
46. Rost S, Fregin A, Ivaskevicius V, et al. Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature* 2004 Feb 5;427(6974):537–541. [PubMed: 14765194]
47. Higashi MK, Veenstra DL, Kondo LM, et al. Association between CYP2C9 genetic variants and anticoagulation-related outcomes during warfarin therapy. *J Am Med Assoc* 2002 Apr 3;287(13):1690–1698.
48. Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. *N Engl J Med* 2005 Jun 2;352(22):2285–2293. [PubMed: 15930419]
49. Weinshilboum RM, Wang L. Pharmacogenetics and pharmacogenomics: development, science, and translation. *Annu Rev Genomics Hum Genet* 2006;7:223–245. [PubMed: 16948615]**An comprehensive review of pharmacogenetics and pharmacogenomics research.
50. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001 Feb 16;291(5507):1304–1351. [PubMed: 11181995]
51. Weiss ST, McLeod HL, Flockhart DA, et al. Creating and evaluating genetic tests predictive of drug response. *Nat Rev Drug Discov* 2008 Jul;7(7):568–574. [PubMed: 18587383]
52. Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. *N Engl J Med* 2008 Feb 7;358(6):568–579. [PubMed: 18256392]
53. Nebert DW, Vesell ES. Advances in pharmacogenomics and individualized drug therapy: exciting challenges that lie ahead. *Eur J Pharmacol* 2004 Oct 1;500(1–3):267–280. [PubMed: 15464039]

54. Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003 Feb 15;361(9357):567–571. [PubMed: 12598142]
55. Ntzani EE, Rizos EC, Ioannidis JP. Genetic effects versus bias for candidate polymorphisms in myocardial infarction: case study and overview of large-scale evidence. *Am J Epidemiol* 2007 May 1;165(9):973–984. [PubMed: 17293603]
56. Lesnick TG, Papapetropoulos S, Mash DC, et al. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* 2007 Jun;3(6):e98. [PubMed: 17571925]
57. Donfack J, Buchinsky FJ, Post JC, Ehrlich GD. Human susceptibility to viral infection: the search for HIV-protective alleles among Africans by means of genome-wide studies. *AIDS Res Hum Retroviruses* 2006 Oct;22(10):925–930. [PubMed: 17067260]
58. McClurg P, Pletcher MT, Wiltshire T, Su AI. Comparative analysis of haplotype association mapping algorithms. *BMC Bioinformatics* 2006;7:61. [PubMed: 16466585]
59. Frazer KA, Eskin E, Kang HM, et al. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* 2007 Aug 30;448(7157):1050–1053. [PubMed: 17660834]
60. Pletcher MT, McClurg P, Batalov S, et al. Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. *PLoS Biol* 2004 Dec;2(12):e393. [PubMed: 15534693]
61. Guo Y, Weller P, Farrell E, et al. In silico pharmacogenetics of warfarin metabolism. *Nat Biotechnol* 2006 May;24(5):531–536. [PubMed: 16680137]*Study that used genetic association among inbred mouse strains to identify a functional polymorphism that affects warfarin metabolism.
62. Guo Y, Lu P, Farrell E, et al. In silico and in vitro pharmacogenetic analysis in mice. *Proc Natl Acad Sci U S A* 2007 Nov 6;104(45):17735–17740. [PubMed: 17978195]*Study demonstrating the utility of using panels of inbred strains for pharmacogenetic analysis.
63. Churchill GA, Airey DC, Allayee H, et al. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat Genet* 2004 Nov;36(11):1133–1137. [PubMed: 15514660]
64. Roberts A, Pardo-Manuel dV, Wang W, et al. The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. *Mamm Genome* 2007 Jul;18(6–7):473–481. [PubMed: 17674098]*Study demonstrating that Collaborative Cross mouse strains are an improvement for use in systems biology research over RI lines and inbred mouse panels.
65. Wang K, Li M, Bucan M. Pathway-Based Approaches for Analysis of Genomewide Association Studies. *Am J Hum Genet* 2007 Oct 26;81(6)
66. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005 Oct 25;102(43):15545–15550. [PubMed: 16199517]
67. Collison M, Glazier AM, Graham D, et al. Cd36 and molecular mechanisms of insulin resistance in the stroke-prone spontaneously hypertensive rat. *Diabetes* 2000 Dec;49(12):2222–2226. [PubMed: 11118030]
68. Karp CL, Grupe A, Schadt E, et al. Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nat Immunol* 2000 Sep;1(3):221–226. [PubMed: 10973279]
69. Hitzemann R, Reed C, Malmanger B, et al. On the integration of alcohol-related quantitative trait loci and gene expression analyses. *Alcohol Clin Exp Res* 2004 Oct;28(10):1437–1448. [PubMed: 15597075]
70. Cook DN, Wang S, Wang Y, et al. Genetic regulation of endotoxin-induced airway disease. *Genomics* 2004 Jun;83(6):961–969. [PubMed: 15177550]
71. Chesler EJ, Lu L, Shou S, et al. Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat Genet* 2005 Mar;37(3):233–242. [PubMed: 15711545]
72. Nebert DW, Jorge-Nebert L, Vesell ES. Pharmacogenomics and "individualized drug therapy": high expectations and disappointing achievements. *Am J Pharmacogenomics* 2003;3(6):361–370. [PubMed: 14672516]
73. Phipps AN, Stewart J, Wright B, Wilson ID. Effect of diet on the urinary excretion of hippuric acid and other dietary-derived aromatics in rat. A complex interaction between diet, gut microflora and substrate specificity. *Xenobiotica* 1998 May;28(5):527–537. [PubMed: 9622854]

74. Schrattenholz A, Ss Kic V. What does systems biology mean for drug development? *Curr Med Chem* 2008;15(15):1520–1528. [PubMed: 18537627]
75. Thomas CE, Ganji G. Integration of genomic and metabonomic data in systems biology--are we 'there' yet? *Curr Opin Drug Discov Devel* 2006 Jan;9(1):92–100.**An comprehensive review of omics data integration strategies and current research.
76. Ippolito JE, Xu J, Jain S, et al. An integrated functional genomics and metabolomics approach for defining poor prognosis in human neuroendocrine cancers. *Proc Natl Acad Sci U S A* 2005 Jul 12;102(28):9901–9906. [PubMed: 15998737]
77. Li Z, Srivastava S, Mittal S, et al. A Three Stage Integrative Pathway Search (TIPS) framework to identify toxicity relevant genes and pathways. *BMC Bioinformatics* 2007;8:202. [PubMed: 17570844]
78. Verhoeckx KC, Bijlsma S, de Groene EM, et al. A combination of proteomics, principal component analysis and transcriptomics is a powerful tool for the identification of biomarkers for macrophage maturation in the U937 cell line. *Proteomics* 2004 Apr;4(4):1014–1028. [PubMed: 15048983]
79. Hirai MY, Klein M, Fujikawa Y, et al. Elucidation of gene-to-gene and metabolite-to-gene networks in arabidopsis by integration of metabolomics and transcriptomics. *J Biol Chem* 2005 Jul 8;280(27):25590–25595. [PubMed: 15866872]
80. Clayton TA, Lindon JC, Cloarec O, et al. Pharmacometabonomic phenotyping and personalized drug treatment. *Nature* 2006 Apr 20;440(7087):1073–1077. [PubMed: 16625200]*Manuscript suggesting that pre-dose metabolomic profiling can predict acetaminophen toxicity responses in a rat model.
81. Fostel JM. Future of toxicogenomics and safety signatures: balancing public access to data with proprietary drug discovery. *Pharmacogenomics* 2007 May;8(5):425–430. [PubMed: 17465705]
82. Yang L, Allen BC, Thomas RS. BMDExpress: a software tool for the benchmark dose analyses of genomic data. *BMC Genomics* 2007;8:387. [PubMed: 17961223]
83. Chesler EJ, Miller DR, Branstetter LR, et al. The Collaborative Cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. *Mamm Genome*. (In Press)

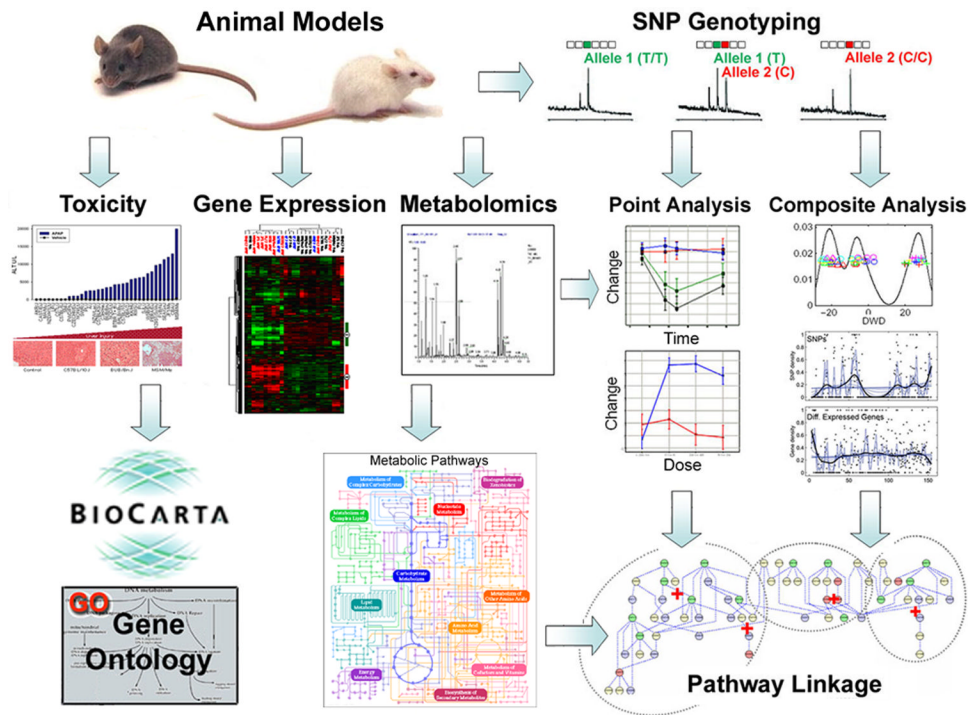


Figure 1.
Information flow in systems toxicology.