代替法学会の前身となる研究会が設立された．これら
の研究を通じて多くの方法が開発されてきた．一方，
EU議会は動物愛護運動と動物実験に対する反対運動
の高まりに対応して，代替法開発の拠点とし，代替法
についてのデータベースを設置・維持するため，また，
行政，産業，生物・医学分野の科学者，消費者，およ
び動物愛護運動グループの対話を促進することを目的
に1991年に代替法バリデーションセンター（European
Center for the Validation of Alternative Methods:
ECVAM）を設立した（1994年開所）．国レベルでも代
替法研究のための機関として，ドイツはThe Centre
for Documentation and Evaluation of Alternatives to
Animal Experiments（ZEBET）を1989年に，オランダ
はNetherlands Centre Alternatives to Animal Use
（NCA，最近National Knowledge Centre on Alternatives
（NKCA）に名称変更）を，イギリスはThe National
Centre for the Replacement, Refinement and Reduction
of Animals in Research（NC3Rs）を2004年に設立した．
米国は，毒性試験法の開発，バリデーション，受入れ，
および国内および国際レベルでのハーモナイゼーショ
ンに関する問題を連邦政府内で調整するために
NICEATM（NTP Interagency Center for the Evaluation
of Alternative Toxicological Methods）の下にNIEHS
（National Institute of Environmental Health Sciences）
を含む14の行政機関および研究機関によりICCVAM
（Interagency Coordinating Committee on the Validation
of Alternative Methods）を1993年に設置した．この
ように，欧米の行政機関においては，新たに開発され
た動物実験代替法を科学的に評価し，可能なものにつ
いては取り入れていこうと，努力している．
　わが国においても，動物福祉への関心の高まりに応
じて，平成17年6月に交付された動物の愛護及び管
理に関する法律の一部を改正する法律（改正動愛法，
平成17年法律第68号）により，動物実験等に関する
3Rの原則を尊重することが義務付けられるとともに，
翌年，動物実験に関する基本指針が文部科学省(5)，厚
生労働省(6)，環境省(7)，および農林水産省(8)から
通知された．また，日本学術会議より動物実験に関す
る詳細指針が通知された．2008年には，主に動物実験
代替法の評価を行う施設として，国立医薬品食品衛生
研究所 安全性生物試験研究センター 薬理部内に新規
試験法評価室（通称JaCVAM）が設立され，活動を始
めた．その後，2010年には，その活動が認められ，
ECVAM，ICCVAMおよびカナダ厚生省（Health
Canada）とともに代替法に関する協力を目的に
International Cooperation on Alternative Test Methods

（ICATM）協定が結ばれ，代替法の評価が進んでいる．
2011年にはこれに韓国FDAに属するKorean Center
for the Validation of Alternative Methods（KoCVAM）
も加わった．

2. 薬理学分野での動物実験代替法への対応について

　薬理学の研究においては，医薬品や化学物質等の生
体に及ぼす影響を検討するため，摘出臓器や組織，細
胞，細胞下分画などを用いたインビトロ試験系が多用
されてきたが，得られた結果が真に意義のあるもので
あることを証明するためには，インビボ実験が欠かせ
ない．実際，表1に示したように，平成23年度に予定
した第84回薬理学会における口頭発表の内，マウス
やラット，ウサギ，およびモルモットなど哺乳動物を
用いる研究は，139件で全体の約52%を占めていた
（なお，第77回年会ではポスター発表のうち約70%が
動物を用いた研究であった．）．このように，依然とし
て，薬理学の研究において動物実験は欠かせないもの
である．
　一方，動物実験に対する一般市民の反応は，その意
義と説明の仕方により異なってくる．たとえば，イギ
リスで行われた調査では，白血病のような重篤な疾病
の治療につながる研究では，80%以上の人がマウスに
よる実験を許容していたが，化粧品のための研究では，
40%弱の人しか動物実験を許容しなかった．また，単
に「科学者は動物実験を許容されるべきである」との
質問では，64%の人が中程度以上の反対を示したが，
「痛みを和らげる薬物の開発や生命を脅かす疾患の治
療法を開発するための研究をしている科学者がいる」
ことを説明すると，その反対が41%に減少した（New
Scientist. 1999 May 22, p. 26-31）．
　すなわち，動物実験を行う薬理学者は自己の研究目
的に動物実験がやむを得ざるものであること，また，
行う動物実験は動物実験倫理委員会により承認された
ものであることを示す必要がある．また，動物実験を
行っている施設は，指針に則り，当該施設が適切な施
設で適切な動物実験を行っていることが第三者評価に
より認証されることが望ましい．
　このようにして，動物実験に関する社会的要請に応
えることが，継続的に薬理学研究を進める上で必須で
ある．このことを広く会員に示し，適正な動物実験を
進めるため，薬理学会では，改正動愛法の趣旨を尊重
し，動物実験に関する日本薬理学会指針を改正し，「日
本薬理学会は本指針に従った動物実験が行われること
を期待するとともに，これに反する研究の成果は本会

表1 薬理学会での動物実験の調査（口頭発表のみ）

		初日	2〜3日	計	合計	%
動物を使用した実験	in vivo 実験*					
	マウス	39	51	90		
	GM マウス#	7	16	23		
	ラット	5	17	22		
	ウサギ	1	1	2		
	モルモット	1	1	2	139	51.87
	カエル	1	2	3		
	ゼブラフィッシュ	1	2	3	6	2.24
	不明##	6	8	14	14	5.22
	摘出組織・細胞実験**					
	マウス	7	5	12		
	ラット	2	12	14	26	9.70
	ヒト	1	8	9	9	3.36
	カエル	1	0	1		
	ゼブラフィッシュ	1	0	1	2	0.75
動物未使用***		10	38	48	48	17.91
不明		7	17	24	24	8.96
合計		90	178	268		

*：in vivo で動物に何らかの処置をしているもの．**：無処置の動物から摘出した組織や細胞等を利用しているもの．
***：細胞株，微生物，線虫および屠殺場に由来する試料を用いているもの．#：GM と wild の両方を用いたものは重複して計算．##：動物を使用しているが，動物種が不明のもの．

の刊行する学術雑誌から排除する所存である」と宣言した（平成19年）(9)．薬理学研究者はこの指針に従い，研究を進めるよう留意すべきである．しかし，表2に示したように，改正動愛法が施行されて1年半が経過した平成19年2〜3月の時点でも，代替法に関する教育は必ずしも十分では無かった．現在は，改善されていることを望むものである．

医薬品開発のための国際的な指針「医薬品の臨床試験及び販売承認申請のための非臨床安全性試験の実施についてのガイダンス（ICH M3 (R2)）」(10) においても，3R の原則への配慮が多々なされており，本指針が「3R（使用動物数の削減／苦痛の軽減／代替法の利用）の原則に従って動物の使用を抑え，医薬品開発のための資源の有効利用に資する」よう期待している．「安全性評価のための新しいインビトロ代替法の利用について考慮すべきである．」とも述べている．これは今後開発されることが予想される iPS 細胞由来標本を用いた試験を配慮して記載されたものであるが，ICH で受け入れるためには，それらのバリデーションが完了し，全ての ICH 規制当局によって認められる必要があるとされた．

表2 動物実験倫理に関する薬学系大学の教育内容

教育内容*	教えている大学の割合
動愛法	80.80%
実験動物の生理，生態，習性	65.40%
実験動物の飼育，管理方法	84.60%
動物実験における 3R の原則	69.20%
動物の苦痛の評価	69.20%
苦痛の軽減方法	73.10%
安楽死の方法	80.80%
動物実験代替法	57.70%
その他	42.30%

*：文部科学省等の指針から教えることが望ましいと思われるもの．平成19年2〜3月にアンケート調査した結果の集計（回答数29，回答施設数24）．

効力を裏づけるためのインビボおよび／またはインビトロにおける薬力学的試験は，目的とする治療標的に対する被験物質の作用機序や効果を調べることを意図している．このような試験は，通常，医薬品開発の探索段階で実施され，通常は GLP に従っては，行われない．しかし，ICH M3 (R2) 指針では，これらの試験結果を非臨床試験および臨床試験の用量設定に役立てることを勧めている．逆に，有効に利用できるよう

にGLPに準じて明確な試験計画をたて，データの記録や保管に心がけるべきである．また，インビボで評価する場合には，使用動物を削減するため，可能な範囲内で，一般毒性試験や他の試験に組み込んで実施することを考慮すべきとした．乱用性試験についても，ヒト以外の霊長類を使用するケースは，霊長類がヒトでの乱用性を予測し得ると考える明確な根拠があり，しかも，げっ歯類のモデルは不適切である場合に限るべきであるとした．なお，実験の際の動物の苦痛の評価については，国立大学法人動物実験施設協議会の作成した「動物実験処置の苦痛分類に関する解説」が参考となる．

おわりに

動物実験を行う者はそれが本当に必要な実験であるか，動物を用いない他の試験法に置き換えることができないか，動物に不必要な苦痛を与えていないかを，施設毎の動物実験委員会で審議することが求められている．薬理学を含む生物科学を教える立場にある者は，動物実験代替法について適正な教育を行う必要がある．また，レギュラトリーサイエンスに携わる者には，ガイドライン等で定められた毒性試験法のように，行政的に要求される試験の妥当性について過去の経験，新しい技術・知見，現時点での社会情勢などに基づき常に見直さなくてはならない．EUで一時大きな問題となった動物の権利を標榜する団体による研究機関の攻撃も，最近は鎮静化している模様であるが，これらは行政機関や大学や企業などの研究機関，および動物実験実施者等による動物福祉への配慮や違法行為に対する毅然たる態度による成果であると考えている．

文　　献

1) Russell WMS, et al. The principles of human experimental technique. Methuen; 1959.
2) 小野宏，大野忠夫編著．動物実験代替法マニュアル．共立出版; 1994. p. 123-133.
3) 大野泰雄．国立衛研報．2004;122:1-10.
4) OECD: Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment, OECD series on testing and assessment 2005; No 34 (18 August 2005).
5) 文部科学省．研究機関等における動物実験等の実施に関する基本指針．文部科学省告示 第71号, 2006.6.1.
6) 厚生労働省．厚生労働省の所管する動物実験等の実施に関する基本指針．厚生労働省通知 科発 0601002号, 2006.6.1.
7) 環境省．実験動物の飼育及び保管並びに苦痛の軽減に関する基準．環境省告示第88号, 2006.6.1.
8) 農林水産省．農林水産省の所管する研究機関等における動物実験等の実施に関する基本指針. 2006.6.1.
8) 日本学術会議．動物実験の適正な実施に向けたガイドライン. 2006.6.1.
9) 大野泰雄．日薬理誌．2007;129:5-9.
10) 厚生労働省．医薬品の臨床試験及び販売承認申請のための非臨床安全性試験の実施についてのガイダンス ICH Consensus Guideline. 薬食審査発 0219 第4号, 2010.2.19.

著者プロフィール

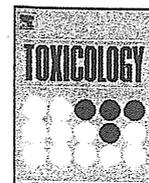**大野　泰雄（おおの　やすお）**

国立医薬品食品衛生研究所所長．
◇1976年東京大学薬学系大学院博士課程修了．同年現在の研究所に採用され，'11年より現職．肝・腎における代謝と毒性研究に従事．カロリンスカ研究所に留学後 in vitro 毒性および動物実験代替法研究に関与．日本動物実験代替法学会長および第6回国際動物実験代替法学会長を歴任．

Identification of a novel set of biomarkers for evaluating phospholipidosis-inducing potential of compounds using rat liver microarray data measured 24-h after single dose administration

Henrik T. Yudate[a,1], Toshihiro Kai[a,*,1], Mikio Aoki[a], Yohsuke Minowa[b], Toru Yamada[a], Toru Kimura[a], Atsushi Ono[c], Hiroshi Yamada[b], Yasuo Ohno[d], Tetsuro Urushidani[b,e]

[a] Genomic Science Laboratories, Dainippon Sumitomo Pharma Co., Ltd., Japan
[b] Toxicogenomics Informatics Project, National Institute of Biomedical Innovation, Japan
[c] Division of Risk Assessment, National Institute of Health Sciences, Japan
[d] National Institute of Health Sciences, Japan
[e] Department of Pathophysiology, Faculty of Pharmaceutical Sciences, Doshisha Women's College of Liberal Arts, Japan

ARTICLE INFO

ABSTRACT

Phospholipid accumulation manifests as an adverse effect of cationic amphiphilic drugs in particular. Detection, however, by histopathology examination is time-consuming and may require repeated administration of compounds for several weeks. To eliminate compounds with potential for inducing phospholipidosis from the discovery pipeline, we have identified and validated a set of biomarkers for predicting the phospholipidosis-inducing potential utilizing a comprehensive rat transcriptome microarray database created by the Japanese Toxicogenomics and Toxicogenomics Informatics Projects (TGP/TGP2) together with in-house data. The set of biomarkers comprising 25 Affymetrix GeneChip probe sets was identified using genetic algorithm optimization on 24-h time-point microarray data from rats treated with single doses of hepatotoxic compounds including amiodarone, clomipramine, haloperidol, hydroxyzine, imipramine, and perhexiline. The set of novel biomarkers represents an early time-point gene-expression pattern characteristic for a condition eventually leading to phospholipidosis. This implies significant advantages in terms of time and resources over currently published biomarkers derived using repeated-dosing late time-point data. The biomarker set was validated by 11 independent compounds. Accuracy, sensitivity, and specificity values were 82%, 67%, and 100%, respectively and the area under the receiver operating characteristic curve was 0.97. These results show that the biomarker set possesses a high classification accuracy for novel compounds. Pathway analysis was carried out for the biomarkers and the detection of pathways related to lipid-metabolism was statistically significant. These pathways most probably reflect lipid metabolism changes associated with phospholipidosis supporting the validity of our novel biomarkers.

© 2012 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Phospholipidosis is an excessive accumulation of phospholipids in lysosomes and may occur in multiple organs and tissues of the body (Anderson and Borlak, 2006; Halliwell, 1997; Hruban, 1984; Reasor and Kacew, 2001). It is characterized by vacuolated cells as seen on light microscopy and by multicentric lamellated membranous inclusions (lamellar bodies) as seen in the cytoplasm on electron microscopy (Anderson and Borlak, 2006; Hruban, 1984). Many cationic amphiphilic drugs (CADs) (including antidepressants, antimicrobial agents, and antiarrhythmic agents) have been reported to induce phospholipidosis, and drug-induced phospholipidosis is a concern of the pharmaceutical industry and regulatory agencies (Berridge et al., 2007; Kruhlak et al., 2008). In the case that phospholipidosis is detected in preclinical studies, a wide margin between drug efficacy and drug toxicity will be required to move forward to clinical development. It is therefore essential to evaluate phospholipidosis-inducing potential of compounds at an early stage in order to reduce the risk of attrition in drug development. Phospholipidosis-inducing potential of compounds is

generally assessed in pharmaceutical companies by histopathological examination in repeated dose toxicity studies, and consequently alternative methods are required for the early prediction of phospholipidosis.

Analysis of genome-wide gene expression data has been shown in numerous studies to contribute to the elucidation of mechanism of actions of novel compounds, adverse effects (Chiang and Butte, 2009), or induced injuries (Blomme et al., 2009). Pharmaceutical research and development divisions are constantly under pressure to discover and develop new compounds, and we have experienced an increase in both time and cost of this process from the early concept to marketing of a successful drug (DiMasi et al., 2003). It is not surprising, therefore, that companies have embraced gene expression analysis in their portfolio of techniques, and several major Japanese pharmaceutical companies and national institutes have jointly created a database of rat in vivo/in vitro and human in vitro transcriptome microarray data in the Japanese Toxicogenomics and Toxicogenomics Informatics Projects (TGP/TGP2) (Urushidani and Nagao, 2005). Whereas microarray data are also publicly available most notably from the GEO and Stanford microarray databases, the current database from the TGP project represents a unique opportunity for the pharmaceutical industry to analyze compound-induced toxicities with its more than 27,000 Affymetrix chip samples covering several dose and time points for approximately 150 compounds, most of which are drugs that have been shown to induce one or more toxicities.

Structure–activity relationship (SAR) models, phospholipid staining, and biomarker-based approaches have been developed to support evaluation of the phospholipidosis-inducing potential of compounds. SAR models are mainly based on pKa (ionization constant) and cLogP (hydrophobicity) (Kuroda and Saito, 2010; Ploemen et al., 2004), but some processes are difficult to explain by simple parameters (e.g. binding of compounds to phospholipids, inhibition of phospholipid degradation) (Halliwell, 1997; Kuroda and Saito, 2010; Taillardat-Bertschinger et al., 2002). The phospholipid staining approach is suitable for high-throughput screening (HTS) using cultured cell lines, but evaluation of drug metabolite toxicities is difficult with this method (Fujimura et al., 2007; Nioi et al., 2007). Furthermore, the same concentration for all compounds is used in this approach although the organ distribution is different. Finally, several biomarker candidates for evaluation of phospholipidosis-inducing potential of compounds have been reported in previous studies. In particular, di-docosahexaenoyl (C22:6)-bis(monoacylglycerol) phosphate (Baronas et al., 2007; Tengstrand et al., 2010) and phenylacetylglycine (Delaney et al., 2004) may have utility as noninvasive urinary phospholipidosis biomarkers. In a recent toxicogenomic approach, 78 probe-sets of rat hepatic genes were extracted from microarray data measured after repeated dosing. The predictive performance, however, of these biomarkers leave room for improvement, and accurate biomarkers, which can be detected in single-dose toxicity studies, are still highly required (Hirode et al., 2008; Reasor et al., 2006; Sawada et al., 2005).

In the present study, we extracted a set of biomarker candidates to predict phospholipidosis-inducing potential using the comprehensive TGP database. The set of biomarkers was optimized on hepatic gene expression data measured 24 h after single dose administration of individual compounds from this database and subsequently assessed by an external validation set that was not available at the time of biomarker optimization.

2. Materials and methods

2.1. Compounds

The chemical name, abbreviation, dosage, vehicle, administration route, and phospholipidosis-inducing potential are summarized in Table 1 for all compounds used in the present study.

2.2. Animal treatment

Animal treatment has been described previously (Takashima et al., 2006). Five-week-old male Sprague-Dawley rats were purchased from Charles River Japan Inc. (Kanagawa, Japan). After a 7-day quarantine and acclimatization period, 6-week-old animals were assigned to dosage groups using a computerized stratified random grouping method based on individual body weight. The animals were individually housed in stainless-steel cages in a room that was lighted for 12 h (7:00–19:00) daily, ventilated with an air-exchange rate of 15 times per hour, and maintained at 21–25 °C with a relative humidity of 40–70%. Each animal was allowed free access to water and pellet food (CRF-1, sterilized by radiation, Oriental Yeast Co., Japan). Rats in each group were orally administered with various drugs suspended or dissolved either in 0.5% methylcellulose solution (MC) or corn oil according to their dispersibility, except that cephalothin was dissolved in saline and administered intravenously. Animals were sacrificed 24 h after the last dose. The experimental protocols were reviewed and approved by the Ethics Review Committee for Animal Experimentation of the National Institute of Health Sciences.

2.3. Microarray analysis

An aliquot of the tissue sample (approximately 30 mg) for microarray analysis was obtained from the left lateral lobe of the liver in each animal immediately after sacrifice, kept in RNAlater (Ambion, Austin, TX, USA) overnight at 4 °C, and then frozen at −80 °C until use. Liver samples were homogenized with the buffer RLT supplied in RNeasy Mini Kit (Qiagen, Valencia, CA, USA), and total RNA was isolated according to the manufacture's instructions. Microarray experiments were performed using GeneChip® RAT Genome 230 2.0 Arrays (Affymetrix, Santa Clara, CA, USA), containing 31,042 probe-sets. The procedure was conducted according to the manufacturer's instructions for cDNA synthesis, purification, and biotin-labeled cRNA synthesis. Ten micrograms of fragmented cRNA was hybridized to a Rat Genome 230 2.0 Array for 18 h at 45 °C at 60 rpm, after which the array was washed and stained by streptavidin–phycoerythrin using Fluidics Station 400 (Affymetrix) and then scanned with a Gene Array Scanner (Affymetrix). The digital image files were processed with an Affymetrix Microarray Suite version 5.0 (MAS5.0) and the intensities were normalized for each chip by setting the mean intensity to 100. Signal intensities were averaged over three replicates for each dosage groups. Microarray data and histopathological findings from TGP/TGP2 as used in this study have been made publicly available in the Open TG-Gates website (http://toxico.nibio.go.jp/).

2.4. Genetic algorithm

Our group hypothesized that compounds with the capability to induce phospholipidosis possibly give rise to a characteristic gene expression pattern leading to phospholipid accumulation and that this pattern may exist at a time point much earlier than the time point at which phospholipidosis can be confirmed in histopathology analysis. We therefore set out to identify a distinguishing pattern by developing a genetic algorithm (Golub et al., 1999) and applied it to single-dose 24-h gene expression data. The gene expression data were initially filtered both for fold change relative to vehicle and for expression level. We used a signal expression threshold of 300.0 and a fold change criteria of 1.6. The genetic algorithm was set up with a population of 50 chromosomes each representing a set of 25 marker candidates, which were initially selected at random. The mutation rate and crossover probability were set to 0.1 and 0.8, respectively, which may be commonly used values. Average fold change values of marker candidates were computed as follows. For each marker candidate, the number of positive training set compounds in which the marker was increased (a), decreased (b), and not changed (c) were counted. If (a) was greater than (b), the fold change value was computed as the average over positive compounds in which the marker was increased. If (b) was greater than (a), the fold change value was computed as the average over positive compounds in which the marker was decreased. If (a) was equal to (b), the marker would not be a candidate for optimization. The phospholipidosis signature was defined as the set of the logarithm of the average fold change values computed as described above for the marker candidates. The fitness function in the genetic algorithm computes the Pearson correlation coefficient between the phospholipidosis signature and the logarithm of fold change values obtained from the training set samples for the set of selected marker candidates (see Table 3). A total of 25,000 generations were generated at which point the fitness values did no longer change. Accordingly, 1,250,000 chromosomes, i.e., sets of marker candidates, were evaluated in a single optimization cycle. The genetic algorithm optimized the fitness function as described above, it was required that all training set compounds should be correctly evaluated, and that the variance of the fitness values over the positive compounds should be minimized. The phospholipidosis signature was subsequently validated for predicting the phospholipidosis-inducing potential on novel in-house data that were prepared only after the final signature had been selected. This signature and the validation are detailed in the following.

2.5. Assessment with external validation set

Following optimization of the biomarker set for phospholipidosis prediction, an external validation set comprising 11 compounds (Table 1) was used to assess

Table 1
Training set and independent validation set compounds.

Compound	Abbreviation	Dose (mg/kg)	Vehicle	Route	Phospholipidosis induction
Training set					
Azathioprine	AZP	30	MC	p.o.	Negative
Benziodarone	BZD	300	MC	p.o.	Negative
Cephalothin	CLT	2000	Saline	i.v.	Negative
Cimetidine	CIM	1000	MC	p.o.	Negative
Clofibrate	CFB	300	Corn oil	p.o.	Negative
Colchicine	COL	15	MC	p.o.	Negative
Cyclosporine A	CSA	300	Corn oil	p.o.	Negative
Danazol	DNZ	2000	MC	p.o.	Negative
Diltiazem	DIL	800	MC	p.o.	Negative
Disopyramide	DIS	400	MC	p.o.	Negative
Enalapril	ENA	600	MC	p.o.	Negative
Famotidine	FAM	1000	MC	p.o.	Negative
Glibenclamide	GBC	1000	Corn oil	p.o.	Negative
Ibuprofen	IBU	400	MC	p.o.	Negative
Mefenamic acid	MEF	300	MC	p.o.	Negative
Metformin	MEM	1000	MC	p.o.	Negative
Omeprazole	OPZ	1000	MC	p.o.	Negative
Phenytoin	PHE	600	MC	p.o.	Negative
Quinidine	QND	200	MC	p.o.	Negative
Rifampicin	RIF	200	MC	p.o.	Negative
Tacrine	TAC	30	MC	p.o.	Negative
Amiodarone	AM	200	MC	p.o.	Positive
Clomipramine	CPM	100	MC	p.o.	Positive
Haloperidol	HPL	30	MC	p.o.	Positive
Hydroxyzine	HYZ	100	MC	p.o.	Positive
Imipramine	IMI	100	MC	p.o.	Positive
Perhexiline	PH	150	MC	p.o.	Positive
Independent validation set					
Cyproterone	CP	100	Corn oil	p.o.	Negative
Diflunisal	DF	100	MC	p.o.	Negative
Doxycycline	DOX	250	MC	p.o.	Negative
Ofloxacin	OFL	200	MC	p.o.	Negative
Procainamide	PROC	50	MC	p.o.	Negative
Chloroquine	CQ	100	MC	p.o.	Positive
Citalopram	CTP	150	Corn oil	p.o.	Positive
Clozapine	CLZ	100	MC	p.o.	Positive
Fluoxetine	FLX	100	MC	p.o.	Positive
Maprotiline	MPT	150	MC	p.o.	Positive
Zimelidine	ZIM	100	MC	p.o.	Positive

MC, 0.5 w/v% methylcellulose; p.o., peroral; i.v., intravenous.

the accuracy, sensitivity, and specificity of the final model. The external validation data were obtained from experiments conducted in-house, and they are therefore independent of the TGP database. We purchased rats that were identical in terms of strain, age, and supplier, and employed the same Affymetrix GeneChip system, whereas the laboratory in which the actual experiments were performed, the lots of reagents, and the technicians in charge differed.

In order to calculate the receiver operating characteristic (ROC) curves, we used the ROC function of the Epi package (Carstensen, 2007) included in the statistical software R (version 2.13.0.).

2.6. Pathway analysis

The Ingenuity Pathways Analysis software (IPA® 8.5, Ingenuity Systems, www.ingenuity.com) was used to identify any presence of significant canonical pathways in the final set of biomarkers.

3. Results

3.1. Sample selection and liver histopathology

Six compounds known to induce phospholipidosis (AM, CPM, HPL, HYZ, IMI, and PH) were selected as the positive part of the training set (Table 1). Vacuolization was observed in the livers after repeated administration of AM (for 7–28 days), CPM (for 28 days), HYZ (for 14–28 days), and IMI (for 7–28 days). In the case of HPL and PH, no vacuolization was seen after repeated administration for 28 days under our assay conditions. These compounds, however, have been confirmed to induce phospholipidosis as reported in the literature (Table 2) (Atienzar et al., 2007; Kasahara et al.,

2006; Pelletier et al., 2007; Tomizawa et al., 2006). Twenty-one compounds, which have not been reported to be related to phospholipidosis, were chosen for the negative part of the training set. The compounds were selected from among the remaining TGP compounds, considering diversity in both molecular structure and toxicity (Table 1).

3.2. Biomarker discovery

In order to develop a biomarker model for prediction of phospholipidosis-inducing potential using hepatic gene expression data measured only 24 h after single dose administration, we applied the genetic algorithm to the training set of 21 negative and 6 positive compounds (Table 1), as described in Section 2. The optimization resulted in a set of biomarkers, i.e., a phospholipidosis signature. The phospholipidosis signature itself is listed in Table 3 as annotated with Affymetrix probe set identifiers, gene symbols and gene titles. A phospholipidosis score for an individual compound is computed as Pearson's correlation coefficient between the phospholipidosis signature and the logarithm of the fold change values obtained from the compound. The calculated values for all positive compounds in the training set are listed in Table 3. The phospholipidosis signature was capable of correctly characterizing each of the training set compounds regarding phospholipidosis-inducing potential (Fig. 1). It was necessary to introduce a threshold for the phospholipidosis score using the training set data in order to be able

Table 2

Histopathological findings and references for 6 positive training set compounds.

Compound	Findings	Topography	24 h	4 day	8 day	15 day	29 day	Reference
Amiodarone	Vacuolization	Bile duct cell	(−)	(−)	3/5 (±)	1/5 (±), 4/5 (+)	4/4 (+)	1, 2, 3, 4
	Vacuolization	Hepatocyte	(−)	(−)	(−)	2/5 (±)	4/4 (±)	
	Vacuolization	Kupper cell	(−)	(−)	(−)	5/5 (±)	4/4 (±)	
Clomipramine	Vacuolization	Hepatocyte	(−)	(−)	(−)	(−)	1/5 (±), 1/5 (+)	1, 2, 3, 4
Haloperidol	Not applicable	Not applicable	(−)	(−)	(−)	(−)	(−)	3, 4
Hydroxyzine	Vacuolization	Hepatocyte	(−)	(−)	(−)	2/5 (±), 2/5 (+)	4/5 (+), 1/5 (2+)	1
Imipramine	Vacuolization	Hepatocyte	(−)	(−)	4/5 (±)	1/5 (±), 2/5 (+)	1/4 (±), 3/4 (+)	1, 2, 3
	Vacuolization	Kupper cell	(−)	(−)	(−)	3/5 (±)	(−)	
Perhexiline	Not applicable	Not applicable	(−)	(−)	(−)	(−)	(−)	1, 2, 4

−, negative; ±, minimal; +, slight; 2+, moderate; 3+, severe.

References: (1) Pelletier et al. (2007); (2) Tomizawa et al. (2006); (3) Kasahara et al. (2006); (4) Atienzar et al. (2007).

to make objective decisions when the phospholipidosis signature is applied to novel data. As seen in Fig. 1, any threshold value in the range from 0.48 to 0.60 separates the scores from the positive and negative compounds and the intermediate value 0.54 is therefore adequate as an objective threshold.

The result of the IPA canonical pathway analysis on the set of 25 biomarkers is listed in Table 4. The most significantly represented pathway is the Metabolism of Xenobiotics by Cytochrome P450, as is also evident from the list of individual biomarkers in Table 3, where several P450 and UGT enzymes are present. The following three pathways "linoleic acid metabolism", "fatty acid metabolism",

and "arachidonic acid metabolism" from the IPA analysis are indeed related to lipid metabolism underscoring the hypothesis that there exist characteristic genes exhibiting an early gene expression pattern leading to phospholipidosis. These pathways most probably reflect the change of lipid metabolism in association with phospholipidosis.

3.3. Validation using the external data set

The predictive performance of the novel set of biomarkers derived in the present study was assessed by data from an

Table 3

List of 25 probe sets for evaluation of phospholipidosis-inducing potential at 24 h following single dose oral administration.

Probe set ID	Gene symbol	Gene title	Phospholipidosis signature (a)	Log 2 of fold-change values (b)					
				AM	CPM	HPL	HYZ	IMI	PH
1367691_at	Prkcdbp	UDP glucuronosyltransferase 2 family, polypeptide B17	−0.872	0.441	0.254	−1.037	0.202	0.381	−0.724
1368155_at	Cyp2c12	Cytochrome P450, family 2, subfamily c, polypeptide 12	−1.315	0.162	−0.925	−1.977	0.142	−0.334	−1.234
1368453_at	Fads2	Fatty acid desaturase 2	−1.088	−0.207	−0.474	−1.442	−0.524	−0.363	−0.803
1370281_at	Fabp5	Fatty acid binding protein 5, epidermal	−1.172	−1.087	−0.613	−1.950	−0.850	−1.240	−0.881
1370350_x_at	Rup2	Urinary protein 2	−0.754	0.361	−0.665	0.303	−0.761	−0.748	−0.039
1370387_at	Cyp3a9	Cytochrome P450, family 3, subfamily a, polypeptide 9	1.301	0.447	1.619	0.816	1.592	0.520	0.662
1370698_at	Ugt2b17	UDP glucuronosyltransferase 2 family, polypeptide B17	1.510	0.321	1.494	0.439	1.480	1.555	0.041
1371076_at	Cyp2b1/2b2	Cytochrome P450, family 2, subfamily b, polypeptide 1/2	2.094	0.399	1.644	1.019	3.244	2.272	0.402
1371390_at	Tubb2c	Tubulin, beta 2c	−0.838	−0.886	−0.136	−0.925	0.261	−0.756	−0.063
1371479_at	Mettl7a	Methyltransferase like 7A	0.810	0.451	0.000	0.914	0.609	0.698	0.514
1371942_at	Gstt3	Glutathione S-transferase theta 3	1.278	0.605	−0.351	1.671	0.673	1.492	0.720
1372585_at	RGD1566254	RGD1566254	0.838	0.749	0.765	−0.740	−0.180	−0.188	−0.132
1372727_at	−	EST	−1.003	−0.089	−1.229	−0.809	−0.359	−0.575	0.630
1373992_at	MGC108823	Similar to interferon-inducible GTPase	1.044	0.863	0.710	−0.361	0.079	−0.654	−0.183
1377950_at	RGD1309362	Similar to interferon-inducible GTPase	−1.606	0.354	−0.644	−3.322	−1.613	0.503	−0.840
1378260_at	Adh1	Alcohol dehydrogenase 1	−0.814	−0.166	0.468	0.621	−0.855	−0.061	−0.773
1379888_at	−	EST	0.990	0.113	−0.771	1.239	−0.242	0.889	0.804
1384580_at	C6	Complement component 6	−0.912	−0.158	0.150	−1.187	−0.298	−0.681	0.340
1387093_at	Slco1a4	Solute carrier organic anion transporter family, member 1a4	−0.912	0.695	0.441	0.010	1.247	0.795	0.815
1387094_at	Slco1a4	Solute carrier organic anion transporter family, member 1a4	1.030	0.713	0.656	−0.030	1.322	1.064	0.928
1387865_at	Dut	Deoxyuridine triphosphatase	−0.844	−0.601	−0.306	−1.014	0.185	−0.692	−0.410
1387892_at	Tubb5	Tubulin, beta 5	−0.940	−0.957	−0.240	−1.092	0.227	−1.055	−0.436
1388318_at	Pgk1	Phosphoglycerate kinase 1	0.848	−0.389	−0.039	0.038	0.764	0.928	−0.077
1389967_at	Arl6ip1	AD-ribosylation factor-like 6 interacting protein 1	−0.753	−0.385	−0.273	−0.741	0.341	−0.764	−0.062
1398307_at	Cyp3a18	Cytochrome P450, 3a18	1.057	0.485	−0.245	0.353	1.179	0.016	0.924
Phospholipidosis scores[a]				0.60	0.67	0.78	0.77	0.74	0.70

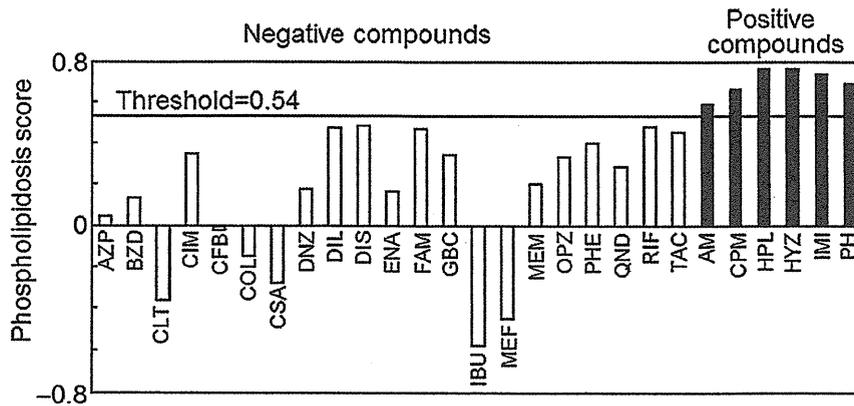[a] Phospholipidosis scores are computed as Pearson's correlation coefficients between (a) and (b).

Fig. 1. Classification of the training set compounds. The phospholipidosis scores computed with the optimized biomarker set (the phospholipidosis signature in Table 3) correctly classified the full training set of compounds (white bars: negative compounds, black bars: positive compounds). For compound abbreviations, see Table 1.

11-compound external validation set, which was independent of the TGP database (Fig. 2a, Table 1). These eleven compounds were selected from lists of phospholipidosis-positive/negative compounds from published studies (Atienzar et al., 2007; Pelletier et al., 2007), purchased from available suppliers, and the corresponding 24 h single-dose microarray data measured in-house as described in Section 2.5. The phospholipidosis scores of all the negative validation compounds (CP, DF, DOX, OFL, and PROC) were correctly lower than the score threshold (0.54), and the scores of 4 out of 6 positive compounds were also correctly larger than the threshold (accuracy = 82%, sensitivity = 67%, and specificity = 100%). Whereas the two compounds (CQ and ZIM) became false negative, their scores were nonetheless close to the threshold. The area under the ROC curve (ROC-AUC) was 0.97 (Fig. 2b). These results suggest that the approach to phospholipidosis prediction as described in this paper can also be applied to data derived from outside of the TGP database.

4. Discussion

The earliest time point at which intracellular vacuolization has been observed among the data for the positive training set was the 7-day repeated administration of AM and IMI. We have hypothesized, however, that histopathologically visible, drug-induced phospholipidosis is preceded by a characteristic gene expression pattern, and that this pattern can be detected already at 24 h following administration of a compound with potential for inducing phospholipidosis. We therefore applied our genetic algorithm to gene expression data measured 24 h after single dose administration, and identified the set of 25 biomarker candidates (Fig. 1,

Table 4
Ingenuity pathways analysis (IPA) of the set of 25 biomarker candidates.

Pathway name	p-Value
Metabolism of xenobiotics by cytochrome P450	**7.59E−08**
Linoleic acid metabolism	**3.02E−05**
Fatty acid metabolism	**8.32E−05**
Arachidonic acid metabolism	**2.69E−03**
Tryptophan metabolism	**2.69E−03**
LPS/IL-1 mediated inhibition of RXR function	**4.90E−03**
Glycolysis/gluconeogenesis	1.74E−02
14-3-3-mediated signaling	1.95E−02
Germ cell–Sertoli cell junction signaling	3.16E−02

Pathways outlined in bold font overlapped with the pathways represented in the biomarkers selected from gene expression data measured after repeated dosing in the Hirode et al.'s study.

Table 3). This set of biomarkers was assessed by the external validation set prepared in-house (Fig. 2).

The mechanisms underlying compound-induced phospholipidosis are not well understood, but several interesting explanations have been published (inhibition of phospholipases, decreased breakdown of phospholipids, increased synthesis of phospholipids) (Halliwell, 1997; Hruban, 1984). Different mechanisms may be also at play according to the characteristics of the individual compounds including in vivo half-life, compound structure, and dose level. We considered, therefore, that an extensive search for biomarkers based on genome-wide expression data would be the most effective approach in this complex situation, and that a multi-marker strategy could possibly improve prediction of compound-specific phospholipidosis-inducing potential. In order to identify such
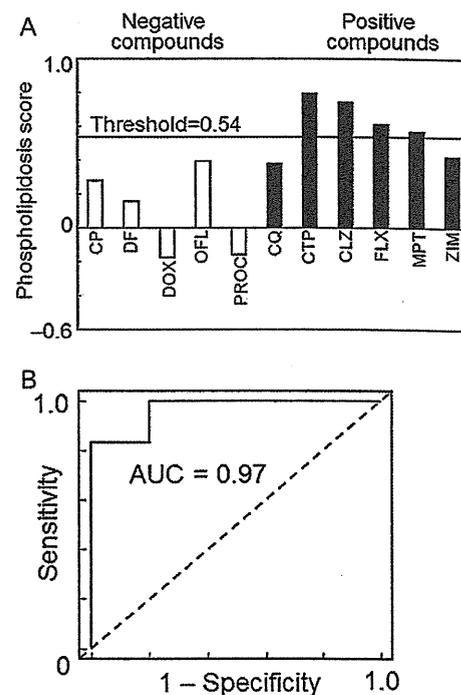

Fig. 2. Predictions for the independent validation set. The phospholipidosis-scores for the independent validation set (white bars: negative compounds, black bars: positive compounds). For compound abbreviation, see Table 1. The ROC curve from prediction of phospholipidosis-inducing potential. ROC, receiver operating characteristic; AUC, area under the curve.

markers, we initially searched for consistently regulated probe-sets in the training set, i.e., the probe-sets that were regulated across all compounds known to induce phospholipidosis, and that were not regulated in the negative compounds. Regulated probe-sets were defined as probe-sets with fold change values above the given threshold. It was, however, not possible to identify consistently regulated probe-sets for phospholipidosis prediction, even if it was possible to identify such markers across for example PPAR agonist compounds in the TGP database (unpublished data). Apparently, the difficulties in extracting marker candidates suggested that there might be different mechanisms, timings, or other factors involved for the selected training set compounds. Therefore, it was necessary to employ less strict criteria that did not require the probe-sets to be regulated in all phospholipidosis-inducing compounds in order to obtain probe-sets for further analysis. The results from the IPA pathway analysis and the 11-compound validation procedure suggest that it was possible in this way to pick up probe-sets for genes implicated in the mechanism of phospholipidosis, as the 25 markers contained in the phospholipidosis signature are related to lipid metabolism pathways (Table 4).

As mentioned in Section 1, alternative and publicly available sets of biomarkers exist for phospholipidosis prediction, notably the set from the Hirode et al.'s study (Hirode et al., 2008), which was also derived based on data from the TGP database. The study, however, employed microarray gene expression data measured after repeated dosing, i.e., at much longer time points. When we compared our set of biomarkers with the set of 78 biomarkers selected from the phenotypically anchored gene expression profiles in the Hirode et al. study, we made the remarkable observation that the top six of the nine pathways listed in Table 4 overlap with the pathways represented in the set of 78 biomarkers (Supplementary Table 1). This suggests that these two biomarker sets reflect to some extent similar biological processes underlying phospholipidosis. There is, however, only an overlap of 3 genes (Ugt2b17, Cyp2b2, and Slco1a4) when the two sets are directly compared. It is therefore thought that our novel set of biomarkers represents early gene expression changes before phospholipidosis is evident, and that we can predict phospholipidosis-inducing potential of compounds by our biomarkers using gene expression data at the 24 h time point following single dose administration.

Drugs may induce phospholipidosis on a more severe level in organs other than the liver, and phospholipidosis screening should preferably be able to correctly evaluate such compounds. In the present study, the compounds CPM and FLX are known to induce phospholipidosis mainly in the lung (Gonzalez-Rothi et al., 1995; Sgaragli et al., 1983), and it is promising that these compounds were correctly characterized by our set of biomarker candidates. Phospholipidosis may occur in multiple organs and tissues as stated above, but most compounds pass through liver and it would be therefore expected that phospholipidosis-inducing potential can be evaluated by analysis of the hepatic gene expression pattern even in the case that liver is not the main target organ for phospholipidosis.

In summary, a set of biomarker candidates have been optimized to predict phospholipidosis-inducing potential of compounds using rat liver microarray data measured 24 h following single dose administration. The set of biomarkers was capable of correctly classifying all training set compounds, and when applied to an external validation set comprising 11 compounds, the phospholipidosis-inducing potential for 9 compounds was correctly predicted. Whereas histopathological changes are only observed after repeated administration, we will be able to assess phospholipidosis-inducing potential of compounds using the set of early time-point biomarkers as described in the present study. We, therefore, conclude that the set of 25 biomarkers will be a valuable tool for phospholipidosis screening in the early phase of drug

development and this should bring about distinct advantages in terms of cost and speed in drug discovery.

Conflict of interest

There is no conflict of interest in this study.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tox.2012.02.015.

References

Anderson, N., Borlak, J., 2006. Drug-induced phospholipidosis. FEBS Lett. 580, 5533–5540.

Atienzar, F., Gerets, H., Dufrane, S., Tilmant, K., Cornet, M., Dhalluin, S., Ruty, B., Rose, G., Canning, M., 2007. Determination of phospholipidosis potential based on gene expression analysis in HepG2 cells. Toxicol. Sci. 96, 101–114.

Baronas, E.T., Lee, J.W., Alden, C., Hsieh, F.Y., 2007. Biomarkers to monitor drug-induced phospholipidosis. Toxicol. Appl. Pharmacol. 218, 72–78.

Berridge, B.R., Chatman, L.A., Odin, M., Schultze, A.E., Losco, P.E., Meehan, J.T., Peters, T., Vonderfecht, S.L., 2007. Phospholipidosis in nonclinical toxicity studies. Toxicol. Pathol. 35, 325.

Blomme, E.A., Yang, Y., Waring, J.F., 2009. Use of toxicogenomics to understand mechanisms of drug-induced hepatotoxicity during drug discovery and development. Toxicol. Lett. 186, 22–31.

Carstensen, B., 2007. Age-period-cohort models for the Lexis diagram. Stat. Med. 26, 3018–3045.

Chiang, A.P., Butte, A.J., 2009. Data-driven methods to discover molecular determinants of serious adverse drug events. Clin. Pharmacol. Ther. 85, 259–268.

Delaney, J., Neville, W.A., Swain, A., Miles, A., Leonard, M.S., Waterfield, C.J., 2004. Phenylacetylglycine, a putative biomarker of phospholipidosis: its origins and relevance to phospholipid accumulation using amiodarone treated rats as a model. Biomarkers 9, 271–290.

DiMasi, J.A., Hansen, R.W., Grabowski, H.G., 2003. The price of innovation: new estimates of drug development costs. J. Health Econ. 22, 151–185.

Fujimura, H., Dekura, E., Kurabe, M., Shimazu, N., Koitabashi, M., Toriumi, W., 2007. Cell-based fluorescence assay for evaluation of new-drugs potential for phospholipidosis in an early stage of drug development. Exp. Toxicol. Pathol. 58, 375–382.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S., 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286, 531–537.

Gonzalez-Rothi, R.J., Zander, D.S., Ros, P.R., 1995. Fluoxetine hydrochloride (Prozac)-induced pulmonary disease. Chest 107, 1763–1765.

Halliwell, W.H., 1997. Cationic amphiphilic drug-induced phospholipidosis. Toxicol. Pathol. 25, 53–60.

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

Hruban, Z., 1984. Pulmonary and generalized lysosomal storage induced by amphiphilic drugs. Environ. Health Perspect. 55, 53–76.

Kasahara, T., Tomita, K., Murano, H., Harada, T., Tsubakimoto, K., Ogihara, T., Ohnishi, S., Kakinuma, C., 2006. Establishment of an in vitro high-throughput screening assay for detecting phospholipidosis-inducing potential. Toxicol. Sci. 90, 133–141.

Kruhlak, N.L., Choi, S.S., Contrera, J.F., Weaver, J.L., Willard, J.M., Hastings, K.L., Sancilio, L.F., 2008. Development of a phospholipidosis database and predictive quantitative structure–activity relationship (QSAR) models. Toxicol. Mech. Methods 18, 217–227.

Kuroda, Y., Saito, M., 2010. Prediction of phospholipidosis-inducing potential of drugs by in vitro biochemical and physicochemical assays followed by multivariate analysis. Toxicol. In Vitro 24, 661–668.

Nioi, P., Perry, B.K., Wang, E.J., Gu, Y.Z., Snyder, R.D., 2007. In vitro detection of drug-induced phospholipidosis using gene expression and fluorescent phospholipid based methodologies. Toxicol. Sci. 99, 162–173.

Pelletier, D.J., Gehlhaar, D., Tilloy-Ellul, A., Johnson, T.O., Greene, N., 2007. Evaluation of a published in silico model and construction of a novel Bayesian model for predicting phospholipidosis inducing potential. J. Chem. Inf. Model. 47, 1196–1205.

Ploemen, J.P., Kelder, J., Hafmans, T., van de Sandt, H., van Burgsteden, J.A., Saleminki, P.J., van Esch, E., 2004. Use of physicochemical calculation of pKa and CLog P to predict phospholipidosis-inducing potential: a case study with structurally related piperazines. Exp. Toxicol. Pathol. 55, 347–355.

Reasor, M.J., Hastings, K.L., Ulrich, R.G., 2006. Drug-induced phospholipidosis: issues and future directions. Expert Opin. Drug Saf. 5, 567–583.

Reasor, M.J., Kacew, S., 2001. Drug-induced phospholipidosis: are there functional consequences? Exp. Biol. Med. (Maywood) 226, 825–830.

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

Sgaragli, G.P., Della Corte, L., Gremigni, D., 1983. Chlorimipramine-induced phospholipidosis: biochemical and pharmacokinetic observations in the rat. Pharmacol. Res. Commun. 15, 231–246.

Taillardat-Bertschinger, A., Martinet, C.A., Carrupt, P.A., Reist, M., Caron, G., Fruttero, R., Testa, B., 2002. Molecular factors influencing retention on immobilized artificial membranes (IAM) compared to partitioning in liposomes and n-octanol. Pharm. Res. 19, 729–737.

Takashima, K., Mizukawa, Y., Morishita, K., Okuyama, M., Kasahara, T., Toritsuka, N., Miyagishima, T., Nagao, T., Urushidani, T., 2006. Effect of the difference in vehicles on gene expression in the rat liver—analysis of the control data in the Toxicogenomics Project Database. Life Sci. 78, 2787–2796.

Tengstrand, E.A., Miwa, G.T., Hsieh, F.Y., 2010. Bis(monoacylglycerol)phosphate as a non-invasive biomarker to monitor the onset and time-course of phospholipidosis with drug-induced toxicities. Expert Opin. Drug Metab. Toxicol. 6, 555–570.

Tomizawa, K., Sugano, K., Yamada, H., Horii, I., 2006. Physicochemical and cell-based approach for early screening of phospholipidosis-inducing potential. J. Toxicol. Sci. 31, 315–324.

Urushidani, T., Nagao, T., 2005. Toxicogenomics: the Japanese initiative. In: Borak, J. (Ed.), Handbook of Toxicogenomics—Strategies and Applications. Wiley-VCH, Weinheim, pp. 623–631.

Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database

Takeki Uehara [a,b,*], Yohsuke Minowa [b], Yuji Morikawa [b], Chiaki Kondo [a], Toshiyuki Maruyama [a], Ikuo Kato [a], Noriyuki Nakatsu [b], Yoshinobu Igarashi [b], Atsushi Ono [b], Hitomi Hayashi [c,d], Kunitoshi Mitsumori [c], Hiroshi Yamada [b], Yasuo Ohno [e], Tetsuro Urushidani [b,f]

[a] Drug Developmental Research Laboratories, Shionogi & Co., Ltd., 3-1-1 Futaba-cho, Toyonaka, Osaka 561-0825, Japan
[b] Toxicogenomics Informatics Project, National Institute of Biomedical Innovation, 7-6-8 Asagi, Ibaraki, Osaka 567-0085, Japan
[c] Laboratory of Veterinary Pathology, Tokyo University of Agriculture and Technology, 3-5-8 Harumi-cho, Fuchu, Tokyo 183-8509, Japan
[d] Pathogenetic Veterinary Science, United Graduate School of Veterinary Sciences, Gifu University, 1-1 Yanagido, Gifu, 501-1193 Gifu, Japan
[e] National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan
[f] Department of Pathophysiology, Faculty of Pharmaceutical Sciences, Doshisha Women's College of Liberal Arts, Kodo, Kyotanabe, Kyoto 610-0395, Japan

ARTICLE INFO

ABSTRACT

The present study was performed to develop a robust gene-based prediction model for early assessment of potential hepatocarcinogenicity of chemicals in rats by using our toxicogenomics database, TG-GATEs (Genomics-Assisted Toxicity Evaluation System developed by the Toxicogenomics Project in Japan). The positive training set consisted of high- or middle-dose groups that received 6 different non-genotoxic hepatocarcinogens during a 28-day period. The negative training set consisted of high- or middle-dose groups of 54 non-carcinogens. Support vector machine combined with wrapper-type gene selection algorithms was used for modeling. Consequently, our best classifier yielded prediction accuracies for hepatocarcinogenicity of 99% sensitivity and 97% specificity in the training data set, and false positive prediction was almost completely eliminated. Pathway analysis of feature genes revealed that the mitogen-activated protein kinase p38- and phosphatidylinositol-3-kinase-centered interactome and the v-myc myelocytomatosis viral oncogene homolog-centered interactome were the 2 most significant networks. The usefulness and robustness of our predictor were further confirmed in an independent validation data set obtained from the public database. Interestingly, similar positive predictions were obtained in several genotoxic hepatocarcinogens as well as non-genotoxic hepatocarcinogens. These results indicate that the expression profiles of our newly selected candidate biomarker genes might be common characteristics in the early stage of carcinogenesis for both genotoxic and non-genotoxic carcinogens in the rat liver. Our toxicogenomic model might be useful for the prospective screening of hepatocarcinogenicity of compounds and prioritization of compounds for carcinogenicity testing.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Carcinogenicity is one of the most serious side effects associated with new drug development. Thus, it is especially important for pharmaceutical companies to know as much as possible about the eventual carcinogenic properties of new drugs, even in the early stages of drug development. However, the current "gold standard" for carcinogenicity testing is a bioassay in which mice and rats are treated with a target compound for their entire 2-year lifespan. This carcinogenicity testing cannot be performed in the early stage of drug development because it is time-consuming and expensive and it requires the use of many animals and large amounts of chemicals.

Additionally, while this assay provides evidence of carcinogenicity of the target chemicals in rodents, it provides only limited mechanistic information about carcinogenesis. Thus, the current strategy of a 2-year bioassay to evaluate in vivo carcinogenicity is not satisfactory. It should be replaced with better test systems that are cheaper and faster, use fewer animals, and provide the appropriate sensitivity and specificity desired for a screen of carcinogenic potential.

Toxicogenomics has been expected as a powerful approach for elucidating mechanisms underlying toxicological endpoints and a useful strategy for the early detection of potential chemical toxicity (Battershill, 2005; Heinloth et al., 2004; Irwin et al., 2004; Kiyosawa et al., 2009; Searfoss et al., 2005). Important scientific breakthroughs have been achieved by applying toxicogenomics to the early detection of chemical carcinogenicity in vivo. Studies have focused on hepatocarcinogenicity because the liver is the most common target organ for carcinogenesis. Kramer et al. (2004) studied microarray-derived comprehensive gene

* Corresponding author at: Developmental Research Laboratories, Shionogi & Co., Ltd., 3-1-1 Futaba-cho, Toyonaka, Osaka 561-0825, Japan. Fax: +81 6 6332 6385.
E-mail address: takeki.uehara@shionogi.co.jp (T. Uehara).

expression data from the livers of rats treated with 10 non-genotoxic hepatocarcinogens and determined that the expression level of NAD(P)H P450 reductase was positively correlated with hepatocarcinogenicity and the expression level of transforming growth factor-β clone 22 was negatively correlated with hepatocarcinogenicity. Nie et al. (2006) demonstrated that 6 biomarker genes predict the carcinogenic potential of non-genotoxic hepatocarcinogens with a prediction accuracy of 88.5% by using expression data from the livers of rats treated with a single dose of 24 non-genotoxic chemicals and 28 non-hepatocarcinogens. Nakayama et al. (2006) successfully separated several isomers with or without hepatocarcinogenicity based on the expression profiles of selected genes, and they identified characteristic gene expression changes in hepatocarcinogenic isomers after up to 28 days of repeated dosing. Ellinger-Ziegelbauer et al. (2008) constructed support vector machine (SVM) prediction models by using gene expression data from rats treated for up to 14 days with 13 different chemicals for the training set and rats treated with 16 independent chemicals for the validation set; the resulting prediction models differentiated between genotoxic hepatocarcinogens, non-genotoxic hepatocarcinogens and non-hepatocarcinogens with up to 88% classification accuracy. Fielden et al. (2007) analyzed hepatic gene expression in rats treated with 25 non-genotoxic hepatocarcinogens and 75 non-hepatocarcinogens for 1 to 7 days. They constructed a SVM model consisting of 37 probes that yielded prediction accuracies for carcinogenicity with 86% specificity and 81% sensitivity. These researchers successfully constructed gene-based prediction models from data obtained from up to 1 month of repeated dosing. In contrast, Auerbach et al. (2010) recently demonstrated further evidence that the dosing period is an important factor in the construction of highly accurate prediction models based on toxicogenomics-derived gene expression profiles, especially in the case of weakly carcinogenic compounds. They concluded that a 90-day exposure period is needed to detect gene expression changes specifically related to carcinogenic compounds.

In Japan, the Toxicogenomics Project (TGP) has established a large-scale toxicogenomics database known as TG-GATEs (Genomics-Assisted Toxicity Evaluation System developed by the Toxicogenomics Project in Japan) (Uehara et al., 2010; Urushidani, 2010). In this project, rats were exposed to 3 different doses of 150 compounds for a period ranging from 1 to 28 days; the gene expression in the livers and kidneys of these animals at 8 different time points was comprehensively analyzed using microarrays. We have used this database to identify several different types of biomarker genes and construct prediction models for hepatotoxicity and nephrotoxicity (Gao et al., 2010; Hirode et al., 2008; Kondo et al., 2009; Uehara et al., 2010). Regarding hepatocarcinogenicity, we previously tried to build a gene-based predictor of non-genotoxic hepatocarcinogenicity in rats (Uehara et al., 2008). Consequently, we have successfully built a model, consisting of 112 probes, for the early detection of hepatocarcinogenesis based on gene expression changes that are commonly induced by compounds with hepatocarcinogenicity in rats. However, since this model was trained to achieve early and sensitive detection of potential carcinogenicity after a single exposure to a compound, false positive predictions occurred in some non-carcinogenic hepatotoxins. Moreover, a limited number of compounds were used for training of the model in the study since our database was under construction. In an effort to make high-quality predictive models based on gene expression data, a fairly extensive data set of several compounds with multiple time points and multiple dose levels is required. Now, our large-scale toxicogenomics database has been completed, and microarray data for all 150 compounds are available. In this research, we hypothesized that our large-scale toxicogenomics database might lead to the construction of a more robust and accurate prediction model of hepatocarcinogenicity in rats. By taking into account the findings of the latest work by Auerbach et al. (2010), we have trained a classifier by using data from our longest dosing period (28 days) to decrease the percentage of false-positive predictions. Consequently, our new SVM-based classifier yielded prediction

accuracies for hepatocarcinogenicity with 99% sensitivity and 97% specificity in a training data set, and false-positive predictions were almost completely eliminated. The usefulness and robustness of our predictor were further confirmed in an independent validation data set obtained from a public database. Interestingly, similar positive predictions were obtained for several genotoxic hepatocarcinogens as well as non-genotoxic hepatocarcinogens. In the present report, we provide reliable candidate gene biomarkers in the early stages of the hepatocarcinogenesis that are predictive for both genotoxic and non-genotoxic hepatocarcinogens. Our present toxicogenomic model might be useful to reduce the dependence on 2-year rodent bioassays by instead using a short-term repeated dosing study.

Materials and methods

Animals and experimental design. Five-week-old male Sprague–Dawley rats were obtained from Charles River Japan, Inc. (Kanagawa, Japan). After a 7-day quarantine and acclimatization period, the 6-week-old animals were assigned to dosage groups (5 rats per group) by using a computerized stratified random grouping method based on individual body weight. The animals were individually housed in stainless-steel cages in an animal room that was lighted for 12 h (7:00–19:00) daily, ventilated with an air-exchange rate of 15 times per hour, and maintained at 21 °C–25 °C with a relative humidity of 40%–70%. Each animal was allowed free access to water and a pellet diet (CRF-1, sterilized by radiation, Oriental Yeast Co., Ltd., Tokyo, Japan).

The compounds used in this study are summarized in Table 1 (for detailed experimental conditions, see Supplemental Table 1). A total of 150 compounds were used for training and testing of models. The training data set consisted of 6 positive compounds (necrogenic hepatocarcinogens with no evidence of genotoxicity, namely non-genotoxic hepatocarcinogens) and 54 negative compounds (non-hepatocarcinogens), and the test data set consisted of remaining 90 compounds (for more detailed information, see Supplemental Table 2). According to the standard protocol in our project (Uehara et al., 2010), 5 rats per group were treated with these compounds at 3 different dose levels (low: L, middle: M, and high: H). The maximum tolerated dose of each compound, which was estimated from a preliminary 7-day repeated dosing study, was chosen as the highest dose level. For single-dose studies, rats were euthanized at 3, 6, 9 and 24 h after dosing. For repeated dose studies, the animals were treated daily for 3, 7, 14 and 28 days and euthanized 24 h after the last dosing (4, 8, 15 and 29D). The animals were euthanized by exsanguination from the abdominal aorta under ether anesthesia, and liver samples were collected from the left lateral lobe of the liver immediately after the animals were euthanized. The experimental protocols were carefully reviewed and approved by the Ethics Review Committee for Animal Experimentation of the National Institute of Health Sciences.

RNA extraction and microarray analysis. An aliquot of the sample (about 30 mg) for microarray analysis was obtained from the left lateral lobe of the liver in each animal immediately after the animals were euthanized. The sample was kept in RNAlater® (Ambion, Austin, TX, USA) overnight at 4 °C and then frozen at −80 °C until use. Liver samples were homogenized with buffer RLT supplied in the RNeasy Mini Kit (Qiagen, Valencia, CA, USA), and total RNA was isolated according to the manufacturer's instructions.

Microarray analysis was conducted on 3 of 5 samples for each group by using Affymetrix Rat Genome 230 2.0 arrays (Affymetrix, Santa Clara, CA, USA). The procedure was basically conducted according to the manufacturer's instructions as previously reported (Uehara et al., 2010). The digital image files were preprocessed by Affymetrix Microarray Analysis Suite version 5.0 (MAS5.0). The expression signal values were scaled by the median of each chip sample. The normalized data sets were then converted into the log-ratio of base 2 to the means of the

Table 1

Compounds and carcinogenicity definitions used in this study.

Compound class	Compound name
Training set (positive)	
Non-genotoxic hepatocarcinogen (hepatotoxic oxidative stressor)	Carbon tetrachloride (CCL4), ethionine (ET), thioacetamide (TAA), methapyrilene (MP), coumarin (CMA), monocrotaline (MCT)
Training set (negative)	
Non-hepatocarcinogen	Acetaminophen (APAP), naphthyl isothiocyanate (ANIT), allyl alcohol (AA), theophylline (TEO), trimethadione (TMD), naproxen (NPX), methotrexate (MTX), aspirin (ASA), labetalol (LBT), ketoconazole (KC), tetracycline (TC), metformin (MFM), methyldopa (MDP), vitamin A (VA), chlorpropamide (CPP), nicotinic acid (NIC), famotidine (FAM), ranitidine (RAN), diltiazem (DIL), captopril (CAP), enalapril (ENA), mexiletine (MEX), meloxicam (MLX), lornoxicam (LNX), cyclosporine A (CSA), isoniazid (INAH), phenylbutazone (PhB), nitrofurantoin (NFT), propylthiouracil (PTU), amiodarone (AM), cimetidine (CIM), flutamide (FT), methimazole (MYZ), iproniazid (IPA), chloramphenicol (CMP), furosemide (FUR), chlorpheniramine (CHL), caffeine (CAF), sulpiride (SLP), simvastatin (SST), chlormadinone (CLM), carboplatin (CBP), bucetin (BCT), perhexiline (PH), pemoline (PML), ibuprofen (IBU), erythromycin ethylsuccinate (EME), nifedipine (NIF), sulindac (SUL), disopyramide (DIS), disulfiram (DSF), tolbutamide (TLB), acarbose (ACA), ajmaline (AJM)
Test set	
Genotoxic hepatocarcinogen	Lomustine (LS), acetamidofluorene (AAF), nitrosodiethylamine (DEN)
Non-genotoxic hepatocarcinogen (enzyme inducer)	Phenobarbital (PB), carbamazepine (CBZ), phenytoin (PHE), rifampicin (RIF), hexachlorobenzene (HCB), sulfasalazine (SS)
Non-genotoxic hepatocarcinogen (peroxisome proliferator)	Clofibrate (CFB), WY-14643 (WY), gemfibrozil (GFZ), fenofibrate (FFB)
Non-genotoxic hepatocarcinogen (hormonal modulator)	Ethinylestradiol (EE)
Non-hepatocarcinogen/unknown (non-genotoxicant)	Ethionamide (ETH), indomethacin (IM), bromobenzene (BBZ), ethambutol (EBU), colchicine (COL), clomipramine (CPM), puromycin aminonucleoside (PAN), methyltestosterone (MTS), valproic acid (VPA), chlorpromazine (CPZ), diclofenac (DFNa), benzbromarone (BBr), allopurinol (APL), fluphenazine (FP), thioridazine (TRZ), adapin (ADP), glibenclamide (GBC), chlormezanone (CMN), moxisylyte (MXS), imipramine (IMI), amitriptyline (AMT), hydroxyzine (HYZ), quinidine (QND), mefenamic acid (MEF), tiopronin (TIO), acetazolamide (ACZ), promethazine (PMZ), dantrolene (DTL), triazolam (TZM), terbinafine (TBF), danazol (DNZ), bendazac (BDZ), benziodarone (BZD), bromoethanamine (BEA), nimesulide (NIM), phenylanthranilic acid (NPAA), cephalothin (CLT), ticlopidine (TCP), gentamicin (GMC), vancomycin (VMC), omeprazole (OPZ), diazepam (DZP), haloperidol (HPL), griseofulvin (GF), tamoxifen (TMX), tannic acid (TAN), triamterene (TRI), ethanol (ETN), ciprofloxacin (CPX), tacrine (TAC), nitrofurazone (NFZ), papaverine (PAP), penicillamine (PEN), azathioprine (AZP), doxorubicin (DOX), cyclophosphamide (CPA), etoposide (ETP), cisplatin (CSP), phenacetin (PCT)
Unknown	K01, K02, K03, K04, K05, K06, K07, K08, K09, K10, K11, K12, K13, K14, K15, K16, K17

K01 to K17 were compounds synthesized in member companies.

corresponding control groups. Raw microarray data (CEL files) are available in Open TG-GATEs (http://toxico.nibio.go.jp/).

Gene selection and supervised classification. Among a total of 150 compounds in our database, we have selected 6 compounds consisting of carbon tetrachloride, ethionine, thioacetamide, methapyrilene, coumarin and monocrotaline, as positive training compounds for modeling, which are non-genotoxic hepatocarcinogens with hepatocellular necrosis/degeneration in histopathology following multiple dosing for up to 28 days in our experimental condition. Individual histopathological data of all compounds are available (http://toxico.nibio.go.jp/datalist.html). High-dose 29D groups treated with these compounds were used for the positive training data set, with the exception for monocrotaline; the middle-dose group of monocrotaline at 29D was included in the positive training set because all animals in the high-dose group were dead at 29D in the current experimental condition. High- or middle-dose groups at all time points (3 to 24H for single-dose studies, 4 to 29D for repeated-dose studies) of randomly selected 54 non-hepatocarcinogens were selected as the negative training data set. To exclude genes being transiently regulated by the treatment of non-carcinogenic compounds, data obtained from all time points were used as the negative training set. The remaining compounds were used as the independent test set as follows: (i) genotoxic hepatocarcinogens; (ii) non-genotoxic hepatocarcinogens whose carcinogenic mechanisms are thought to be related to hepatic enzyme induction, peroxisomal proliferation and hormonal modulation; and (iii) non-hepatocarcinogens (for more detailed information, see Supplemental Table 2).

SVM combined with wrapper-type gene selection algorithms was used to build a prediction model, as previously reported (Kondo et al., 2009). First, the probes that were judged as absent in all samples of the training data set using MAS5.0 P/A-call were excluded.

Next, the following 3 statistical parameters were calculated for each probe: (i) fold change of gene expression between positive and negative training data sets; (ii) Mann–Whitney U value; (iii) confident margin of SVM classifier (normal margin corrected by classification accuracy). The probes were filtered by following criteria: (i) fold change >2 or <0.5; (ii) p-value<0.01; and (iii) confident margin >0.05. Then, the combined gene ranking based on the 3 parameters was calculated by using a layer ranking algorithm (Chen et al., 2007). To estimate the performance of our classifier, 5-fold cross-validation was executed as described in our previous report (Kondo et al., 2009). Finally, 9 of the 82 top-ranked probes (7 genes) were selected to maximize the classification accuracy and the area under the curve (AUC) of the receiver operating characteristic curve (ROC).

Ingenuity pathways analysis. The 82 top-ranked probes were analyzed by using Ingenuity Pathways Analysis software (v. 7.1; Ingenuity Systems, Redwood City, CA) to determine the biological networks that were enriched in selected feature genes.

Independent validation of our classifier by using an external data set from NEDO. An external microarray data set from the NEDO project, another toxicogenomics consortium in Japan, was used for independent validation of our classifier (Matsumoto et al., 2009). In this project, the NEDO-ToxArray III consists of 6709 genes, and hepatic gene expression data was comprehensively obtained from F344 rats treated with 88 compounds for up to 28 days. All of the microarray data are available in the public microarray database of Gene Expression Omnibus (GEO). SVM modeling and principal component analysis (PCA) were performed by using the expression data at 3 different time points (4, 15 and 29D).

Predictions using published biomarker genes. For comparison of prediction accuracy with previously published models, we built

SVM models with our training data set by using published gene lists and then compared the prediction performance of all models (Auerbach et al., 2010; Ellinger-Ziegelbauer et al., 2008; Fielden et al., 2007; Nakayama et al., 2006; Uehara et al., 2008). To estimate the performance of each model, 5-fold cross-validation was executed using the training data set.

Results

Gene selection and supervised classification. We trained a binary classifier by using an SVM algorithm combined with wrapper-type gene selection to construct a statistically reliable model without over-fitting to the profiles of training samples according to our previous report with minor modifications (Kondo et al., 2009). By applying statistical analysis for feature selection, 82 probes passed current statistical criteria, and these top-ranked probes are summarized in Supplemental Table 3. 1 to 82 of the top-ranked probes were used to construct the classifiers with further feature selection. A ROC curve and its AUC are plotted in Fig. 1. Although there were no big differences in the prediction accuracy of each classifier, a classifier consisting of 9 probes (7 genes; Table 2) achieved the best classification accuracy under the 5-fold cross-validation. The sensitivity and specificity of the prediction of the optimized classifier was 99% and 97%, respectively.
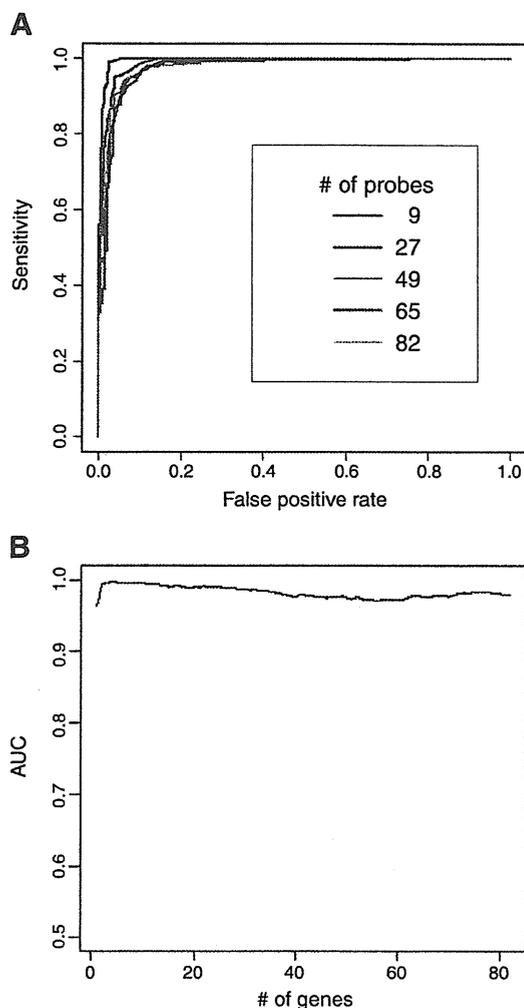
Fig. 1. Receiver operating characteristic analysis of prediction models. ROC curve (A) and area under the ROC curve (B) of prediction models are plotted. A model consisting of 9 probes was selected as the best model in our training data set.

Table 2
The 9 top-ranked probes in the optimized model.

Affymetrix probe ID	Gene title	Gene abbreviation	GO biological process (selected)
1370583_s_at	ATP-binding cassette, sub-family B (MDR/TAP), member 1A/1B	Abcb1a/1b	Nucleotide binding/transporter activity/protein binding/ATP binding/ATPase activity/drug transporter activity/hydrolase activity/nucleoside-triphosphatase activity
1395737_at /1374198_at	Cd276 molecule	Cd276	Receptor binding/protein binding
1367787_at	Islet cell autoantigen 1	Ica1	Protein binding/protein domain specific binding
1379419_at	Transmembrane protein 184C	Tmem184c	Unknown
1379262_at	Acyl-CoA thioesterase 9	Acot9	Unknown
1383401_at	Testis derived transcript	Tes	Zinc ion binding/metal ion binding
1375719_s_at/1373102_at	Cadherin 13	Cdh13	Calcium ion binding/protein binding/low-density lipoprotein binding/protein homodimerization activity/cadherin binding/adiponectin binding

Gene expression profiles of selected feature genes. Fig. 2 shows a heat map of the selected 9 probes and all 82 top-ranked probes. Among the 82 probes, 44 probes were upregulated, and 38 probes were downregulated in positive training compounds. All 9 probes involved in the optimized model were upregulated in positive training compounds. The extent of upregulation of genes in positive compounds was clearly higher than that in negative compounds and changed in a time- and dose-dependent manner (Fig. 3).

Ingenuity pathways analysis. To further characterize the biological significance of alterations in gene expression, functional pathway analysis was performed by using the 82 selected probes. The mitogen-activated protein kinase p38 (*p38 Mapk*)- and phosphatidylinositol-3-kinase (*PI3k*)-centered interactome (Fig. 4A) and the v-myc myelocytomatosis viral oncogene homolog (*Myc*)-centered interactome (Fig. 4B) were the 2 most significant networks. Among 9 probes (7 genes) of the best classifier, the following 5 genes were involved in these networks: ATP-binding cassette, sub-family B (MDR/TAP), member 1A/B (*Abcb1a/b*), Cd276 molecule (*Cd276*), islet cell autoantigen 1 (*Ica1*), testis-derived transcript (*Tes*), and cadherin 13 (*Cdh13*).

Prediction results of all compounds. The SVM classification scores of all 150 compounds are summarized in Supplemental Table 2. All 3 dose groups of 90 test compounds that had not been used as the training set and the remaining groups of the 60 training compounds were used as a test data set for external validation of the classifier. The classifier predicted the following samples as positive: thioacetamide (H: 8D and 15D; M: 29D), methapyrilene (H: 8D and 15D), carbon tetrachloride (M: 29D) and monocrotaline (H: 15D). As expected, positive predictions for several hepatocarcinogens were observed only after long-term repeated dosing. There were no positive predictions in the low-dose groups of these positive-training compounds. Among the
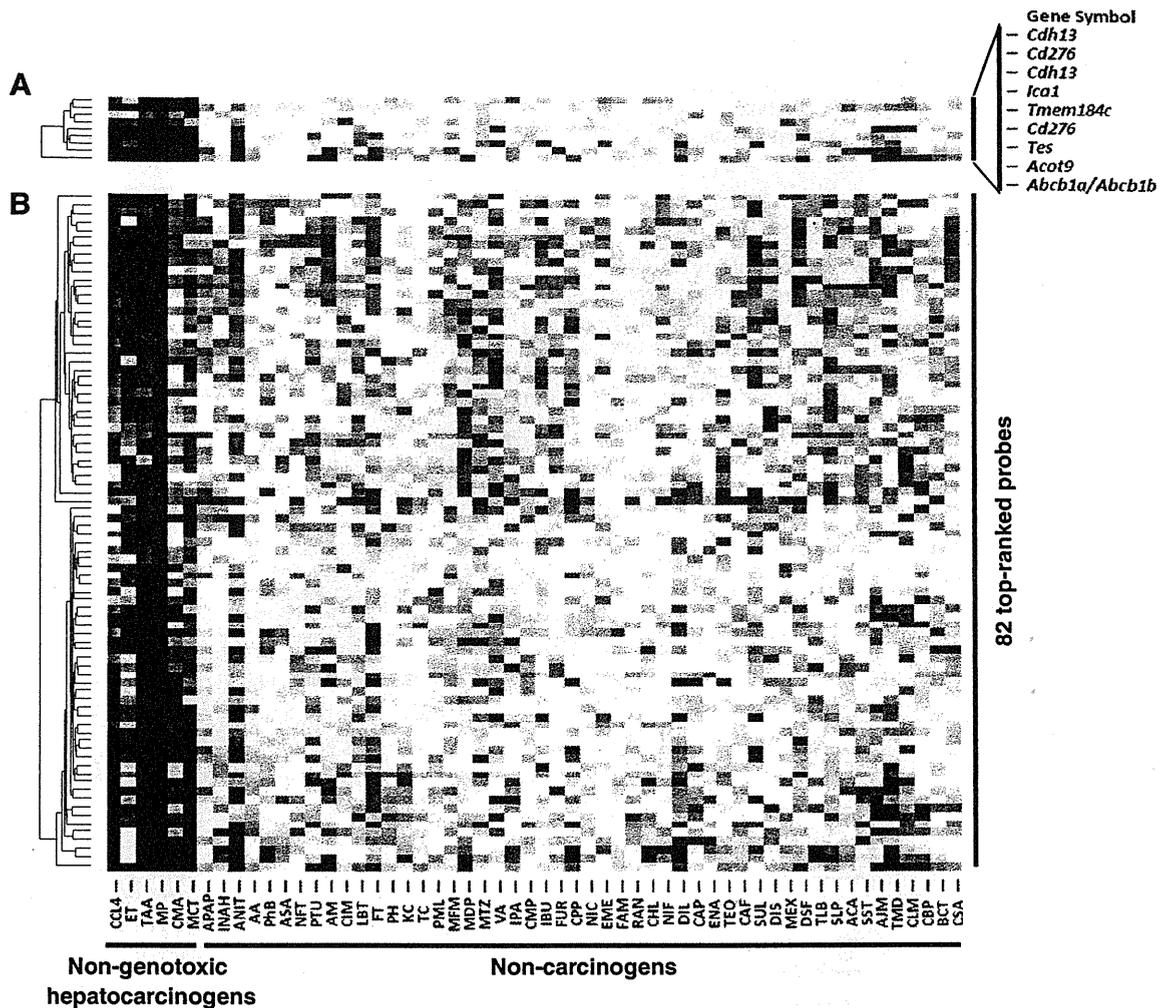
Fig. 2. Hierarchical clustering analysis of expression changes of selected feature genes. Gene expression changes of the 9 (A) and 82 top-ranked probes (B) were analyzed by hierarchical clustering. Clustering method: UPGMA (unweighted average); similarity measure: Euclidean distance; ordering function: average value. Heat map shows log-ratio of base 2 to the means of the corresponding control groups (red: 5-fold higher, blue: -5-fold lower expression of log 2 ratio). Symbols of selected genes are shown on the right side of the heat map.

genotoxins, lomustine (H: 29D), acetamidofluorene (L, M and H: 15 and 29D) and nitrosodiethylamine (M: 29D and H: 8 and 15D), for which the target of carcinogenicity is the liver, were correctly predicted as positive. Several genotoxic compounds that have potential to elicit cancer in other target organs, such as cyclophosphamide, etoposide, cisplatin and carboplatin, colchicine, phenacetin and doxorubicin, were predicted as negative at all time points of the 3 dose groups. Non-genotoxic hepatocarcinogens, including enzyme inducers (phenobarbital, carbamazepine, phenytoin, rifampicin, hexachlorobenzene and sulfasalazine), peroxisome proliferators (clofibrate, Wy-14643, gemfibrozil and fenofibrate) and hormonal modulators (ethinylestradiol), were classified as negative in all groups. The priority focus on this study was to build a prediction model with improved false-positive classification of non-hepatocarcinogens reflecting temporal gene expression changes at early time points after single or short-term repeated dosing. Among the 46 compounds in the negative/unknown test set, almost all compounds were classified as negative except for 2 compounds that were falsely classified as positive, ethionamide (H: 24H, 4D, and 15D; M: 24H) and etambutol (H: 8D).

Further independent validation of the optimized predictor. Since interlaboratory difference is an important issue in the field of toxicogenomics research (Fielden et al., 2008), a further independent

validation was performed by using microarray data obtained by independent laboratories. In the present study, we used an external data set obtained in the NEDO project for further validation (Matsumoto et al., 2009). The NEDO study used custom microarrays consisting of 6709 probes, and gene expression data was comprehensively obtained from the livers of rats treated with several compounds for up to 28 days. Among their data set, we used following 14 compounds commonly involved in our data set: carbon tetrachloride, ethionine, thioacetamide, methapyrilene and nitrosodiethylamine (positive test set), and phenobarbital, phenytoin, hexachlorobenzene, clofibrate, ethinylestradiol, indomethacin, acetaminophen, aspirin and tannic acid (negative test set). For the purpose of a comprehensive comparison, we used 82 top-ranked probes for this analysis. Due to differences in the microarray platform, only 53 out of 82 probes were shared by both microarray platforms. By using the gene expression data for 53 probes, a SVM classifier was built without further feature selection using our training data set and then analyzed prediction accuracy of these test chemicals. As a result, all hepatocarcinogens included in the test data set were correctly predicted as positive regardless of using microarray data obtained by the different platform in independent laboratories. The overall sensitivity and specificity of the prediction by this classifier consisting of 53 probes was 100% and 89%, respectively (Supplemental Table 4).
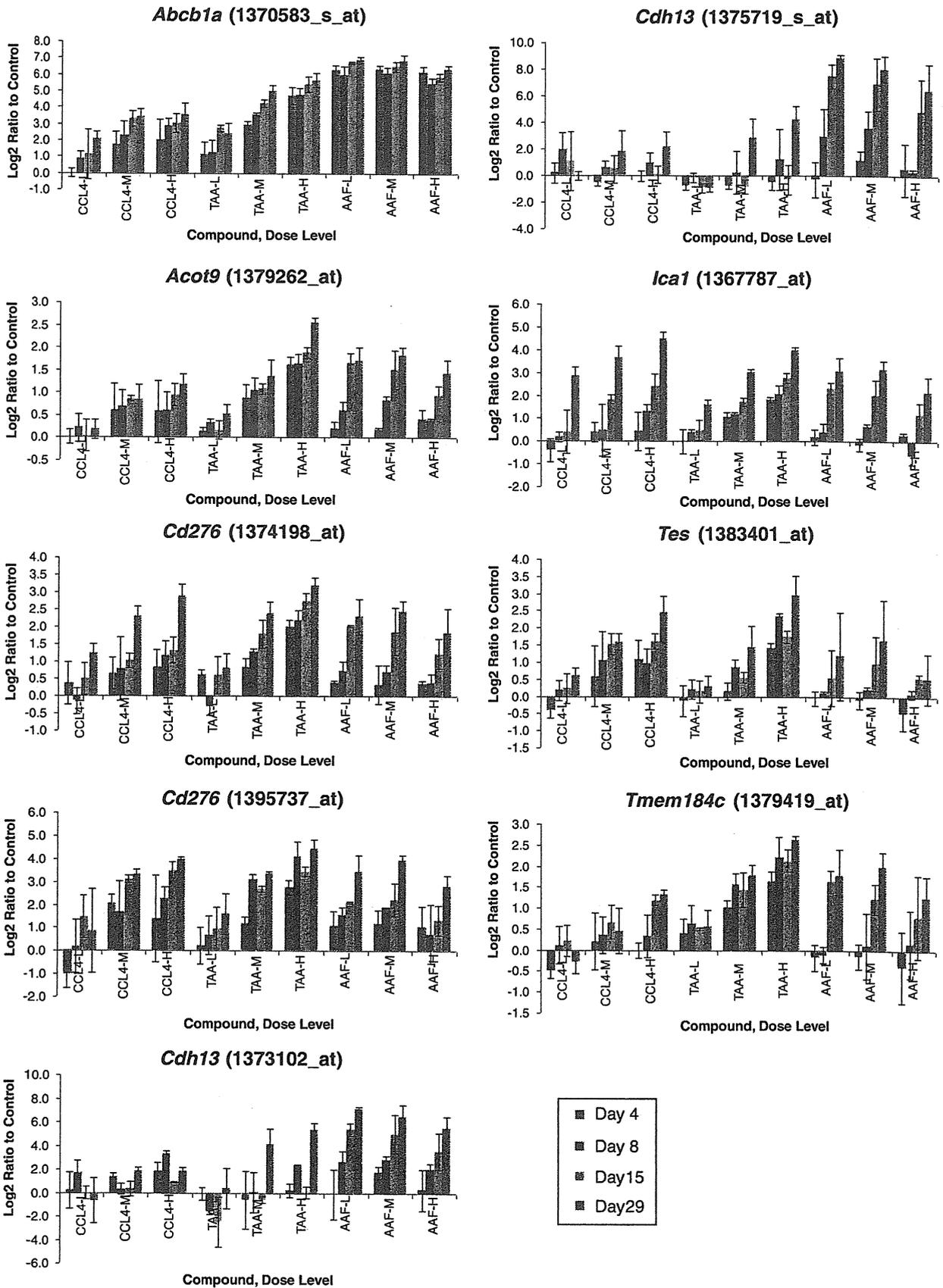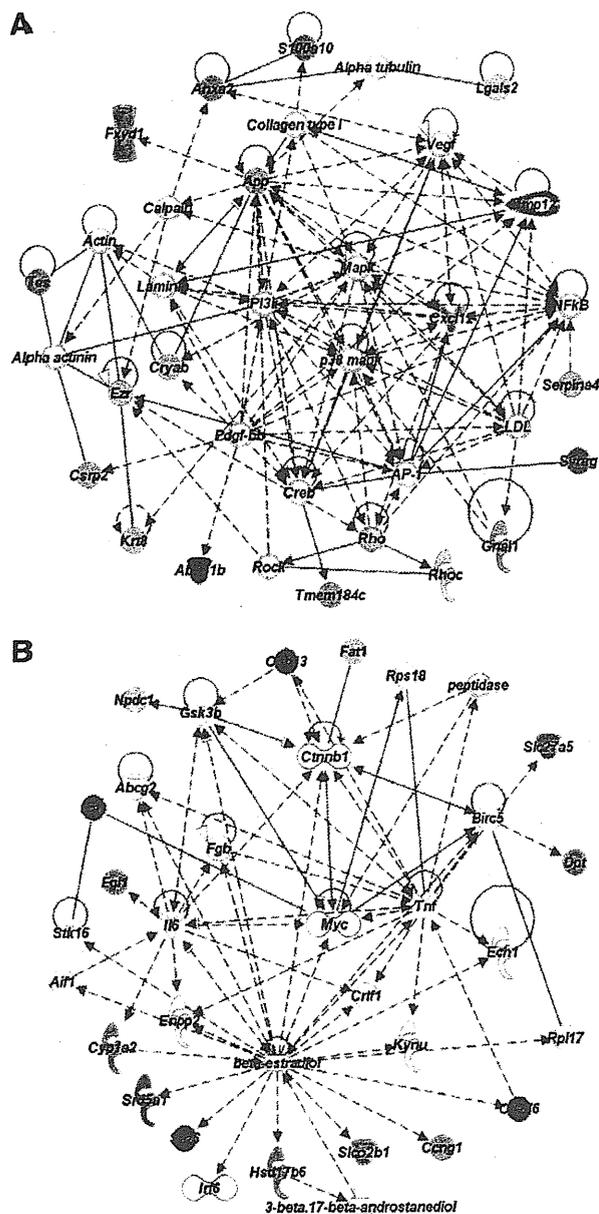
Fig. 3. Effect of dose and exposure duration on the expression of candidate biomarker genes for hepatocarcinogenesis. Log-ratio of base 2 to the means of the corresponding control groups with standard deviation are shown in 3 representative hepatocarcinogens, carbon tetrachloride, thioacetamide and acetamidofluorene.

A

B

Fig. 4. Molecular networks representing selected feature genes. p38 Mapk- (A) and Myc-centered interactomes (B) were selected as significant network in the selected 82 genes. Red and green represent molecules upregulated or downregulated in positive compounds compared to negative compounds, respectively. Molecules incorporated into the network are shown in white. Ellipse, square, triangle, trapezoid, lozenge and circle represent transcription regulator, cytokine, kinase, transporter, enzyme and other molecules, respectively. Arrows connecting molecules indicate one molecule acts on another, and lines indicate one molecule binds to another. Dashed arrows or lines indicate indirect interactions of 2 molecules.

Furthermore, unsupervised analysis, PCA, was also performed based on the expression profiles of 53 probes in all of the NEDO data set. The scores of the first and second principal components (PC1 and PC2) at 4, 15 and 29D were separately plotted in Fig. 5. Several genotoxic hepatocarcinogens, such as nitrosodiethylamine, N-nitrosomorpholine, 2-nitropropane, furan and N-nitrosopiperidine, which show positive findings in in vitro genotoxicity assays and in vivo carcinogenicity assays in rats, were separately plotted from non-hepatocarcinogens toward the PC1 direction. In addition, non-genotoxic hepatocarcinogens, methapyrilene, thioacetamide and carbon tetrachloride, which are

also included in our data set, showed clear separations in the PC1 direction. In almost all of these compounds, a distinct separation was observed in the data set after long-term repeated dosing.

Predictions using published biomarker genes. ROC curves of all models are plotted in Fig. 6. Although there was no large overlap of genes among the models, all models showed high prediction performance in our data set. Among all models, our new model achieved the best classification accuracy under the 5-fold cross-validation (Table 3).

Discussion

The goal of the present study was to develop a robust toxicogenomic model for the early assessment of potential hepatocarcinogenicity of chemicals based on gene expression changes stored in our toxicogenomics database, TG-GATEs. Carcinogenesis occurs via multiple mechanisms. Based on the properties of DNA damage, carcinogens are classified as genotoxins or non-genotoxins. Genotoxic carcinogens induce direct DNA damage, that is detectable by in vitro genotoxicity assays (e.g., Ames test and chromosomal aberration assays) and in vivo micronucleus test. Non-genotoxic carcinogens act through various modes of action that do not involve direct DNA damage (e.g., hepatocellular necrosis and regenerative proliferation, xenobiotic receptor agonists, peroxisome proliferation, or hormonal-mediated processes). Although it would be useful to detect the potential hepatocarcinogenicity of these different classes of non-genotoxic hepatocarcinogens by a single model, it is generally believed that a mechanism-based approach would be a promising strategy for robust toxicogenomic modeling since different classes of compounds generally show different gene expression profiles. In the present study, we have tried to establish reliable candidate gene biomarkers to assess the potential risks of hepatocarcinogenicity in the early stage of drug development with a particular focus on liver necrogenic compounds. Consequently, we have successfully identified robust biomarker genes specifically upregulated in necrogenic compounds with hepatocarcinogenicity. In contrast, no positive predictions were observed in the other class of non-genotoxic hepatocarcinogens involved in hepatic enzyme inducers, peroxisome proliferators and hormonal modulators. Therefore, these characteristics of prediction profile might help to elucidate the mechanisms of action involved in non-genotoxic hepatocarcinogenesis.

In contrast to our previous classifier (Uehara et al., 2008), our current model successfully achieved robust detection of non-genotoxic hepatocarcinogenicity of compounds by using hepatic gene expression data after 28 days of repeated dosing. There were some differences in the prediction properties of these 2 models. Namely, the current model enables us to detect robust gene expression changes possibly related to hepatocarcinogenic process following chronic doses in contrast to the previous model, which is useful to detect early temporal signals after a single dose as well as repeated doses; therefore, the combined use of both models for comprehensive judgment is thought to be the best strategy for sensitive and robust detection of hepatocarcinogenicity in short-term repeated dosing studies.

Interestingly, our current classifier as well as the previous classifier (data not shown) provided positive prediction results for not only non-genotoxic hepatocarcinogens but genotoxic hepatocarcinogens as well. This observation was further confirmed by the external test data set from the NEDO project (Matsumoto et al., 2009). It is generally believed that genotoxic compounds directly induce DNA damage by themselves or their reactive metabolites and that sufficient initiation is achieved after a single dose or short-term repeated doses due to the strong potency of their genotoxic stimulus. In contrast, non-genotoxic carcinogens generally require repeated dosing for sufficient initiation. Although the mechanisms involved in hepatocarcinogenic process are not the same between genotoxic and non-genotoxic hepatocarcinogens, our selected genes showed similar

A

B

C
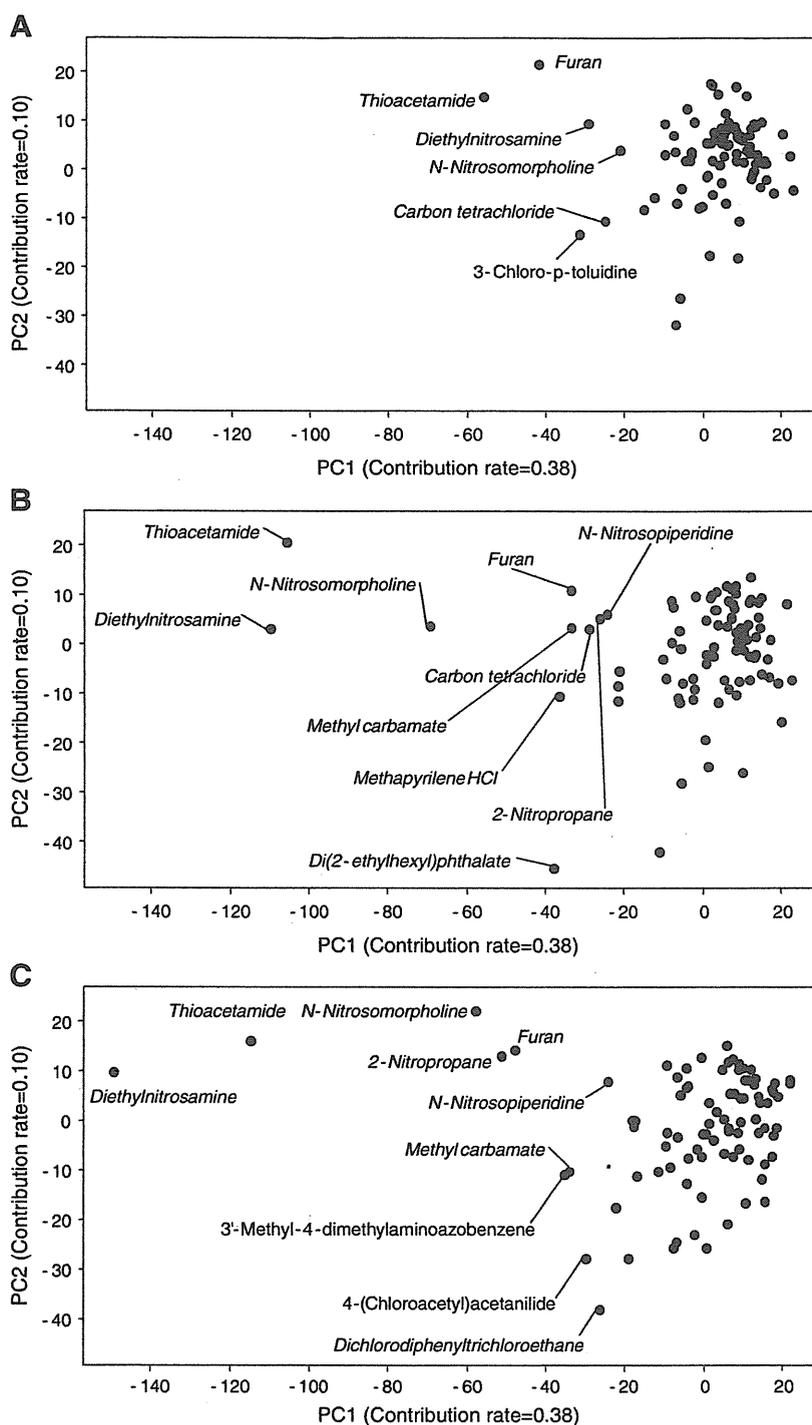
Fig. 5. Principal component analysis in independent validation. An external microarray data set obtained from the NEDO project was used for independent validation of our model. PCA was performed on all compounds by using the expression data at 3 different time points of 4 (A), 15 (B) and 29D (C). Italic font indicates hepatocarcinogens. Several hepatocarcinogens were separated from other chemicals including non-carcinogens toward the direction of PC1.

expression profiles after repeated dosing. Our results indicate that the expression profiles of our newly selected candidate biomarker genes might be common characteristics in the early stage of hepatocarcinogenic process, regardless of the type of carcinogens.

Among the test compounds, ethionamide and etambutol showed positive predictions in the current model, although there is no direct evidence supporting the hepatocarcinogenicity of these compounds in rodents. In our experiments, repeated doses of ethionamide caused

anisokaryosis of centrilobular hepatocytes. In the etambutol-treated liver, karyomegaly was observed in hepatocytes with distinct nucleoli. Although it is hard to conclude that these morphological changes in the nucleus of hepatocytes are directly connected to hepatocarcinogenicity, these morphological changes might be early indicative changes of hepatocarcinogenesis and reflect nuclear DNA damage caused by these compounds. Thus, gene expression changes after further long-term repeated dosing as well as a 2-year bioassay study would be of particular
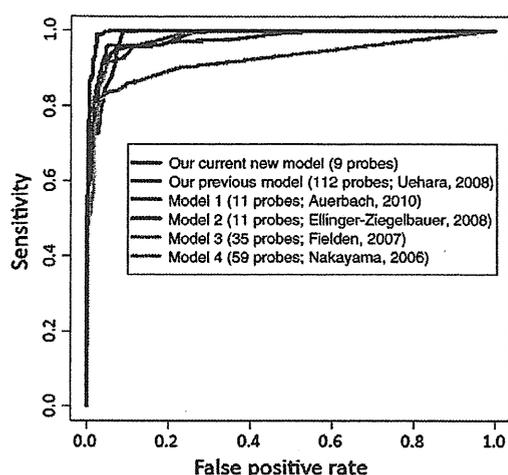
Fig. 6. Receiver operating characteristic analysis of previously published prediction models. ROC curves of prediction models are plotted. Our current new model: prediction results obtained by the best model consisting of 9 probes are shown. Model 1: a model consisting of 11 probes in the 90-day optimal-model, which showed the best prediction performance in our data set among their models, was used for this analysis. Model 2: among their models, commonly selected 101 probes were used for modeling. Models 3 and 4: 35 and 56 probes were used for modeling. Lists of genes used for each modeling are shown in Supplemental Table 5.

interest to determine the carcinogenic potentials of these compounds as well as to confirm the biological significance of these gene expression changes.

This study also provided the novel biological information that modulation of *p38 Mapk-* and *Myc-*centered networks is a common characteristic of gene expression in the early stage of hepatocarcinogenesis. There is increasing evidence that *p38 Mapk*, which is a stress-activated kinase, also participates in cell-cycle regulation, functioning as a suppressor of cell proliferation and tumorigenesis (Campbell et al., 2011). *p38* plays essential roles in modulating chronic inflammation-related cytokine expression, such as tumor necrosis factor α, interleukin 6 and 12, which might act as promoters of cancer growth and progression (Karin et al., 2006; Kumar et al., 2003; Naugler et al., 2007). *p38alpha-*deficient mice are susceptible to carcinogens, and *p38alpha* negatively regulates cell proliferation by antagonizing the *JNK-c-Jun* pathway in multiple cell types and in liver cancer development (Hui et al., 2007a, 2007b). Therefore, upregulation of genes involved in the *p38 Mapk-*network might be a tumor-suppressive action against DNA damage stimulus in the liver following repeated doses of carcinogens. Furthermore, these genes might be useful indicators in the early stage of hepatocarcinogenesis. The *Myc* oncogene is a transcription factor that plays an important role in the pathogenesis of hepatocellular carcinoma (Calvisi and Thorgeirsson, 2005; Thorgeirsson and Grisham, 2002). *Myc* overexpression is believed to exert its neoplastic function via several mechanisms, such as inducing autonomous cell proliferation and growth, blocking differentiation, and causing genomic destabilization (Dang, 1999; Felsher and Bishop, 1999; Grandori et al., 2000; Oster et al., 2002; Pelengaris et al., 2002). While the detailed biological significance should be determined in further studies, several genes

were up or downregulated in this network. Among the genes involved in these networks, there is increasing evidence supporting their involvement in carcinogenesis. *Cdh13* (T-cadherin), an atypical member of the cadherin family, is thought to affect cellular biological processes largely via its signaling properties. It is often downregulated in cancerous cells, which is associated with a poor prognosis in various human carcinomas (Andreeva and Kutuzov, 2010). It is also reported that *Cdh13* re-expression in cancer cell lines inhibits cell proliferation and invasiveness, increases susceptibility to apoptosis, and reduces tumor growth (Andreeva and Kutuzov, 2010). Approximately 50% of human hepatocellular carcinomas overexpress *Cdh13* (Riou et al., 2006). While *Cdh13* expression is restricted to endothelial cells from large blood vessels in normal livers, it is upregulated in sinusoidal endothelial cells in invasive hepatocellular carcinomas (Riou et al., 2006). B7-H3 (*Cd276* protein), a surface immunomodulatory glycoprotein, inhibits the functions of natural killer cells and T cells. Whereas the B7-H3 transcript is ubiquitously expressed in various types of human solid tumors as well as normal tissues, the B7-H3 protein is preferentially expressed only in tumor tissues (Xu et al., 2009). While additional biological studies would be required for further confirmation of the biological significance of the modulation of genes involved in these networks, these gene expression changes might play pivotal roles in the early stage of hepatocarcinogenesis.

It is well-known that carcinogenicity is dependent on total exposure levels of compounds; therefore, we have compared total exposure levels of compounds between our experimental conditions and the 2-year bioassay for a better understanding of the toxicological significance of positive results by our prediction model. Due to the abundant literature on carcinogenesis studies, we selected 2 representative hepatocarcinogens, nitrosodiethylamine and acetamidofluorene, for detailed discussions. In our experimental condition, positive predictions for the nitrosodiethylamine and acetamidofluorene groups were observed at cumulative total doses of 240 to 450 mg/kg for nitrosodiethylamine and 420 to 8400 mg/kg for acetamidofluorene. Williams et al. (2004) performed a series of dose-response investigations with these 2 compounds to characterize differences in hepatocarcinogenic effects with respect to exposures. In their two-stage hepatocarcinogenesis study of rats, nitrosodiethylamine was weekly dosed at 40.9 mg/kg for 10 weeks, and a cumulative nitrosodiethylamine dose of 409 mg/kg yielded an 80% hepatocellular tumor incidence followed by 4 weeks of recovery and 24 weeks of promotion with phenobarbital. In contrast, another study demonstrated that a cumulative dose of 409 mg/kg induced a 45% hepatic tumor incidence after 130 weeks of exposure in drinking water (Peto et al., 1991a, 1991b). Regarding acetamidofluorene, a 100% incidence of hepatic tumors was reported at cumulative doses of 282 mg/kg (3.4 mg/kg/day for 12 weeks) followed by 24 weeks of promotion with phenobarbital, while the dose of 1772 mg/kg (3.3 mg/kg/day for 76 weeks) was needed without the promotion (Williams et al., 1991). Although there were some differences in the susceptibility to tumor occurrence in these 2 reports, the cumulative doses of nitrosodiethylamine and acetamidofluorene in our present study were sufficiently above the cumulative dose needed for hepatocarcinogenesis in rats. Taken together, it might be reasonable to conclude that sufficient initiation effects have been achieved under our current dosing conditions of hepatocarcinogens with positive prediction. For building

Table 3
Prediction performance of previously published models.

Models	# of probes	AUC	Sensitivity	False positive rate	Specificity	Correct classification rate
Our current new model	9	0.996	0.990	0.033	0.967	0.978
Our previous model	112	0.983	0.997	0.087	0.913	0.955
Model 1 (Auerbach)	11	0.977	0.960	0.053	0.947	0.953
Model 2 (Ellinger)	101	0.980	0.937	0.057	0.943	0.940
Model 3 (Fielden)	35	0.923	0.830	0.040	0.960	0.895
Model 4(Nakayama)	56	0.982	0.917	0.053	0.947	0.932

models, we have anchored on available carcinogenicity results from 2-year rat bioassays of each compound. Our toxicogenomics project does not only focus on the carcinogenicity of compounds; therefore, SD rats were used as experimental animals in our project, while carcinogenicity tests generally use F344 rats. In addition, rats received orally or intravenously administered compound at the 1-month maximum tolerated doses. However, other dosing methods, generally in the diet or water, are also used for repeated exposure during the 2-year lifespan of rats. Therefore, differences in experimental conditions should be taken into account for a precise comparison. Since we hypothesized that expression changes in our feature genes might reflect the initiated condition of the liver following large doses of carcinogens for up to 1 month, we are conducting further biological studies by using a 2-step carcinogenicity study model in rats. Further data will be published in the near future.

In conclusion, we constructed a new toxicogenomic model for early prediction for both genotoxic and non-genotoxic hepatocarcinogens by using comprehensive gene expression data stored in our large-scale toxicogenomics database. The usefulness and robustness of our predictor were further confirmed in an independent validation data set obtained from a public database. Our toxicogenomic model might be useful for the prospective screening of hepatocarcinogenicity of compounds and prioritization of compounds for carcinogenicity testing. Furthermore, these characteristics of gene expression changes would help to elucidate the mechanisms of action involved in hepatocarcinogenesis.

Conflict of interest statement

The authors have declared no conflict of interest.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10. 1016/j.taap.2011.07.001.

References

Andreeva, A.V., Kutuzov, M.A., 2010. Cadherin 13 in cancer. Genes Chromosomes Cancer 49, 775–790.
Auerbach, S.S., Shah, R.R., Mav, D., Smith, C.S., Walker, N.J., Vallant, M.K., Boorman, G.A., Irwin, R.D., 2010. Predicting the hepatocarcinogenic potential of alkenylbenzene flavoring agents using toxicogenomics and machine learning. Toxicol. Appl. Pharmacol. 243, 300–314.
Battershill, J.M., 2005. Toxicogenomics: regulatory perspective on current position. Hum. Exp. Toxicol. 24, 35–40.
Calvisi, D.F., Thorgeirsson, S.S., 2005. Molecular mechanisms of hepatocarcinogenesis in transgenic mouse models of liver cancer. Toxicol. Pathol. 33, 181–184.
Campbell, J.S., Argast, G.M., Yuen, S.Y., Hayes, B., Fausto, N., 2011. Inactivation of p38 MAPK during liver regeneration. Int. J. Biochem. Cell Biol. 43, 180–188.
Chen, J.J., Tsai, C., Tzeng, S., Chen, C., 2007. Gene selection with multiple ordering criteria. BMC Bioinformatics 8, 74.
Dang, C.V., 1999. c-Myc target genes involved in cell growth, apoptosis, and metabolism. Mol. Cell. Biol. 19, 1–11.
Ellinger-Ziegelbauer, H., Gmuender, H., Bandenburg, A., Ahr, H.J., 2008. Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. Mutat. Res. 637, 23–39.
Felsher, D.W., Bishop, J.M., 1999. Transient excess of MYC activity can elicit genomic instability and tumorigenesis. Proc. Natl. Acad. Sci. U.S.A. 96, 3940–3944.
Fielden, M.R., Brennan, R., Gollub, J., 2007. A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. Toxicol. Sci. 99, 90–100.
Fielden, M.R., Nie, A., McMillian, M., Elangbam, C.S., Trela, B.A., Yang, Y., Dunn II, R.T., Dragan, Y., Fransson-Stehen, R., Bogdanffy, M., Adams, S.P., Foster, W.R., Chen, S.J., Rossi, P., Kasper, P., Jacobson-Kram, D., Tatsuoka, K.S., Wier, P.J., Gollub, J., Halbert, D.N., Roter, A., Young, J.K., Sina, J.F., Marlowe, J., Martus, H.J., Aubrecht, J., Olaharski, A.J., Roome, N., Nioi, P., Pardo, I., Snyder, R., Perry, R., Lord, P., Mattes, W., Car, B.D., Predictive Safety Testing Consortium, Carcinogenicity Working Group, 2008.

Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. Toxicol. Sci. 103, 28–34.
Gao, W., Mizukawa, Y., Nakatsu, N., Minowa, Y., Yamada, H., Ohno, Y., Urushidani, T., 2010. Mechanism-based biomarker gene sets for glutathione depletion-related hepatotoxicity in rats. Toxicol. Appl. Pharmacol. 247, 211–221.
Grandori, C., Cowley, S.M., James, L.P., Eisenman, R.N., 2000. The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu. Rev. Cell. Dev. Biol. 16, 653–699.
Heinloth, A.N., Irwin, R.D., Boorman, G.A., Nettesheim, P., Fannin, R.D., Sieber, S.O., Snell, M.L., Tucker, C.J., Li, L., Travlos, G.S., Vansant, G., Blackshear, P.E., Tennant, R.W., Cunningham, M.L., Paules, R.S., 2004. Gene expression profiling of rat livers reveals indicators of potential adverse effects. Toxicol. Sci. 80, 193–202.
Hirode, M., Ono, A., Miyagishima, T., Nagao, T., Ohno, Y., Urushidani, T., 2008. Gene expression profiling in rat liver treated with compounds inducing phospholipidosis. Toxicol. Appl. Pharmacol. 229, 290–299.
Hui, L., Bakiri, L., Stepniak, E., Wagner, E.F., 2007a. p38alpha: a suppressor of cell proliferation and tumorigenesis. Cell Cycle 6, 2429–2433.
Hui, L., Bakiri, L., Mairhorfer, A., Schweifer, N., Haslinger, C., Kenner, L., Komnenovic, V., Scheuch, H., Beug, H., Wagner, E.F., 2007b. p38alpha suppresses normal and cancer cell proliferation by antagonizing the JNK-c-Jun pathway. Nat. Genet. 39, 741–749.
Irwin, R.D., Boorman, G.A., Cunningham, M.L., Heinloth, A.N., Malarkey, D.E., Paules, R.S., 2004. Application of toxicogenomics to toxicology: basic concepts in the analysis of microarray data. Toxicol. Pathol. 32, 72–83.
Karin, M., Lawrence, T., Nizet, V., 2006. Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. Cell 124, 823–835.
Kiyosawa, N., Ando, Y., Manabe, S., Yamoto, T., 2009. Toxicogenomic biomarkers for liver toxicity. J. Toxicol. Pathol. 22, 35–52.
Kondo, C., Minowa, Y., Uehara, T., Okuno, Y., Nakatsu, N., Ono, A., Maruyama, T., Kato, I., Yamate, J., Yamada, H., Ohno, Y., Urushidani, T., 2009. Identification of genomic biomarkers for concurrent diagnosis of drug-induced renal tubular injury using a large-scale toxicogenomics database. Toxicology 265, 15–26.
Kramer, J.A., Curtiss, S.W., Kolaja, K.L., Alden, C.L., Blomme, E.A., Curtiss, W.C., Davila, J.C., Jackson, C.J., Bunch, R.T., 2004. Acute molecular markers of rodent hepatic carcinogenesis identified by transcription profiling. Chem. Res. Toxicol. 17, 463–470.
Kumar, S., Boehm, J., Lee, J.C., 2003. p38 MAP kinases: key signalling molecules as therapeutic targets for inflammatory diseases. Nat. Rev. Drug Discov. 2, 717–726.
Matsumoto, H., Yakabe, Y., Saito, K., Sumida, K., Sekijima, M., Nakayama, K., Miyaura, H., Saito, F., Otsuka, M., Shirai, T., 2009. Discrimination of carcinogens by hepatic transcript profiling in rats following 28-day administration. Cancer Inform. 7, 253–269.
Nakayama, K., Kawano, Y., Kawakami, Y., Moriwaki, N., Sekijima, M., Otsuka, M., Yakabe, Y., Miyaura, H., Saito, K., Sumida, K., Shirai, T., 2006. 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.
Naugler, W.E., Sakurai, T., Kim, S., Maeda, S., Kim, K., Elsharkawy, A.M., Karin, M., 2007. Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science 317, 121–124.
Nie, A.Y., McMillian, M., Parker, J.B., Leone, A., Bryant, S., Yieh, L., Bittner, A., Nelson, J., Carmen, A., Wan, J., Lord, P.G., 2006. Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. Mol. Carcinog. 45, 914–933.
Oster, S.K., Ho, C.S., Soucie, E.L., Penn, L.Z., 2002. The myc oncogene: Marvelously Complex. Adv. Cancer Res. 84, 81–154.
Pelengaris, S., Khan, M., Evan, G., 2002. c-MYC: more than just a matter of life and death. Nat. Rev. Cancer 2, 764–776.
Peto, R., Gray, R., Brantom, P., Grasso, P., 1991a. Effects on 4080 rats of chronic ingestion of N-nitrosodiethylamine or N-nitrosodimethylamine: a detailed dose-response study. Cancer Res. 51, 6415–6451.
Peto, R., Gray, R., Brantom, P., Grasso, P., 1991b. Dose and time relationships for tumor induction in the liver and esophagus of 4080 inbred rats by chronic ingestion of N-nitrosodiethylamine or N-nitrosodimethylamine. Cancer Res. 51, 6452–6469.
Riou, P., Saffroy, R., Chenailler, C., Franc, B., Gentile, C., Rubinstein, E., Resink, T., Debuire, B., Piatier-Tonneau, D., Lemoine, A., 2006. Expression of T-cadherin in tumor cells influences invasive potential of human hepatocellular carcinoma. FASEB J. 20, 2291–2301.
Searfoss, G.H., Ryan, T.P., Jolly, R.A., 2005. The role of transcriptome analysis in preclinical toxicology. Curr. Mol. Med. 5, 53–64.
Thorgeirsson, S.S., Grisham, J.W., 2002. Molecular pathogenesis of human hepatocellular carcinoma. Nat. Genet. 31, 339–346.
Uehara, T., Hirode, M., Ono, A., Kiyosawa, N., Omura, K., Shimizu, T., Mizukawa, Y., Miyagishima, T., Nagao, T., 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., Urushidani, T., 2010. The Japanese toxicogenomics project: application of toxicogenomics. Mol. Nutr. Food Res. 54, 218–227.
Urushidani, T., 2010. Toxicogenomics project and drug safety evaluation. Nippon Yakurigaku Zasshi 136, 46–49.
Williams, G.M., Tanaka, T., Maruyama, H., Maeura, Y., Weisburger, J.H., Zang, E., 1991. Modulation by butylated hydroxytoluene of liver and bladder carcinogenesis induced by chronic low level exposure to 2-acetylaminofluorene. Cancer Res. 51, 6224–6230.
Williams, G.M., Iatropoulous, M.J., Jeffrey, A.M., 2004. Thresholds for the effects of 2-acetylaminofluorene in rat liver. Toxicol. Pathol. 32, 85–91.
Xu, H., Cheung, I.Y., Guo, H.F., Cheung, N.K., 2009. MicroRNA miR-29 modulates expression of immunoinhibitory molecule B7-H3: potential implications for immune based therapy of human solid tumors. Cancer Res. 69, 6275–6281.