A NOVEL COMPUTATIONAL APPROACH FOR TOXICOGENOMICS BIOMARKER DISCOVERY IN DRUG DEVELOPMENT PIPELINE

MM Rana[1][*], MN Hasan[1,2], MS Ahmed[1] and MNH Mollah[1]

[1]*Bioinformatics Laboratory, Department of Statistics, University of Rajshahi, Bangladesh*
[2]*Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Bangladesh*

Abstract

In the early stage of drug development process, it is urgent to judge the toxicity effect of some common chemical compounds (CCs) that is not yet well investigated. Biomarker genes (BGs) and dose of CCs can help to draw a deduction about a drug for safety assessment. Classical toxicology method uses large number of samples to extract clinical results which is both time consuming and costly. However, conventional molecular methods can perform to identify only BGs and fail to detect source factor influencing these BGs. The aim of this study is to propose a suitable algorithm that can identify more promising and essential toxicity biomarkers related to some common CCs for safety assessment of new drugs. The *glutathione* is an effective metabolite of detoxification process in liver. *Glutathione* depletion analysis is one of the major key research areas in drug development pipeline. In this paper, we studied *glutathione* depletion analysis of some reported CCs (acetaminophen, methapyrilene and nitrofurazone). We develop an algorithm combining ANOVA and principal component analysis (PCA) using visualization technique to find biomarker genes and associated *glutathione* depleting CCs and their corresponding doses. There are numerous numbers of genes in the *glutathione* metabolism pathway regulated as differentially expressed (DE) genes due to the toxic effect of these CCs and proposed algorithm identify only five genes (Mgst2, Gclc, G6pd, Gsr and Srm) that are also foremost genes in the *glutathione* metabolism pathway. Proposed algorithm states that high dose of all the CCs are responsible for *glutathion*e depletion, nevertheless middle dose of acetaminophen and nitrofurazone also cause *glutathion*e depletion. The proposed algorithm has an additional benefit over the conventional method to discover new chemical entities toxicity.

Key words: ANOVA, biomarker, *glutathione*, microarray, PCA, toxicogenomics

Introduction

The chemicals whose have an adverse effect in human body are regarded as toxic agent but this adverse effects are extensively dependent on the chemicals exposure level. For a pharmaceutical drugs, the effectiveness and the toxicity depends on it dose level, overdose may cause toxicity. Moreover, sometimes the dose makes the thing toxic according to an old cliché "the dose makes the poison". These toxic compounds are disposed through four stages like absorption, distribution, metabolism and excretion in a biological system (Davis and Riley et al. 2004, Ekins et al. 2005). The liver is the largest internal and main organ in the human body for metabolism and detoxification of drugs and environmental chemicals (Klaassen et al. 2007). Hepatotoxicity is one of the major causes for withdrawals of some CCs from the market. However, most of the preclinical studies are designed based on liver organ. Classical toxicological research depends on the animal testing to determine the hazard of toxic chemical compound to humans based on well-established cytological, physiologic, metabolic, and morphologic endpoints (Suter et al. 2004) and require a large number of animals and huge time for drawing statistically significant conclusions to assess

---

[*]Author for correspondence: ringku_740@yahoo.com

the potential of tumarogenicity in animals and relevant risk in humans (Ulrich and Friend 2002, Waters and Fostel 2004). Therefore, the toxicological testing is of high cost in terms of time, labor, compound synthesis, and animals used. On the other side, toxicogenomics that has attracted widespread attention as an alternative means to study the underlying molecular mechanisms of toxicity and address challenges that are difficult to overcome by conventional toxicology methods (Nuwaysir et al. 1999) as well as gene expression (GE) analysis of animals' target organs after drug administration can help assessing potential toxicity before phenotypic appearance through the use of BGs (Fielden et al. 2007, Uehara et al. 2008). That's why toxicogenomics find its great application in the case of drug development and has become a powerful tool for revealing mechanisms underlying toxicological endpoints and a useful approach for the early detection of potential chemical toxicity (Heinloth et al. 2004, Irwin et al. 2004, Battershill et al. 2005, Searfoss et al. 2005, Kiyosawa et al. 2009).

The Japanese Toxicogenomics Project (TGP) systematically collected chemically-induced GE data from 2002 to 2010 and its first outcome was open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation Systems (TG-GATEs). In this study, we use TG-GATEs database that stores GE profiles and traditional toxicological data derived from rats (liver, kidney, and primary hepatocytes) and human (primary hepatocytes) cells after administrating 170 compounds at the combination of multiple dose and time levels (Igarashi et al. 2015).

In the liver, detoxification process is always continuing and glutathione plays the major role in this process by conjugating target toxic compounds and exports the conjugated compounds into bile ducts. To analyze drug induced glutathione depletion from TG-GATEs database, we discover common DE genes for some drugs that are affected by different factors likewise, doses and measurement time levels or combinations thereof. The ultimate goal of toxicogenomics study is to identify genes that are differentially expressed in the conditions (CCs-dose-time) being studied. Comparison statistics, for instance, fold-change, Mann–Whitney U test and t-test are frequently used to discover DE genes in toxicogenomics (Nyström-Persson et al. 2013). Fold-change is the log ratio between two conditions and a threshold value have been used to choose two fold difference whereas Mann–Whitney U test is a nonparametric test also produce same result as like as t-test, a statistical test in normal conditions. The result obtained from fold-change is not so suitable since it does not give any level of confidence to define a DE gene and it's identify wrong gene in case of low intensity with high variance (Newton et al. 2001, Rocke et al. 2001). In toxicogenomic experiment, we have different factors like dose and measurement time points and t-test is compare the expression level of control dose (samples) with the others dose levels by pooling all samples or make many comparisons, one for each different pair of conditions. For toxicogenomics data, t-test is not an efficient approach, because it does not fully utilize all of the information existing in the data and it produces low power because of small sample in each condition (Cui et al. 2003). ANOVA technique considers the variability of the expression levels in different factors. If the variability is present in the expression levels of a gene by the combined effect of dose and time, this indicates that the gene is differentially expressed (Pavlidis et al. 2003). In this study, we consider two-way ANOVA technique that provide interaction effect between dose and time levels of a compound on the expression value of a gene, thereby, the main effects of dose and time levels on that gene are no longer straight forward to interpret (Sahai et al. 2000). PCA can be used to separate dose-time depended cluster of treated CCs group from the control group (Hirode et al. 2008). In this study, combining two-way ANOVA and data visualization based on PCA result, data visualization based on PCA result using common DE genes enables one to identify a group of BGs and to select risk factor (dose of CCs in some toxic mechanisms). Using proposed algorithm, we can easily identify BGs and associated source factors. However, proposed algorithm result permit the histopathological results obtained from the same sample conditions and it is also possible to use proposed algorithm to discover new chemical entities toxicity which is describe in detail in conclusion section.

Materials and Methods

Computational approach

Proposed computational approach to analyses toxicogenomics experiment data are delineated in Fig. 1. Microarray experimental data contain thousands of variables, (i.e. genes) and only relatively few samples, so it is necessary to reduce the number of genes considered for further analysis. This can be done by restricting on a group of genes or by selecting a pathway. However, some compound are influencing a given pathway or biological process of particular interest. For some given compounds, we can reduce the data dimension by selecting a set of genes involving those compound induced pathway. Suppose, we have c compounds data $D_1, D_2, \cdots, D_c$ each have same variables or genes (N) and either same or different number of samples depending on dose-time selections of compound. Let us consider the sample size of each compounds are same (m), it seems a balance condition. Again, consider there are n genes are involved in a given pathway. To analyze this pathway, we diminish variables size N to n (after removing others gene from variable list that are not involved in this pathway).

Identification of DE genes

In this study, to identify DE genes for toxicity prediction, we have used two-way ANOVA model. Let us consider Y_{ijk} is a random variable of GE and y_{ijk} is the observed GE value of Y_{ijk} at i^{th} replication of the jth dose of a chemical and k^{th} measurement time level. In two-way ANOVA models, all these variables as varying independently and normally around a mean expression μ_{jk} with a constant variance σ^2. The two-way ANOVA is therefore given in equation (1).

$$Y_{ijk} = \mu_{jk} + \varepsilon_{ijk} \tag{1}$$

The mean expression value, $\mu_{jk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$; where, μ is the grand mean, α_j is the jth dose level effect, β_k is the kth time level effect, $(\alpha\beta)_{jk}$ is the interaction effect of the dose and k^{th} measurement time level and ε_{ijk} is the random error. Now the equation (1) can be written in a vector for a GE measure as

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \tag{2}$$

Where, $i = 1, 2, \cdots, n; j = 1, 2, \cdots, a; k = 1, 2, \cdots, b.$ The main objective of the above ANOVA model is to test whether all the main effects (α_j) of the dose levels, the main effects (β_k) of time levels or the interaction effects $(\alpha\beta)_{jk}$ of dose and time levels are statistically significant. Under null hypothesis, for testing the mentioned statement the F statistic are used. On the basis of this statistic, if the null hypothesis becomes rejected (i.e., the main or the interaction effects are significantly different). For each CCs, every gene are tested weather the expression level is changed by the effect of dose and time levels of given CCs. DE genes for each CCs are selected by following the condition of interaction effect p-value <0.05.
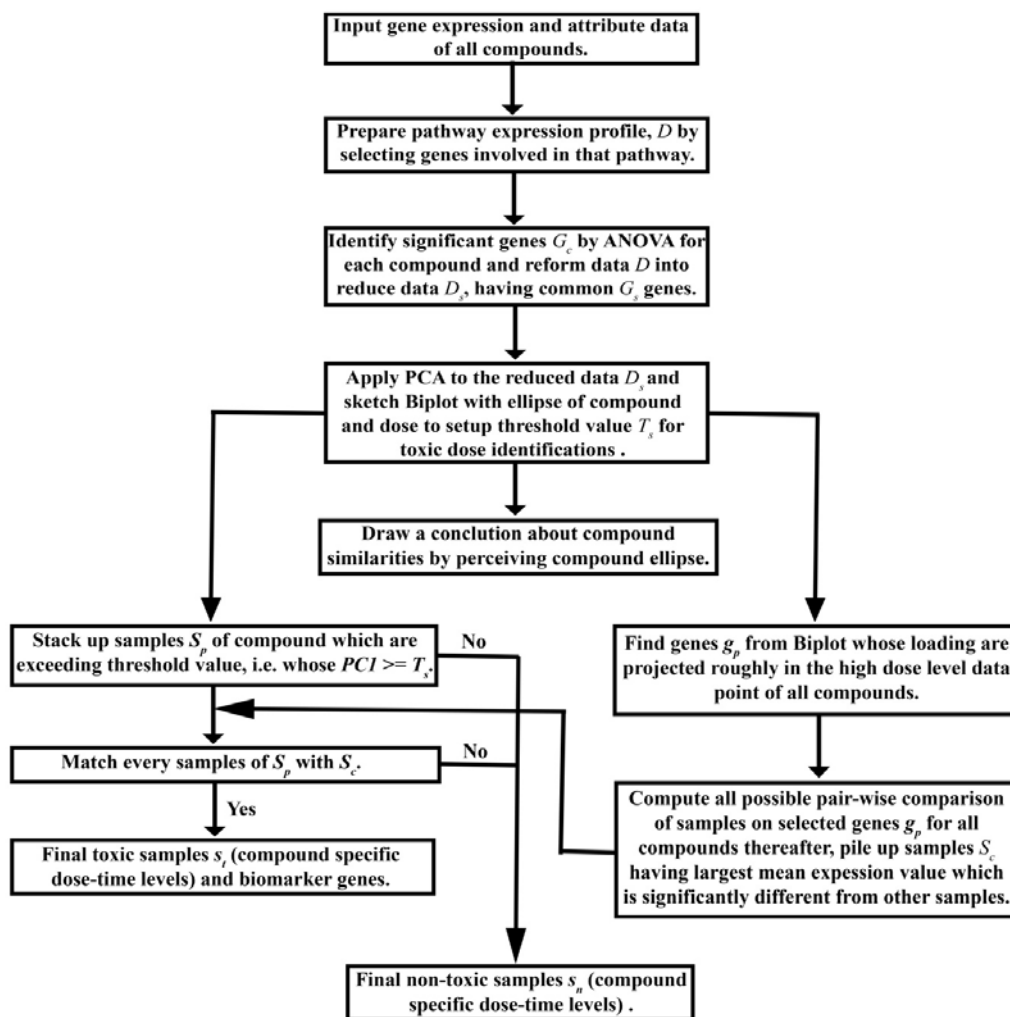
Fig. 1. Overview of the proposed Algorithm.

Identification of BGs and associated source factor

After identifying significant genes from each compound, we extracted common significant genes n_c from the gene list of all CCs which shows much alteration in their expression due to exposure chemicals. Now, apply PCA on the reduce data matrix D_s with dimensions $n_c \times m$. PCA is a promising multivariate technique which is widely used for dimension reduction but in this paper, we used PCA as a classifier or to measure the associations between genes and sample conditions (SCs) i.e., compound specific dose (-time). Combining data visualization and post-hoc testing, multiple comparison tests enables one to differentiate SCs and to find SCs considering expression similarities having highest mean. PCA biplot on significant GE data can be used

to extract the loading sources of variation from toxicogenomics data, to visualize associations between genes and SCs and select BGs from this relationship. We restricts PCA to the first principal component (PC) for selecting probable toxic CCs dose by setting threshold value as smallest absolute loading according to the PC1. It is also possible to make deductions about the groups or similarities of Compound as well as dose by perceiving the scatterings on biplot, similar compound induces GE value might scattered together in the biplot. It is well known that high dose of CCs are more toxic than other dose. Therefore, genes which are projected roughly in the high dose level of CCs are reported as key genes or BGs for toxicogenomic discovery. A gene is DE when it expression patterns shows significant differences in different dose, time or dose-time level. For a specific gene, some doses levels have similar as like control dose and some have different expression value. In the expression level of BGs, SCs are suspected as a toxic compound-dose (-times) whose have high mean expression value. Applying multiple comparisons (Tukey-HSD) post-hoc testing with the condition adjustment of p-value (Holm's), p-adj. $<$0.05 to the BGs, we identify putative SCs. Finally, SCs obtained from PCA and post-hoc test are matched and matching SCs are considered as source factor of these DE genes or the desirable compound-dose-time for toxicogenomics discovery. Results from proposed method are tested by the histopathological results to measure the performance of the proposed method.

Source of data

Chemically induced GE data or microarray data refers to toxicogenomics data. According to Nyström-Persson et al. (2013) acetaminophen, methapyrilene and nitrofurazone are the *glutathione* (a major metabolite in detoxification process) depleting compounds. CEL files for microarray GE data after administrating these three CCs were downloaded from TG-GATEs (http://toxico.nibiohn.go.jp/). TGP has been taken the scheme collecting high dimensional toxicogenomic data systematically since 2002 as a joint government-privet sector project (Uehara et al. 2010). *In vivo* and *in vitro* are the two main types of data have produced by the TGP, the *in vivo* data, which was collected from *Rattus Norvegicus* at four time points (3, 6, 9 and 24 hrs) for each of four dose levels (control, low, middle, high) from two organs (liver, kidney). In this study, we consider *Rattus Norvegicus*'s *in vivo* liver GE data of chemicals acetaminophen, methapyrilene and nitrofurazone.

Pre-processing and normalization of CEL file

CEL files are mainly image files, prior to computational analysis, it is necessary to quantify image file to expression values. Using the Bioconductor Affy package for R, raw files were pre-processed and mas5 function is used for normalization the GE values.

Results and Discussion

Glutathione depletion analysis

In this study, we have considered three reported glutathione depleting compounds acetaminophen, methapyrilene and nitrofurazone and their toxicity effect under different conditions (dose and time levels) were studied on the 42 probes that are belonging to the *glutathione* metabolism pathway. Examining p-values in the interaction effects between dose and time level, 15, 14 and 17 DE probes out of 42 probes were found for acetaminophen, methapyrilene and nitrofurazone respectively. The p-values of common DE genes in compounds are presented in Table 1. Based on gene ontology, the content of gene related to Glutathione metabolism, glutathione biosynthetic process, glutathione binding, aging, response to cadmium ion, glutathione metabolic process, cytosol were significantly high (Table 2). This feature might reveal that the GE changes in *glutathione* metabolism strongly associated with glutathione depletion or conjugation.
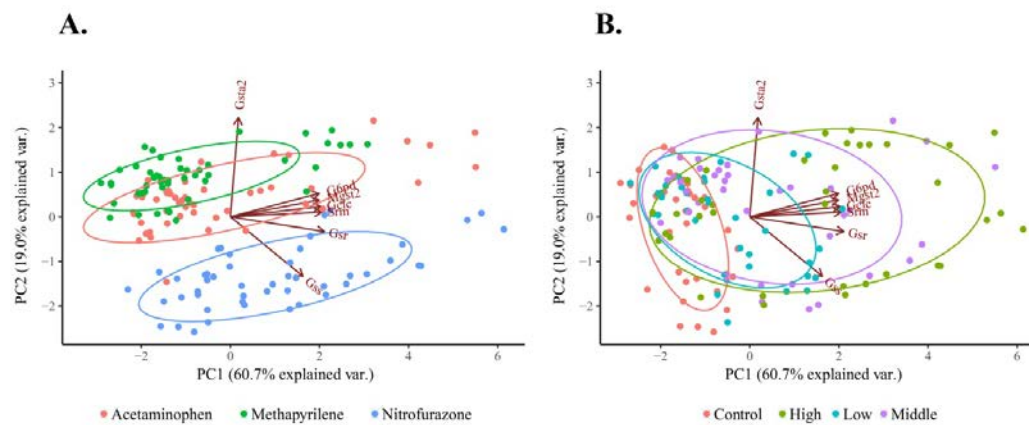
Fig. 2. PCA for common DE genes of all CCs. Biplot of chemical-dose-time data points and common DE genes loading to the first two principal components of the PCA. Different color represents different group. Classification using common DE genes for CCs 'A' (left) and Dose of all chemicals 'B' (right).

Now, we compute PCA on the reduced data matrix based on these seven genes and same SCs. Unless, PCA is generally applied for dimension reduction but in this study, we use PCA for association study. However, the first two components were highly significant contain almost 80% explained variance. The biplot with significant common genes in CCs can be used to correlate SCs and genes. Fig. 2 shows a biplot of SCs and common significant genes with normal ellipse for both CCs and doses. As portrayed in Fig. 2A, acetaminophen and methapyrilene nearly similar compound whereas nitrofurazone is totally different and it's imitate that the mechanism of nitrofurazone may be different from acetaminophen and methapyrilene. From Fig. 2B, we observed that the variation in the data is increases by the increment of doses but doses level don't have any separate cluster. We also observed that most of the low dose data points are scattered with control dose data point and increasing dose data points are scattered far away (especially high dose levels). The biplot enables the association of common DE genes with SCs. The genes, Gsr, Mgst2, Gclc, Srm and G6pd loadings directed according to the first two PCs towards high dose levels of CCs (Fig. 2).

Table 1. Common DE genes for all the CCs in glutathione metabolism pathway.

Gene ID	Gene name	Acetaminophen	Methapyrilene	Nitrofurazone
Gsr	Glutathione reductase	1.91E-07 (4.60E-05)	1.00E-06 (2.30E-05)	8.06E-11 (2.72E-07)
Mgst2	Microsomal glutathione S-transferase 2	1.00E-06 (5.70E-05)	2.00E-06 (0.014261)	3.67E-10 (1.47E-07)
Gclc	Glutamate-cysteine ligase - catalytic subunit	2.50E-05 (0.007903)	1.90E-05 (0.008478)	2.17E-08 (5.00E-06)
Srm	Spermidine synthase-like	0.000408 (0.017384)	4.20E-05 (0.008105)	1.60E-05 (0.993068)
G6pd	Glucose-6-phosphate dehydrogenase	0.001366 (0.026589)	0.000235 (0.22635)	0.000463 (0.700015)
Gss	Glutathione synthetase	0.009570 (0.032616)	0.013921 (0.200432)	0.000617 (0.047296)
Gsta2	Glutathione S-transferase alpha 2	0.017522 (0.006135)	0.023948 (0.341182)	0.001859 (0.906843)

P-values obtained from *t*-test are presented in bracket.
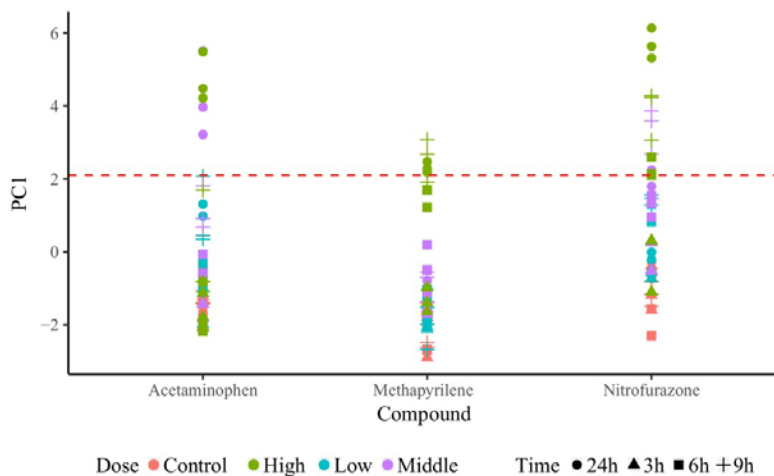
Fig. 3. Score plot for PC1. For each CCs, each data point is duplicated by the combination of different color and shape as shown bottom panel.

The vector of these five genes represent the loadings indicate that they have strong association with the high dose levels (Fig. 2). Therefore, by the proposed algorithm, genes Gsr (glutathione reductase), Mgst2 (microsomal glutathione S-transferase 2), Gclc (glutamate-cysteine ligase), Srm (spermidine synthase) and G6pd (glucose-6-phosphate dehydrogenase) are the biomarker genes for this study as well as glutathione depletion analysis. Gclc is known to accelerate glutathione synthesis and Mgst2 is glutathione-S-transferase, which is the main enzyme for detoxification of toxic compounds by the conjugation reaction. Gsr, G6pd and Srm is a central enzyme of cellular antioxidant defense and reduce oxidized glutathione disulfide (GSSG) to the sulfhydryl form glutathione. The loadings or the vectors indicate that these genes are highly correlated to each other and have strong association with toxic dose. To identify suspected toxic dose of CCs, we setup a threshold value for PC1 score by considering smallest length of these five genes vector. In this study, we chose the threshold value for PC1 score is 2.1 (Fig. 2).

Table 2. Go analysis of identified common seven DE genes.

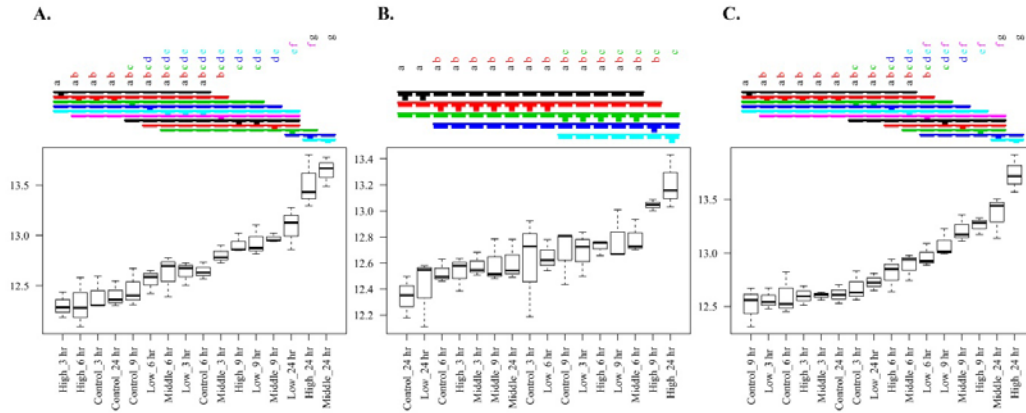| Term	Count	Percent	P-value
Glutathione metabolism	6	85.7	1.9E-10
Glutathione biosynthetic process	3	42.9	2.4E-4
Glutathione binding	3	42.9	2.5E-4
Aging	4	57.1	2.3E-3
Response to cadmium ion	3	42.9	1.8E-3
Glutathione metabolic process	3	42.9	1.8E-3
Cytosol	5	71.4	4.0E-3

Fig. 4. Boxplot in different SCs for Mgst2 gene. Lettering in the top panel represent similarity among SCs. Gene expression patterns for only gene mgst2 in different SCs of acetaminophen 'A', methapyrilene 'B' and nitrofurazone 'C' are displayed.

Only 24 h time point of high and middle doses in acetaminophen, 9 h and 24 h time points of high dose in methapyrilene, and 6h, 9h and 24 h of high dose and only 9h and 24 h of middle dose in nitrofurazone are exceeds the threshold value in PC1 score which data points might be responsible for glutathione depletion (Fig. 3). To validate PCA result, we now apply post-hoc testing criterion by the proposed method using these five significant genes. Applying post-hoc test, we found 9 h of low and middle doses, 24 h of high and middle doses in acetaminophen, 6 h, 9 h and 24 h time points of high dose in methapyrilene, and both 9 h and 24 h of high and middle doses in nitrofurazone have comparatively high and similar expression patterns (a boxplot for the expression patterns of gene Mgst2 from all CCs also shown in Fig. 4).

Table 3. Biomarker genes for all the CCs in glutathione metabolism pathway.

CCs	Dose	Time	Score	Mgst2	Gclc	G6pd	Gsr	Srm
Nitrofurazone	High	24h	5.70	13.74	13.05	12.50	11.81	11.24
Acetaminophen	High	24h	4.73	13.51	13.36	12.47	11.16	10.9
Acetaminophen	Middle	24h	4.23	13.64	13.07	12.04	11.06	10.6
Nitrofurazone	High	9h	3.86	--	13.04	11.54	--	10.91
Nitrofurazone	Middle	9h	3.38	--	13.16	--	--	--
Methapyrilene	High	9h	2.56	13.05	13.38	11.36	10.27	10.81
Methapyrilene	High	24h	2.31	13.21	--	11.76	10.26	10.21
Nitrofurazone	Middle	24h	2.05	13.36	--	--	--	--
Acetaminophen	Middle	9h	1.82	--	12.82	--	--	--
Methapyrilene	High	6h	1.70	--	13.10	10.81	10.16	10.71

-- = not significant (comparatively low expression value than the source factor). Score represent the PC1 score and the value corresponding genes are the mean expression value.

After matching both results from PCA and post-hoc test we found, only high dose of methapyrilene and middle and high doses of acetaminophen and nitrofurazone, which are considered as the responsible source factor for glutathione depletion. The pathological result of these three CCs permits the result obtained from proposed algorithm (http:/toxico.nibio.go.jp/). We can also rank CCs with their dose levels by PC1 score which may reflect the effectiveness of glutathione depletion. The rank score and significant mean expression of biomarker genes are presented in Table 3.

Conclusion

Biomarker discovery in toxicogenomics research is crucial for assessing adverse effect of chemical in the early stage. Large scale chemical-induced GE data help to study adverse effect of chemicals. Basically, GE changes imitate the pathological condition changes. Therefore, to measure toxic potential of unknown drug by measuring GE fingerprint with some known drug give additional benefits which save both cost and time. From the literature it is known that the compounds, acetaminophen, methapyrilene and nitrofurazone are *glutathione* depleting compound and proposed method identify corresponding dose levels which caused glutathione depletion. Proposed algorithm also identifies five BGs (Mgst2, Gclc, G6pd, Gsr and Srm) that's effectively control glutathione depletion. In mechanistic point of view by the proposed method, nitrofurazone shows slightly different mechanism then the others CCs. It would be possible to identify new chemical toxicity (for instance, *glutathione* depletion) by comparing its expression pattern of BGs with the known CCs. Therefore, further study is obviously needed to rally and simplify the candidacy of BGs suggested in the work.

References

Battershill JM (2005). Toxicogenomics: regulatory perspective on current position. Human and Experimental Toxicology 24: 35-40.

Cui X and Churchill GA (2003). Statistical tests for differential expression in cDNA microarray experiments. Genome Biology 4(4): 210.

Davis AM and Riley RJ (2004). Predictive ADMET studies, the challenges and the opportunities. Current Opinion Chemical Biology 8: 378-386.

Ekins S, Nikolsky Y and Nikolskaya T (2005). Techniques: Application of systems biology to absorption, distribution, metabolism, excretion and toxicity. Trends Pharmacology Science 26: 202-209.

Fielden MR, Brenna R and Gollub J (2007). A GE biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. Toxicological Sciences 99: 90-100.

Heinloth AN, Irwin RD, Boorman GA, Nettesheim P, Fannin RD, Sieber SO, Snell ML, Tucker CJ, Li L and Travlos GS (2004). GE profiling of rat livers reveals indicators of potential adverse effects. Toxicological Sciences 80: 193-202.

Hirode M, Ono A, Miyagishima T, Nagao T, Ohno Y and Urushidani T (2008). Gene expression profiling in rat liver treated with compounds inducing phospholipidosis. Toxicology and Applied Pharmacology 229: 290-299.

Igarashi Y, Nakatsu N, Yamashita T, Ohno Y, Ono A, Urushidani T and Yamada H (2015). Open TG-GATEs: a large-scale toxicogenomics. Nucleic Acids Research 43: 921-927.

Irwin RD, Boorman GA, Cunningham ML, Heinloth AN, Malarkey DE and Paules RS (2004). Application of toxicogenomics to oxicology: Basic concepts in the analysis of microarray data. Toxicologic Pathology 32: 72-83.

Kiyosawa N, Ando Y, Manabe S, Yamoto T (2009). Toxicogenomic Biomarkers for Liver Toxicity. Journal of Toxicologic Pathology 22(1): 35-52.

Klaassen C (2007). Casareet and Doull's toxicology: The basic science of poisons. McGraw Hill Professional, New Work.

Newton MA, Kendziorski CM, Richmond CS, Blattner FR and Tsui KW (2001). On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. Journal of Computational Biology 8: 37-52.

Nuwaysir EF, Bittner M, Trent J, Barrett JC and Afshari CA (1999). Microarrays and toxicology: The advent of toxicogenomics. Molecular Carcinogenesis 24: 153-159.

Nyström-Persson J, Igarashi Y, Ito M, Morita M, Nakatsu N, Yamada H and Mizuguchi K (2013). Toxygates: interactive toxicity analysis on a hybrid microarray and linked data platform. Bioinformatics 23: 3080-3086.

Pavlidis P (2003). Using ANOVA for gene selection from microarray studies of nervous system. Methods 31: 282-289.

Rocke DM and Durbin B (2001). A model for measurement error for gene expression arrays. Journal of Computational Biology 8: 557-569.

Sahai H and Agell MI (2000). The Analysis of Variance, *Birkhauser*, Boston.

Searfoss GH, Ryan TP and Jolly RA (2005). The role of transcriptome analysis in pre-clinical toxicology. Current Molecular Medicine 5: 53-64.

Suter L, Babiss LE and Wheeldon EB (2004). Toxicogenomics in predictive toxicology in drug development. Chemistry and Biology 11: 161-171.

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

Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y and Urushidani T (2010). The Japanese toxicogenomics project: Application of toxicog nomics. Molecular Nutrition Food Research 54: 218-227.

Ulrich R and Friend SH (2002). Toxicogenomics and drug discovery: Will new technologies help us produce better drugs? Nature Reviews Drug Discovery 1: 84-88.

Waters MD and Fostel JM (2004). Toxicogenomics and systems toxicology: Aims and prospects. Nature Reviews Genetics 5: 936-948.