【原　著】

Global gene expression analysis in the liver of mice administered carcinogens and non-carcinogens

—For the development of a new rapid and reliable carcinogenicity test—

Satoshi KOKURA[*1, *2], Tomohisa TAKAGI[*2], Katsura MIZUSHIMA[*3]
[*1] Health and Medical Sciences, Kyoto Gakuen University
[*2] Biomedical Safety Science, Kyoto Prefectural University of Medicine
[*3] Medical Proteomics, Kyoto Prefectural University of Medicine

発がん性物質および非発がん性物質をマウスに投与した場合のマウス肝臓での網羅的遺伝子発現解析
—新規迅速発がん性テストの開発を目指して—

古倉　聡[*1, *2]，高木　智久[*2]，水島　かつら[*3]
＊1 京都学園大学　健康医療学部　看護学科
＊2 京都府立医科大学　医学部　生体安全医学講座
＊3 京都府立医科大学　医学部　プロテオミクス医学講座

Abstract

Although much has been elucidated about the mechanism of cellular carcinogenesis, evaluation of the carcinogenicity of substances is still difficult. Since a number of new synthetic materials for drugs and food have been developed, it is necessary to sufficiently evaluate their carcinogenicities. In carcinogenicity tests performed at present, the test substance is administered to rodents for 2 years (their life span) and its carcinogenicity in each organ is evaluated by macroscopic and pathological examinations. Such carcinogenicity tests require considerable time and resources. In this study, we attempted to develop a new carcinogenicity test by global gene expression analysis, which is independent of target organ-specific activity of carcinogens. All food and chemicals are primarily transported to the liver via the portal vein, both of which are independent of the routes of intake. Therefore, we selected the liver as the best organ with which to assess carcinogenicity in gene expression analysis. The present report indicates the possibility that global gene analysis by using DNA microarray in the liver 3 hours after the administration of chemicals could establish a rapid bioassay to identify carcinogens for various organs.

和文抄録

「目的」細胞のがん化のメカニズムは，かなり具体的に明らかとされつつあるが，ある物質に発がん性が有るかどうかの判定は容易ではない．その一方で，今日，医薬品・食品分野では，新規合成物質が盛んに開発されており，これらの発がん性については，十分に検討する必要がある．今日施

行されているがん原性試験は，齧歯類にその生涯期間に相当する2年間，試験する化学物質を投与し発がん性の有無を検討する安全性試験が基本であるが，この試験には膨大な時間・資源が必要である．そこで，我々は，DNA microarrayによる肝臓での網羅的遺伝子発現解析を行なうことにより，被検物質の発がん性を評価する方法の開発を試みている．

「方法」既知の発がん物質と非発がん物質を7週齢雄性BALB/cマウスに投与し，その3時間後および24時間後に，マウスの肝臓を摘出しmRNAを抽出し，Mouse Genome 430-2.0 Arrayにより遺伝子発現解析を行なった．得られたデータは，1）階層的クラスター分析，2）Weighted Voting（重み付け投票）algorithmにより，発がん性のリスクを予測した．

「成績」1）化学物質投与3時間後と24時間後で発現変動する遺伝子発現パターンは一致しない．2）化学物質投与後3時間，24時間で発がん物質のみにみられる共通した動きを示す遺伝子群の絞り込みが達成された．3）Weighted Voting algorithmによる検討では，遺伝子発現プロファイルのなかで，発現が低下した遺伝子よりも，発現が亢進した遺伝子に注目した方が，発がん性予測精度が高い可能性が示唆された．

「結論」既知の発がん物質と非発がん物質をマウスに投与した後の肝臓での遺伝子発現を網羅的に解析したところ，発がん物質と非発がん物質の遺伝子発現プロファイルのパターンに違いを認めた．本検討より新規発がん性予測法の開発の可能性が期待できる．

キーワード：発がん性、DNAマイクロアレイ、網羅的遺伝子発現解析、予測、発がん物質
Key words: carcinogenicity, DNA microarray, global gene expression analysis, prediction, carcinogens

Ⅰ Introduction

In recent years, the incidence of cancer is increasing not only in elderly people but also in relatively young people. There are various causes of cancer, but a major cause may be the influences of daily foods [1]. In this situation, examinations of food additives and newly developed medicine for the presence/absence of carcinogenicity are very important. The most reliable method for the prediction of carcinogenic potential is a long-term *in vivo* test, which extends for over 2 years in two rodent species, rats and mice, requires a large number of experimental animals, and demands both a large space for animal testing and enormous cost [2, 3]. Moreover, gross observations of animals are essential for target organ-specific carcinogenic activity. When we investigate the carcinogenic potential of test samples, we cannot predict the possibility of target organ-specific activity of a sample in advance. Therefore, a conventional long-term carcinogenicity test needs gross observations of animals.

Recent advances in DNA microarray technology have allowed examination of a large number of gene expression profiles in a short time. Some studies have evaluated the carcinogenicity of test substances by DNA microarrays [4–6]. In this method, the procedure is simple because the amount of gene expression is used as a parameter and many samples can be simultaneously processed. However, this method is applicable only to the organ that was analyzed by DNA microarray and not to other organs. Moreover, it is impossible to perform global gene expression analysis of all organs in mice.

In this study, we aimed to provide a basis for a rapid and reliable carcinogenicity test that is independent of target organ-specific activity of carcinogens. All food and chemicals are primarily transported to the liver via the portal vein, both of which are independent of the routes of intake. Therefore, we selected the liver as the best organ with which to assess carcinogenicity in global gene expression analysis by DNA microarray.

Ⅱ Materials and Methods

1. Reagents

All chemicals were prepared immediately before use. An RNA assay Mini kit was purchased from Qiagen (Valencia, CA) and a Mouse Genome430-2 array and Eukaryotic Small Sample

Target Labeling Assay kit were purchased from Affymetrix (Santa Clara, CA). All other chemicals used were of reagent grade.

2. Test chemicals

We selected 12 substances (Table 1) that have been well characterized for carcinogenicity, including 8 chemicals demonstrating carcinogenicity in multiple organs. The remaining 4 chemicals were non-carcinogens. A diverse range of carcinogens was selected to cover the widest possible range of target organs of carcinogens.

3. Experimental design

The experimental design is quite simple. We used 7-week-old male Balb/c mice. Three hours and 24 hours after the administration of test samples, the mice were killed and each liver was extracted and frozen immediately in liquid nitrogen for DNA microarray analysis. We used three animals at each time point for one test sample. As a control, three untreated mice were killed and their livers were extracted at both time points. The Animal Care Committee of the Kyoto Prefectural University of Medicine (Kyoto, Japan) approved all of the experimental procedures described below.

4. Microarray data analysis

Array data analysis was carried out using Affymetrix GeneChip operating software (GCOS) Ave.500. This software analyzes image data and computes an intensity value for each probe cell. Briefly, mismatch probes act as specificity controls that allow the direct subtraction of both background and cross-hybridization signals. To determine the quantitative RNA abundance, the average difference values (i.e., gene expression levels) representing the perfect match-mismatch for each gene-specific probe family were calculated and the fold changes in average difference values were determined according to the Affymetrix algorithms.

We evaluated the gene expression profile by using the Affimetrix Genechip, Mouse Genome 430-2 array (Affimetrix, Santa Clara, CA), containing 45,101 gene probes. A diagrammatic representation of the analysis is shown in Figure 1. Therefore, genes that were either up-regulated ≥2-fold or down-regulated ≤0.5-fold relative to untreated mice were first selected. This gene selection was conducted for both the carcinogen and non-carcinogen groups.

III　Results and Discussion

Table 2 shows the number of genes that were twofold up-regulated or 0.5 -fold down-regulated by carcinogens and non-carcinogens as compared with that of untreated mice 3 hours (Table 2a, b) and 24 hours (Table 2c, d) after the administration of samples, respectively.

Subsequently, among the genes that were either up-regulated or down-regulated in the groups administered carcinogens, genes for which expression levels changed significantly relative to groups administered non-carcinogens were selected. Specifically, "multiple genes for which expression levels

Table 1. Carcinogen category of chemicals used in this study.

No.	chemical	dose	administration
CA1	4-Nitroquinoline-1-oxide (4NQO)	15 mg/kg	sc
CA2	Cyclophosphamide (Cyc)	30 mg/kg	gavage
CA3	N.N-Diethylnitrosoamine (DEN)	90 mg/kg	ip
CA4	Ethyl acrylate (Ethyl)	200 mg/kg	gavage
CA5	N-methyl-N-nitroures (MNU)	75 mg/kg	ip
CA6	Procarbazine (Pro)	12 mg/kg	ip
CA7	Thiotepa (Thi)	2 mg/kg	ip
CA8	Vinyl carbamate (UR)	500 mg/kg	ip
NCA1	α–tocopherol (Toc)	10 mg/kg	gavage
NCA2	Rebamipide (Reba)	30 mg/kg	gavage
NCA3	Teprenone (Tep)	200 mg/kg	gavage
NCA4	Xylene (Xyl)	1000 mg/kg	gavage

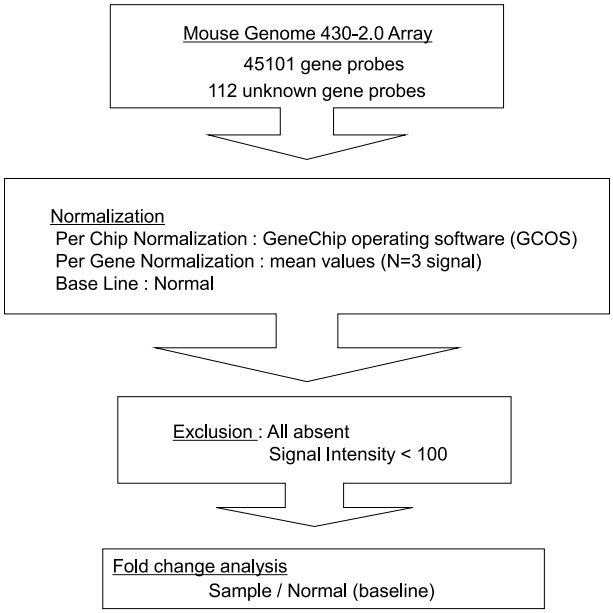

Figure 1. Data analysis diagram. Microarray expression data from 45101 gene probes were used for GCOS analysis. Main categories were carcinogen, non-carcinogen, and control. Three analyses, carcinogen versus control, non-carcinogen versus control, and carcinogen versus non-carcinogen, were carried out in this study.

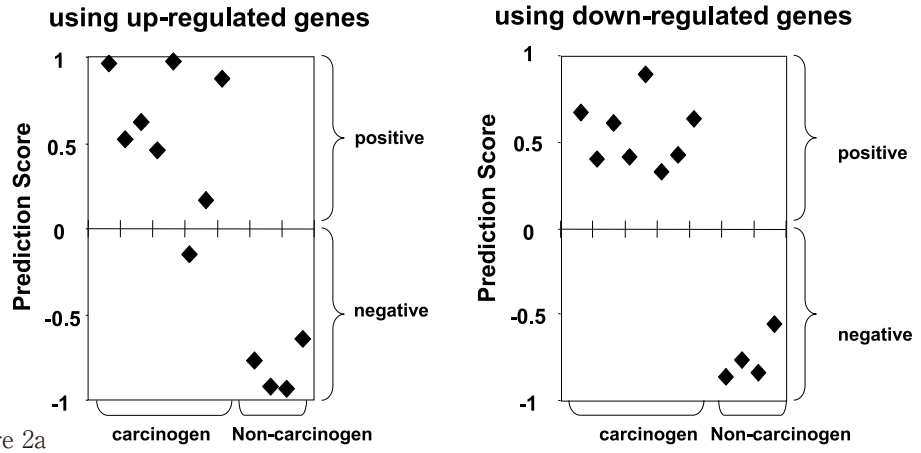

Figure 2a

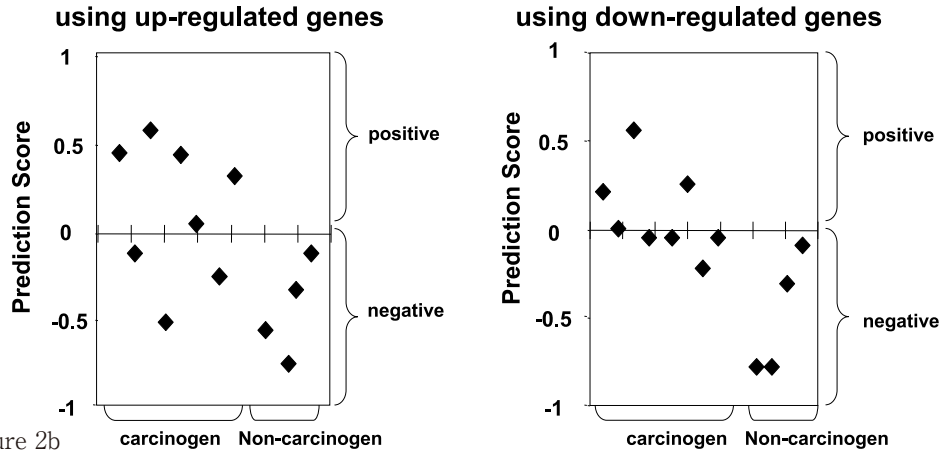

Figure 2b

Figure 2. Prediction of carcinogenicity by weighted voting algorithm.
a) Prediction score by using gene expression profile 3h after administration of test samples. The scatter plots show the prediction scores for 12 samples in cross-validation.
b) Prediction score by using gene expression profile 24h after administration of test samples. The scatter plots show the prediction scores for 12 samples in cross-validation.

Table 2. The number of genes that were twofold up-regulated or 0.5-fold down-regulated by carcinogens and non-carcinogens as compared with that of untreated mice 3 hours (Table 2a, b) and 24 hours (Table 2c, d) after the administration of samples.

Table 2a

Group	Carcinogens		Number of total probes	selected genes	two fold up-regulated genes	0.5 fold down-regulated genes
1	CA1	4NQO	45101	19763	1132	905
2	CA2	Cyc	45101	19777	967	434
3	CA3	DEN	45101	19584	907	1146
4	CA4	Ethyl	45101	19467	657	822
5	CA5	MNU	45101	19389	728	951
6	CA6	Pro	45101	19647	632	575
7	CA7	Thi	45101	19430	696	373
8	CA8	UR	45101	19740	941	874
			Common genes		46	12
			All genes		3614	3450

Table 2b

Group	Non-carcinogens		Number of total probes	selected genes	two fold up-regulated genes	0.5 fold down-regulated genes
9	NCA1	Toc	45101	19690	905	850
10	NCA2	Reba	45101	19830	969	771
11	NCA3	Tep	45101	20159	1220	808
12	NCA4	Xyl	45101	19725	741	570
			Common genes		87	39
			All genes		2668	2040

Table 2c

Group	Carcinogens		Number of total probes	selected genes	two fold up-regulated genes	0.5 fold down-regulated genes
1	CA1	4NQO	45101	19176	387	618
2	CA2	Cyc	45101	19422	508	418
3	CA3	DEN	45101	19354	648	913
4	CA4	Ethyl	45101	20046	1334	1786
5	CA5	MNU	45101	19230	298	638
6	CA6	Pro	45101	19854	1021	1456
7	CA7	Thi	45101	19374	490	367
8	CA8	UR	45101	19249	396	724
			Common genes		12	6
			All genes		3200	3782

Table 2d

Group	Non-carcinogens		Number of total probes	selected genes	two fold up-regulated genes	0.5 fold down-regulated genes
9	NCA1	Toc	45101	19539	576	705
10	NCA2	Reba	45101	19281	414	784
11	NCA3	Tep	45101	19394	652	589
12	NCA4	Xyl	45101	19360	513	751
			Common genes		51	60
			All genes		1547	1910

in the liver changed in response to administration of carcinogens with carcinogenicity in organs other than the liver in mice" were selected. As a result, the following genes were selected: 1247 genes that were significantly up-regulated after 3 hours (gene group A, up-regulated genes), 1388 genes that were significantly down-regulated after 3 hours (gene group B, down-regulated genes), 1319 genes that were significantly up-regulated after 24 hours (gene group C, up-regulated genes), and 1345 genes that were significantly down-regulated after 24 hours (gene group D, down-regulated genes). Based on these expression data, we adopted a weighted-voting (WV) algorithm generally used in gene expression profiling [7–9]. Using the previously selected gene groups A, B, C, and D, we developed an equation for predicting carcinogenicity and conducted leave-one-out cross-validation (LOO). Specifically, for a group administered test substance X, weighted votes were calculated for each selected gene in the following manner. The weighted vote Vxa for gene a in a group administered test substance X was calculated using the following equation (I).

$$Vxa = Sa \{Xa-(Mca+Mnca)/2\} \quad (I)$$

In the above equation, Xa is the expression level (expression intensity) of gene a in a group administered test substance X, Mca is the mean expression level of gene a in groups administered known carcinogens, and Mnca is the mean expression level of gene a in groups administered known non-carcinogens.

The coefficient Sa in equation (I), which assigns the weight of expression of gene a, is calculated using the following equation (II).

$$Sa=(Mca-Mnca)/(SDca+SDnca) \quad (II)$$

In the above equation, SDca and SDnca are standard deviations of the expression level of gene in groups administered known carcinogens and in groups administered known non-carcinogens, respectively.

Weighted vote Vxi for each gene i in a group administered test substance X was calculated, as was the sum of weighted votes ΣVi. Test substance X can be considered likely to be a carcinogen if $\Sigma Vi \geq 1$ and unlikely to be a carcinogen if $\Sigma Vi < 1$ (primary assessment).

Prediction score (PS) for test substance X was calculated using the following equation (III).

$$PS=\{VXc-|VXnc|\}/\{VXc+|VXnc|\} \quad (III)$$

In the above equation, VXc is the sum of positive V values (positive V) among V values calculated for test substance X, VXnc is the sum of negative V values (negative V) among V values calculated for test substance X, and |VXnc| is the absolute value of VXnc.

The PS for test substance X was calculated using equation (III). Test substance X can be considered likely to be a carcinogen if $0<PS\leq 1$ and unlikely to be a carcinogen if $-1<PS<0$ (secondary assessment).

As described above, in gene groups C and D, ΣVi and PS were positive only for certain carcinogens and negative for all non-carcinogens, indicating that these variables can be adequately used for primary screening.

In addition, comparison of results for gene groups A and B (3 hours after administration) and gene groups C and D (24 hours after administration) showed that following administration of test substances to mice, more accurate predictions can be made when the liver is extracted 3 hours after administration than when the liver is extracted 24 hours after administration.

The carcinogenicity test developed in the present study evaluates the carcinogenicity of test substances based on gene expression analysis, and thus has advantages such as the following: 1) evaluation does not require long-term administration of test substances to nonhuman animals and enables fast and simple assessment of carcinogenicity, 2) evaluation involves gene expression analysis, and thus enables detection of early stages of carcinogenicity that cannot be detected macroscopically or pathologically, and 3) evaluation involves analysis of expression levels of multiple genes that undergo changes in expression level (up-regulation or down-regulation) in the liver in response to the administration to nonhuman animals of carcinogens that have carcinogenicity in organs other than the liver, and thus enables simultaneous evaluation of carcinogenicity for a wide range of organs by simply analyzing expression levels of applicable genes in the liver.

Recently, methods that showed the discrimination or prediction of carcinogenic potential by a toxicogenomic approach using human or rat hepatoma

cell lines were reported [10, 11]. These methods are reliable and relatively simple because they are conducted *in vitro* and may thus be used as a screening method in the future. Meanwhile, our evaluation method, which involves the administration of test substances to animals, is characterized by assessment that includes absorption and metabolic processes, particularly when test substances are orally administered, and thus enables evaluation of factors that cannot be assessed *in vitro*.

To confirm the reliability of this method for the prediction of carcinogenicity, we are in the middle of investigating another set of chemicals by this method. In conclusion, we performed global gene analysis by using DNA microarray in the liver at 3 and 24 hours after the administration of chemicals and indicated the possibility that our new carcinogenicity test using the weighted voting algorithm could establish a rapid bioassay to identify carcinogens for various organs.

References

1) Doll R.: Prospects for prevention. Br Med J (Clin Res Ed), **286**: 445–453, 1983
2) Huff J. : Long-term chemical carcinogenesis bioassays predict human cancer hazards. Ann N Y Acad Sci., **895**: 56-79, 1999
3) Schramm T, Teichmann B.: Chemical carcinogens: screening, testing, risk assessment for man. Neoplasma., **28**: 129-131, 1981
4) Nakayama K, Kawano Y, Kawakami Y, et al.: Differences in gene expression profiles in the liver between carcinogenic and non-carcinogenic isomers of compounds given to rats in a 28-day repeat-dose toxicity study. Toxicol Appl Pharmacol. **217**: 299-307, 2006
5) Delker D, Hatch G, Allen J, et al. : Molecular biomarkers of oxidative stress associated with bromate carcinogenicity. Toxicology. **221**: 158-165, 2006
6) Hester SD, Barry WT, Zou F, et al.: Transcriptomic analysis of F344 rat nasal epithelium suggests that the lack of carcinogenic response to glutaraldehyde is due to its greater toxicity compared to formaldehyde. Toxicol Pathol. **33**: 415-424, 2005
7) Golub TR, Slonim DK, Tamayo P, et al.: Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science, **286**:531–537, 1999
8) Slonim D, Tamayo P, Mesirov J, et al. : Class Prediction and Discovery Using Gene Expression Data. Proceedings of the Fourth Annual Conference on Computational Molecular Biology (RECOMB), pp : 263–272, 2000
9) Pomeroy SL, Tamayo P, Gaasenbeek M, et al. : Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature. **415**: 436-442, 2002
10) van Delft JH, van Agen E, van Breda SG, et al. : Comparison of supervised clustering methods to discriminate genotoxic from non-genotoxic carcinogens by gene expression profiling. Mutat Res. **575** :17-33, 2005
11) Tsujimura K, Asamoto M, Suzuki S, et al. : Prediction of carcinogenic potential by a toxicogenomic approach using rat hepatoma cells. Cancer Sci., **97**: 1002-1010, 2006