Tissue	UP	DOWN
HC	29	13
T	22	15
G	5	110
L	122	246
K	2	318
SI	707	1
Y	10	4

他の条件(Wild-RIFやSXRki-PCNなど)での変動の程度を調べ、さらに絞り込む必要がある

Figure14　SXR 活性化パターン候補プローブセット数

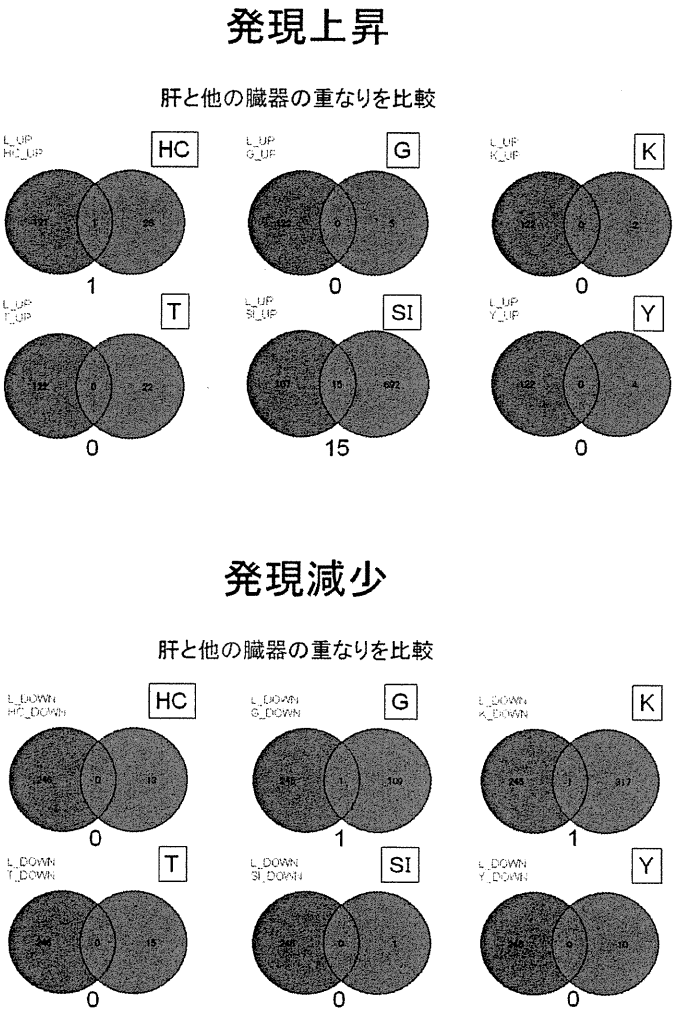

Figure15　SXR 活性化パターン候補プローブセットの臓器間の重なり(肝との比較)

別添４

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書　籍　名	出版社名	出版地	出版年	ページ
Yasushi Okuno, Yohsuke Minowa Hiroshi Yamada, Yasuo Ohno and Tetsuro Urushidani.	In Silico Toxicology Prediction Using Toxicogenomics Data	Daniel A. Casciano, Saura C. Sahu	“Handbook of Systems Toxicology”	John Wiley & Sons	USA	2011	591-598
漆谷 徹郎	トキシコゲノミクス	金子周一、堀池靖浩	バイオチップ実用化ハンドブック	NTS	東京	2010	268-274
T. Urushidani.	Prediction of Heptatotoxicity Based on the Toxicogenomics Database.	S. C. Sahu	"Heptatotoxicity: from Genomics to in vitro and in vivo Models"	John Wiley & Sons	USA	2008	507-529

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Sumida K, Igarashi Y, Toritsuka N, Matsushita T, Abe-Tomizawa K, Aoki M, Urushidani T, Yamada H, Ohno Y.	Effects of DMSO on gene expression in human and rat hepatocytes.	Hum. Exp. Toxicol.	30(10)	1701-1709	2011
Uehara T, Minowa Y, Morikawa Y, Kondo C, Maruyama T, Kato I, Nakatsu N, Igarashi Y, Ono A, Hayashi H, Mitsumori K, Yamada H, Ohno Y, Urushidani T.	Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database.	Toxicol Appl Pharmacol.	255(3)	297-306	2011

Low Y, Uehara T, Minowa Y, Yamada H, Ohno Y, Urushidani T, Sedykh A, Muratov E, Kuz'min V, Fourches D, Zhu H, Rusyn I, Tropsha A.	Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches.	Chem Res Toxicol.	24(8)	1251-1262	2011
Weihua Gao, Yumiko Mizukawa, Noriyuki Nakatsu, Yosuke Minowa, Hiroshi Yamada, Yasuo Ohno and Tetsuro Urushidani.	Mechanism-based biomarker gene sets for glutathione depletion-related hepatotoxicity in rats.	J. Toxicol. Appl. Pharmacol	247(3)	211-221	2010
Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y, Urushidani T.	The Japanese toxicogenomics project: Application of toxicogenomics.	Mol. Nutr. Food Res.	54(2)	218-227	2010
漆谷徹郎	トランスレーショナルリサーチ④　トキシコゲノミクスプロジェクトと安全性試験.	日本薬理学会雑誌	136 (1)	46-49	2010
大野泰雄	マイクロドーズ臨床試験に必要な非臨床試験データ	臨床薬理	41	9-16	2010
M. Hirode, A. Horinouchi, T. Uehara, A. Ono, T. Miyagishima, H. Yamada, T. Nagao, Y. Ohno, T. Urushidani.	Gene expression profiling in rat liver treated with compounds inducing elevation of bilirubin.	Human Exp. Toxicol.	28	231-244	2009
Mitsuhiro Hirode, Ko Omura, Naoki Kiyosawa, Takeki Uehara, Toshinobu Shimuzu, Atsushi Ono, Toshikazu Miyagishima, Taku Nagao, Yasuo Ohno and Tetsuro Urushidani.	Gene expression profiling in rat liver treated with various hepatotoxic-compounds inducing coagulopathy.	J. Toxicol. Sci.	34	281-293	2009

Chiaki Kondo, Yosuke Minowa, Takeki Uehara, Yasushi Okuno, Noriyuki Nakatsu, Atsushi Ono, Toshiyuki Maruyama, Ikuo Kato, Jyoji Yamate, Hiroshi Yamada, Yasuo Ohno and Tetsuro Urushidani.	Identification of genomic biomarkers for concurrent diagnosis of drug-induced renal tubular injury using a large-scale toxicogenomics database.	Toxicology	265	15-26	2009
漆谷徹郎	トキシコゲノミクス	日薬理誌	133	112-114	2009
A. Sanbuissho, M. Yoshida, S. Hisada, F. Sagami, S. Kudo, T. Kumazawa, M. Ube, S. Komatsu, Y. Ohno.	Collaborative work on evaluation of ovarian toxicity by repeated-dose and fertility studies in female rats.	J. Toxicol Sci.	34	SP1-SP22	2009
M. Yoshida, A. Sanbuissho, S. Hisada, M. Takahashi, Y. Ohno, A. Nishikawa.	Morphological characterization of the ovary under normal cycling in rats and its viewpoints of ovarian toxicity detection.	J. Toxicol Sci	34	SP189-SP197	2009
Takeki Uehara, Atsushi Ono, Mitsuhiro Hirode, Naoki Kiyosawa, Ko Omura, Toshinobu Shimizu, Yumiko Mizukawa, Toshikazu Miyagishima, Taku Nagao and Tetsuro Urushidani.	A Toxicogenomics approach for early assessment of potential non-genotoxic hepatocarcinogenicity of chemicals in rats	Toxicology	250	15-26	2008
T. Uehara, N. Kiyosawa, M. Hirode, K. Omura, T. Shimizu, A.Ono, Y. Mizukawa, T.Miyagishima, T.Nagao, T. Urushidani.	Gene Expression Profiling of Methapyrilene-Induced Hepatotoxicity in Rat.	J. Toxicol. Sci.	33	37-50	2008

T. Uehara, N. Kiyosawa, T. Shimizu, K. Omura, M. Hirode, T. Imazawa, Y. Mizukawa, A. Ono, T. Miyagishima, T. Nagao, T. Urushidani..	Species Differences in Coumarin-Induced Hepatotoxicity as an Example of How Toxicogenomics Help Assessing Risks for Human.	Hum. Exp. Toxicol.	27	23-35	2008
M. Hirode, A. Ono, T. Miyagishima, T. Nagao, Y. Ohno, and T. Urushidani.	Gene expression profiling in rat liver treated with compounds inducing phospholipidosis.:	Toxicol Appl Pharmacol.	229(3)	290-299	2008
S. Sakai, R. Matsuda, R. Adachi, H. Akiyama, T. Maitani, Y. Ohno, M. Oka, A. Abe, K. Seiki, H. Oda, K. Shiomi, A. Urisu.	Interlaboratory evaluation of two enzyme-linked immunosorbent assay kits for the determination of crustacean protein in processed food.	J. AOAC International	91	123-129	2008
T. Oguchi, M. Onishi, Y. Chikagawa, Y. Minegichi, T. Kodama, H. Akiyama, Y. Ohno, S. Futo, A. Hino, S. Furui, K. Kitta.	Development of event-specific quantitation method for GA21 maize, which is a GM event without CaMV35S promoter.	J. Food Hyg. Soc. Japan	49	16-22	2008
T. Ashikaga, H. Sakaguchi, K. Okamoto, M. Mizuno, J. Sato, T. Yamada, M. Yoshida, N. Ota, S. Hasegawa, T. Kodama, Y. Okamoto, H. Kuwahara, N. Kosaka, S. Sono, and Y. Ohno.	Assessment of the human cell line activation test (h-CLAT) for skin sensitization; Results of the first Japanese inter-laboratory study.	AATEX	13	27-35	2008

H. Kojima, T. Ando, K. Inagaki, M. Ohhira, T. Kosaka, Y. Nakamura, H. Torishima, N. Morikawa, J. Kanno, M. Kuboki, M. Genno, M. Nokata, T. Harada, T. Morimoto, I. Yoshimura, Y. Ohno:	Validation of human skin models for skin corrosivity tests in Japan	AATEX	13	36-44	2008
漆谷徹郎	医薬品安全性研究の動向～マイクロドーズ試験を含めて～.	新薬展望	44	229-234	2008
K. Omura, N. Kiyosawa, T. Uehara, M. Hirode, T. Shimizu, T.Miyagishima, A. Ono, T. Nagao, T. Urushidani.	Gene Expression Profiling of Rat Liver Treated with Serum Triglyceride-Decreasing Compounds.	J. Toxicol. Sci.	32	387-399	2007
N. Kiyosawa, T. Uehara, W. Gao, K. Omura, M. Hirode, T. Shimizu, Y. Mizukawa, A. Ono, T. Miyagishima, T. Nagao, T.Urushidani.	Identification of Glutathione Depletion-Responsive Genes Using Phorone-Treated Rat Liver.	J. Toxicol. Sci.	32	469-486	2007
漆谷徹郎	レクチャーノート　トキシコゲノミクスプロジェクト（2）	Drug Metab. Pharmacokinet.	22(2)	13-15	2007
大野泰雄	日本薬理学会の奨める動物実験– 苦痛の評価と軽減– 「はじめに」および日本薬理学会の新動物実験指針	日本薬理学雑誌	129	5-9	2007
Y. Shinozaki, Y. Sato, S. Koizumi, Y. Ohno, T. Nagao and K. Inoue	Retinoic acids acting through retinoid receptors protect hippocampal neurons from oxygen-glucose deprivation-mediated cell death by inhibition of JNK and p38 mitogen-activated protein kinase.	Neuroscience	147(1)	153-63	2007

Sato K, Akaishi T, Matsuki N, Ohno Y, Nakazawa K.	beta-Estradiol induces synaptogenesis in the hippocampus by enhancing brain-derived neurotrophic factor release from dentate gyrus granule cells.	Brain Res	1150	108-120	2007
大野泰雄	動物福祉と動物実験代替法への考慮の必要性について	Biophilia	3	4-5	2007
大野泰雄、小野俊介	マイクロドーズ試験ガイダンスの検討について	医薬品研究	38	623-638	2007
大野泰雄	動物実験代替法の国際動向	Fragrance Journal	10	20-28	2007
内藤真策、古田盛、吉田武美、北田光一、笛木修、海野修、大野泰雄、小野寺博志、河村信之、黒川美佐男、佐上文郎、篠田和俊、中澤隆弘、山崎恒義	医薬品開発における代謝物の安全性評価についての考え方	医薬品研究	38	495-498	2007
Naito S., Furuta S., Yoshida T., Kitada K., Fueki O., Unno T., Ohno Y., Onodera H., Kawamura N., Kurokawa M., Sagami F., Shinoda K., Nakazawa T, Yamazaki T.	Current opinon: Safety evaluation of drug metabolites in development of pharmaceuticals.	J. Toxicol. Sci.	32	329-341	2007
Seiki K., Oda H., Yoshioka H., Sakai S., Urisu A., Akiyama H., Ohno Y.	A reliable and sensitive immunoassay for the determination of crustacean protein in processed foods.	J. Agric. Food Chem.	55(23)	9345-50	2007
Arase S, Ishii K, Igarashi K, Aisaki K, Yoshio Y, Matsushima A, Shimohigashi Y, Arima K, Kanno J, Sugimura Y.,	Endocrine Disrupter Bisphenol A Increases In Situ Estrogen Production in the Mouse Urogenital Sinus.	Biol Reprod.	84 (4)	734-42	2011

Yoshida T, Sekine T, Aisaki KI, Mikami T, Kanno J, Okayasu I.,	CITED2 is activated in ulcerative colitis and induces p53-dependent apoptosis in response to butyric acid.	J. Gastroentero l	46(3)	339-49.	2011
Katsuhide Igarashi, Satoshi Kitajima, Ken-ichi Aisaki, Kentaro Tanemura, Yuhji Taquahashi, Noriko Moriyama, Eriko Ikeno, Nae Matsuda, Yumiko Saga, Bruce Blumberg, and Jun Kanno	Development of Humanized Steroid and Xenobiotic Receptor Mouse by homologous knock-in of the human Steroid and Xenobiotic Receptor Ligand Binding Domain sequence.	J Toxicol Sci.	Vol37, No2	373-380	2012
Fujiki R, Hashiba W, Sekine H, Yokoyama A, Chikanishi T, Ito S, Imai Y, Kim J, He HH, Igarashi K, Kanno J, Ohtake F, Kitagawa H, Roeder RG, Brown M, Kato S.	GlcNAcylation of histone H2B facilitates its monoubiquitination.	Nature	VOL480	557-560	2011
Matsukura H, Aisaki K, Igarashi K, Matsushima Y, Kanno J, Muramatsu M, Sudo K, Sato N.	Genistein promotes DNA demethylation of the steroidogenic factor 1 (SF-1) promoter in endometrial stromal cells.	Biochem Biophys Res Commun.	412	366-372	2011
Suzuki A, Igarashi K, Aisaki KI, Kanno J, Saga Y.	NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs.	Proc Natl Acad Sci U S A.	107 (8)	3594-9	2010
Matsunaga N, Kanno J, Hamada C, Yoshimura I.	An experimental design for judging synergism on consideration to endocrine disruptor animal experiments.	Environmetri cs	20	1-13.	2009

Ishimaru N, Takagi A, Kohashi M, Yamada A, Arakaki R, Kanno J, Hayashi Y.	Neonatal exposure to low-dose 2,3,7,8-tetrachlorodibenzo-p-dioxin causes autoimmunity due to the disruption of T cell tolerance.	J Immunol.	182(10)	6576-6586.	2009
Upham BL, Park JS, Babica P, Sovadinova I, Rummel AM, Trosko JE, Hirose A, Hasegawa R, Kanno J, Sai K.	Structure-activity-dependent regulation of cell communication by perfluorinated fatty acids using in vivo and in vitro model systems.	Environ Health Perspect.	117(4)	545-51	2009
Nakamura T, Imai Y, Matsumoto T, Sato S, Takeuchi K, Igarashi K, Harada Y, Azuma Y, Krust A, Yamamoto Y, Nishina H, Takeda S, Takayanagi H, Metzger D, Kanno J, Takaoka K, Martin TJ, Chambon P, Kato S.	Estrogen prevents bone loss via estrogen receptor alpha and induction of Fas ligand in osteoclasts.	Cell	130	811-823	2007
Nakatsu N, Nakamura T, Yamazaki K, Sadahiro S, Makuuchi H, Kanno J, Yamori T.	Evaluation of action mechanisms of toxic chemicals using JFCR39, a panel of human cancer cell lines.	Mol Pharmacol.	72(5)	1171-80	2007
Aisaki K, Aizawa S, Fujii H, Kanno J, Kanno H.	Glycolytic inhibition by mutation of pyruvate kinase gene increases oxidative stress and causes apoptosis of a pyruvate kinase deficient cell line.	Exp Hematol.	35	1190-1200	2007

Kato Y, Ikushiro S, Takiguchi R, Haraguchi K, Koga N, Uchida S, Sakaki T, Yamada S, Kanno J, Degawa M.	A novel mechanism for polychlorinated biphenyl-induced decrease in serum thyroxine level in rats.	Drug Metab Dispos	35	1949-1955	2007
菅野 純，北嶋 聡，相崎健一，五十嵐勝秀，中津則之，高木篤也，小川幸男，児玉幸夫	Percellome Projectによる毒性トランスクリプトミクスの新しい試み	細胞工学	26	71-77	2007
菅野 純，相崎健一，五十嵐勝秀，北嶋 聡，中津則之，児玉幸夫，高木篤也	トキシコゲノミクスの新展開：Percellomeプロジェクトによる2,3,7,8-TCDD－2,3,7,8-TCDF比較	細胞工学	26	1391-1396	2007

36

In Silico Toxicology Prediction Using Toxicogenomics Data

Yasushi Okuno[1], Yohsuke Minowa[2], Hiroshi Yamada[2], Yasuo Ohno[2] and Tetsuro Urushidani[2,3]

[1]*Department of Systems Bioscience for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Japan,* [2]*Toxicogenomics-Informatics Project, National Institute of Biomedical Innovation, Ibaraki, Osaka, Japan* and [3]*Department of Pathophysiology, Faculty of Pharmaceutical Sciences, Doshisha Women's College of Liberal Arts, Kodo, Kyoto, Japan*

1 INTRODUCTION

Toxicogenomics holds the promise of unprecedented advances in two broad, overlapping fields: mechanistic or investigative toxicology, and predictive toxicology. Mechanistic or investigative toxicology is the area of toxicology focused on biological responses to drug toxicity, and provides important perspectives on risk assessments of various compounds. In contrast, predictive toxicology focuses on identifying the potential toxicity of compounds (Cunningham and Lehman-McKeeman, 2005; Boverhof and Zacharewski, 2006). In previous studies, toxicogenomics has contributed to in-depth investigations of molecular mechanisms and the modes of toxin/chemical/environmental stressor action that was difficult to be achieved by conventional toxicological approaches. At the same time, expanding toxicogenomic data has promoted a valuable platform for the establishment of biomarkers to predict a compound's toxicity.

The progress of toxicogenomics has been supported by DNA microarray technology, a powerful tool for directly monitoring patterns of cellular perturbations through the identification and quantification of global shifts in gene expression resulting from pathological alterations within cells and tissues. Microarrays provide a large amount of transcriptional expression data for thousands of individual genes under various experimental conditions. Bioinformatics technologies can determine which genes are meaningful, facilitating the analysis of huge pools of toxicogenomics data in mechanistic and predictive toxicology. Meaningful genes are referred to as "signature genes" with characteristic gene expression profiles for exposure or toxicological response to specific classes of toxic compounds. This chapter is devoted to computational approaches for the data mining of biomarker genes from toxicogenomics data, leading to toxicity prediction.

2 MICROARRAY DATA FOR TOXICOGENOMICS

Microarray techniques have been applied not only to toxicogenomics but also to various clinical purposes, such as disease classification and patient diagnostics. These applications have revealed the challenges of achieving reproducibility and stability of outcomes with microarray technologies caused by cross-platform chips and differing experimental

Handbook of Systems Toxicology. Edited by Daniel A. Casciano and Saura C. Sahu.
 ISBN 978-0-470-68401-6.

conditions. Therefore, data sets for constructing reliable prediction models must be obtained with the same platform and experimental conditions. For this purpose, many public databases for toxicologically relevant microarray data have been developed, including the Comparative Toxicogenomics Database (CTD) (Mattingly *et al.*, 2006), Environment, Drugs and Gene Expression database (EDGE) (Hayes *et al.*, 2005), and Chemical Effects in Biological System (CEBS) knowledgebase (Waters *et al.*, 2003), among others.

In 2002, we began the toxicogenomics project (TGP), a public–private collaborative project of the National Institute of Health Sciences, the National Institute of Biomedical Innovation, and 15 pharmaceutical companies in Japan (Urushidani and Nagao, 2005). With an emphasis on the uniformity of data quality, TGP has generated a large-scale toxicology database of transcriptomes intended to predict the toxicity of new chemical entities in the early stages of drug development. Drug reactions such as efficacy and toxicity are associated with the dosage and time course after treatment, and so precisely monitoring drug reactions requires multiple dose- and time-dependent experiments for each drug. Thus far, about 150 chemicals, primarily medicinal compounds, have been selected for the database. Over 27 000 gene expression profiles have been compiled for multiple doses and times in rat livers and kidneys, as well as rat and human hepatocytes, through comprehensive analysis using the Affymetrix GeneChip® Affymetrix, Inc., Santa Clara, CA, USA. These gene expression profiles, conjugated with histopathological changes, blood biochemical examination results, and the other phenotypic profiles, are stored in our database with a web-based tool for statistical analysis, genomics-assisted toxicity evaluation system developed by the toxicogenomics project in Japan (TG-GATEs). Thirteen of the 150 chemicals were typical nephrotoxicants or drugs showing clinical side-effects (e.g., cisplatin, carboplatin, gentamicin, vancomycin, phenacetin, and bucetin), and 20 chemicals exhibited nephrotoxicity in addition to hepatotoxicity (e.g., phenylbutazone, ethionine, and indomethacin).

3 IDENTIFICATION OF BIOMARKER GENES WITH TOXICOGENOMICS DATA

The first step from toxicogenomics to mechanistic and predictive toxicology is the identification of an individual gene or a cluster of genes detective or predictive of certain types of toxicity; these "signature genes" are employed as biomarkers. A biomarker is defined by the International Programme on Chemical Safety (IPCS) of the WHO as any substance or its product, structure, or process that can be measured in the body and that can influence or predict the incidence of disease outcome. The ideal biomarker provides a sensitive, informative, and reproducible indicator of potential adverse effects at times or doses preceding overt tissue damage, toxicity, or disease initiation. The discovery and validation of biomarkers is useful for application in high-throughput experimental systems to characterize target organ effects and to detect specific toxicity end-points in the early steps of a compound's development. The identification and utilization of biomarkers through toxicogenomics have several further applications: from current use in pre-clinical toxicology to risk characterization and risk assessment of chemicals; from early clinical stages of drug development to the later stages; and even into daily clinical use in diagnostics, disease classification, and therapeutic monitoring.

4 GENE SELECTION FROM MICROARRAY DATA

Identifying biomarker genes in huge sets of microarray data is referred to as the gene selection problem. In selecting genes from a microarray with good separation between toxic and non-toxic drug-treated samples, one seeks the significant genes that are affected by the adverse drug effects, or even those that caused the adverse reaction. This is a key step toward understanding mechanistic and predictive toxicology through the underlying biological process.

Gene selection is also relevant in the classification problems in machine learning, in which the class of toxic response (including non-toxic responses) of a sample (e.g., drug-treated organ or tissue) is determined by a classifier. A sample is represented as a feature vector $\mathbf{x}$. Each dimension in the feature vector $\mathbf{x}$ holds the expression value of a particular gene, which is obtained from a DNA microarray experiment. The classifier is constructed by inputting N feature vectors (called training data) with known toxic response outcomes into the machine-learning algorithms. However, because of the low number of samples (N: the number of the feature vectors)

and the high number of observed genes [dim(**x**): dimension of each feature vector], using all genes to classify the samples into good and bad outcomes incurs a high risk of over-training (or over-fitting). Over-training, in this case, means including the noise in the data, which may increase the generalization error (the error rate of the resulting predictor on samples that were not used during the training phase). This may yield a result that achieves high accuracy levels for the training data, but does not generalize to new data. The underlying problem is that if the sample size is much smaller than the number of genes, then one can distinguish different types of toxicity based on the noise present in these measurements, rather than on distinct biological characteristics of their gene expression levels. One approach to overcome this serious problem is to reduce the number of genes by removing irrelevant and redundant features, a method known as "feature selection" in computer science.

5 FEATURE SELECTION ALGORISMS

Feature selection methods can be broadly categorized into the filter model and the wrapper model. The filter model separates feature selection steps from the machine-learning algorithm, and relies on the general characteristics of the training data to select features (Figure 1a). The wrapper model uses the predictive accuracy of a pre-determined learning algorithm to determine the quality of a selected subset (Figure 1b).

5.1 Filter Methods

Traditional methods in gene selection fall within the filter model. Filter approaches remove irrelevant features based on a ranking of all the genes according to their individual relevance or discriminative power to the target class. Filter approaches generally follow three simple steps (Figure 1a), they: (a) rank all the genes in the microarray data using a filter method, (b) choose the top-ranked $n-1$ genes as the best feature gene subset, and (c) construct a classifier in learning algorithm using the selected $n-1$ genes. The filter-type feature selection (steps 1 and 2) is independent of the classifier-learning algorithm (step 3). One major problem with the filter model is the selection of a threshold for discarding irrelevant features. All the features are given a score

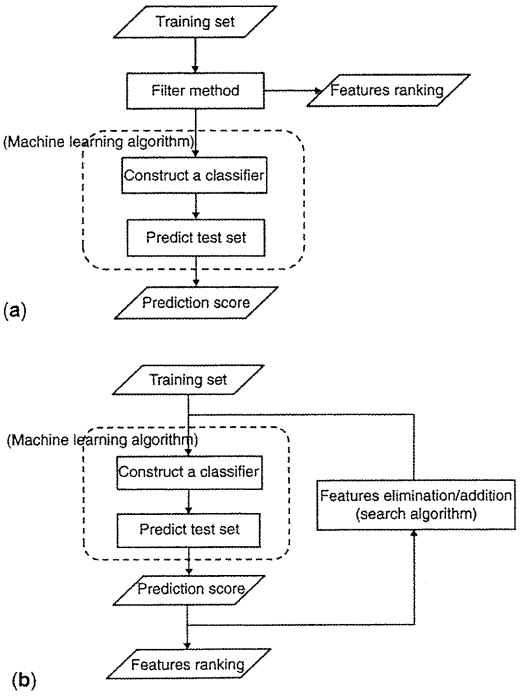

Figure 1. Calculation procedures of filter method (a) and wrapper method (b).

by the filter algorithm, but determining the optimal threshold "n" for the data is difficult. We must choose the threshold "n" using some appropriate criteria.

Filter methods are selected to produce the most relevant possible ranking. Many methods have been developed based on statistical tests. They are:

1. t-Statistic: This criterion measures the worth of a feature gene by computing the value of the t-statistic between toxic and non-toxic classes (Thomas *et al.*, 2001; Tsai, Chen and Chen, 2003).
2. ReliefF: This is a feature-weighting algorithm sensitive to feature interactions. The key idea of ReliefF is to rate features according to how well their values distinguish among instances of different classes and how well they cluster instances of the same class (Kononenko, 1994). To this end, ReliefF repeatedly chooses a single instance at random from the data, and then locates the nearest instances of the same class and the nearest instances pertaining to different

classes. The feature values of these instances are used to update the scores for each feature.

3. Correlation-based feature selection (CFS): This method evaluates a subset of features by considering the individual predictive ability of each feature as well as the degree of redundancy among them (Hall, 2000),

$$\mathrm{CFS} = \frac{k\bar{\tau}_{cf} f}{\sqrt{k + k(k-1)\bar{\tau}_{ff}}} \quad (1)$$

where CFS is the score of a feature subset containing k features, $\bar{\tau}_{cf}$ is the average of the correlation coefficients between features and classes, and $\bar{\tau}_{ff}$ is the average intercorrelation among features of the subset.

5.2 Wrapper Methods

Wrapper approaches to gene selection evaluate the prediction performance of a learning machine trained for each given feature subset using a search algorithm to explore possible combinations of features. The strategy of wrapper approaches is composed of four steps (Figure 1b):

1. Choose a machine-learning algorithm to evaluate the prediction score of a feature subset. Many machine-learning algorithms have been developed, including the Naïve Bayes classifier, linear discriminant methods, support vector machines (SVM) (Vapnik, 1998), and artificial neural networks (Bishop, 1995). Cross-validation (CV) is often used for scoring prediction performance.
2. Choose a search algorithm.
3. Perform the search algorithm and note the best subset encountered.
4. Output the encountered subset with the best score.

Search algorithms compare all possible feature combinations, which are usually too numerous to be exhaustively explored. Of these search methods, greedy methods (forward selection or backward elimination) are the most popular. At each round of CV, forward selection adds the best feature; backward elimination deletes the worst feature.

5.3 Recursive Feature Elimination with Support Vector Machine (RFE-SVM)

This is a type of wrapper algorithm that couples recursive feature elimination with linear SVM (Guyon *et al.*, 2002). In a linear SVM, the decision function (classifier) is given as $f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b$ or $f(\boldsymbol{x}) = \sum_{k=1}^{n} w_k x_k + b$. For a given feature x_k, the absolute value of its weight w_k indicates the significance of its contribution to the margin of the linear SVM and to the construction of a linear classifier. Hence, the weight w_k is used as a feature-ranking coefficient in RFE-SVM. This algorithm first constructs a linear SVM classifier from the microarray data with n genes; then, the gene with the lowest w_k^2 is removed and another classifier is trained on the remaining $n-1$ genes. This process is repeated until only one gene remains.

A gene ranking is produced from the order in which the genes were eliminated, and the last remaining gene is the most relevant gene. However, because of computational cost considerations, the algorithm is often implemented in such a way that several features are reduced at once. In such implementations, the method produces a feature subset ranking instead of a feature ranking. The optimal signature genes are the gene sets with the best prediction performance throughout the elimination process.

5.4 Filter Approaches Versus Wrapper Approaches

The filter methods are useful in practice because they are much faster than wrapper methods. Filters exclude irrelevant genes but cannot remove redundant genes because methods that select important genes based on individual gene information fail to account for mutual information among genes. The problem of redundancy among selected genes is twofold: the selected gene set may less comprehensively represent the target class than one of the same size without redundant genes would and in including all representative genes, redundant genes unnecessarily increase the size of the selected gene set, which will in turn affect the mining performance of the small sample.

In contrast, wrapper methods were developed to select discriminative genes while decreasing gene

redundancy. In theory, the wrapper model should provide more accurate classification results than the filter model (Langley, 1994). Wrappers use classifiers to estimate the usefulness of feature subsets. The use of optimal feature subsets should provide corresponding classifiers with better classification accuracy because the features are selected according to their contribution to the classification accuracy of the classifiers. The disadvantage of the wrapper approach is its computational expense because the classifier must be repeatedly constructed to evaluate a subset during the CV process. Moreover, wrapper-type feature selection is sensitive to training data and runs the risk of over-fitting, leading to a lack of robustness in the selected gene set.

6 ROBUSTNESS OF THE SELECTED GENE LIST

Although many methods have been developed for conducting feature selection on microarrays, these selection methods produce selected gene lists that are insufficiently robust. If the predictor developed with one selected gene set would work well on data from other studies, then we would not have had to worry about list diversity. However, a lack of transferability of predictive power is often observed as a result of the same reason that causes instability of gene lists. Ein-Dor, Zuk and Domany 2006 found that the gene lists developed from microarrays using different methods in prognostic cancer studies differ significantly, even for different subsets of the same microarray data sets. They concluded that thousands of samples are needed for robust gene selection. As generating a more stable gene list will lead to more robust predictors, we must evaluate not only prediction performance but also the robustness of gene sets in feature gene selection. Assessing the stability of selected gene lists is crucial to guarantee their controlled and reliable utilization.

7 APPLICATION OF FILTER AND WRAPPER METHODS TO DIAGNOSTICS OF DRUG TOXICITY USING TOXICOGENOMICS DATA

Our previous study on concurrent diagnosis of drug-induced tubular injury using TGP data demonstrates the application of filter- and wrapper-type methods (Kondo *et al.*, 2009). Drug-induced renal tubular injury is a major concern in pre-clinical safety evaluations. In this study, we analyzed 33 nephrotoxicants and eight non-nephrotoxic hepatotoxicants to elucidate time- and dose-dependent global gene expression changes associated with proximal tubular toxicity. The compounds were administered orally or intravenously once daily to male Sprague–Dawley rats. The animals were exposed to four different doses of the compounds, and kidney tissues were collected on days 4, 8, 15, and 29. High-dose groups of 23 compounds that caused necrosis, degeneration, or regeneration in the renal tubules during chronic exposure were defined as the positive set (other high-dose groups of 10 nephrotoxicants were used as the external test set). Low-dose groups of all 41 compounds and high-dose groups of the eight hepatotoxicants that had no histopathological findings were defined as the negative set. To perform supervised classification algorithms after selecting differentially expressed genes, the microarray samples treated with nephrotoxicants and hepatotoxicants were divided into positives and negatives of the training set according to their histopathological findings.

Both filter- and wrapper-type gene selection algorithms with SVM-learning algorithms were used to extract biomarker candidates and construct classifiers using the selected genes. RFE methods were used for the wrapper-type gene selection algorithms and intensity-based moderated *t*-statistics (IBMT; Sartor *et al.*, 2006) was used as a filter-type gene selection algorithm (SVM was used as the classifier in this case).

Fivefold CV was executed for optimization of the classifiers and to calculate their prediction accuracies. First, the whole positive and negative training data sets were randomly divided into five subsets of roughly equal size. The SVM was trained with a selection of optimal genes on four subsets and then applied to the fifth subset as the test data set. Before the SVM was trained, optimal genes were selected from the training sets with both RFE-SVM and IBMT-SVM. The 99 top-ranked genes from each selection strategy were used to construct the classifiers. The prediction model using the top-ranked 99 feature genes exhibited saturated prediction performance (Figures 2 and 3). Thus, by tracing the prediction performance of each constructed classifier for the test sets, we determined an appropriate

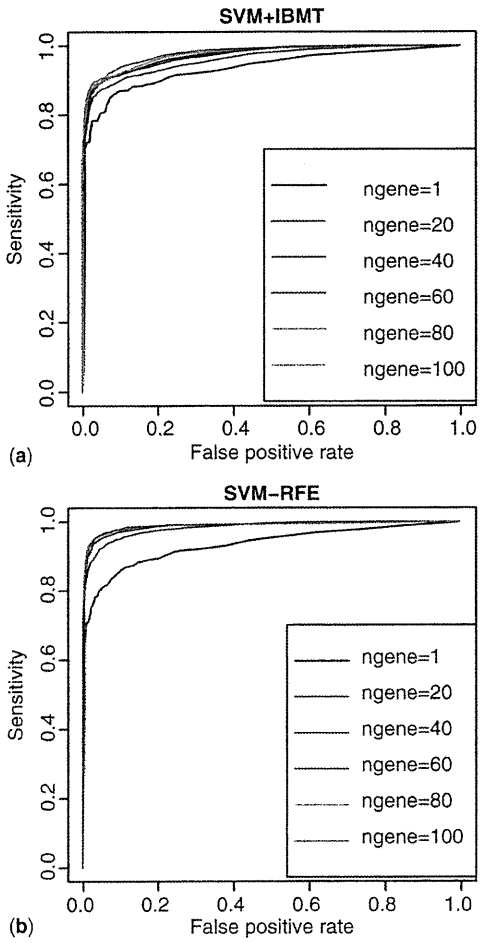

Figure 2. Receiver operating characteristic (ROC) curves for feature gene subsets selected using SVM-IBMT (a) and SVM-RFE (b).

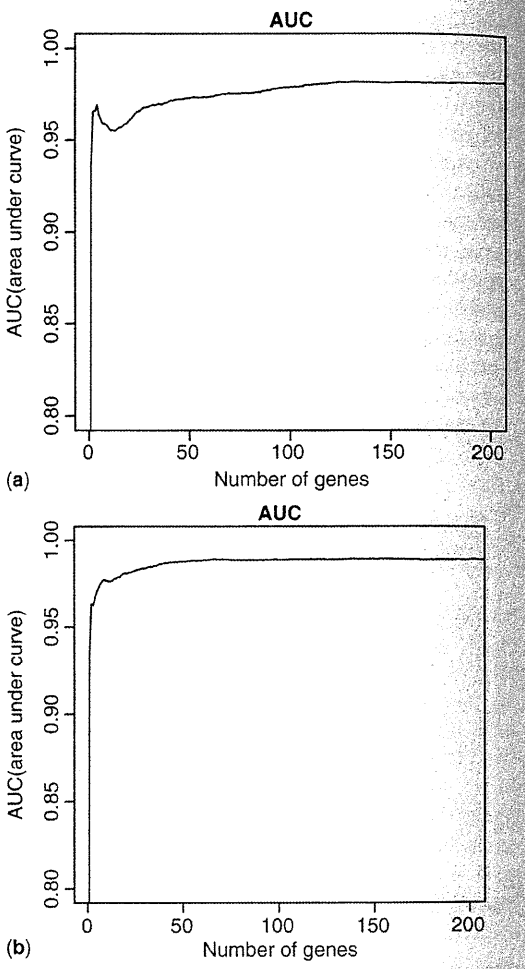

Figure 3. Area under curves (AUC) of the ROC curves by SVM-IBMT (a) and SVM-RFE (b), along with the number of selected feature genes.

threshold number of feature genes, which is one of the challenges in executing search algorithms.

As the result of fivefold CV, we determined the sensitivity of each classifier to be 94% (RFE-SVM; 99 probes) and 93.8% (IBMT-SVM; 99 probes), when we allowed for 10% false positives (Figure 2). Although SVM-RFE exhibited the higher classification accuracy, as we expected, the concordance rate of the feature gene list selected by RFE between the subtraining sets of fivefold CV was lower than that of the filter-type IBMT-SVM (Figure 4). The difference in concordance rate indicates that the RFE-SVM classifier may have over-fitted to the training set, resulting in an insufficiently robust list of biomarkers. In contrast, the feature genes selected by IBMT-SVM were stable between different training data sets generated through fivefold CV. As noted above, in addition to the prediction performance of the constructed classifier, the robustness (stability) of the selected gene sets is critical to successful feature selection. Therefore, in this case using TGP kidney data sets, we concluded that the filter-type IBMT-SVM method was the preferable gene selection and classification algorithm.

Consequently, we determined the appropriate maker gene set using the following criteria: (a) the prediction accuracy was saturated (Figure 3); (b) the

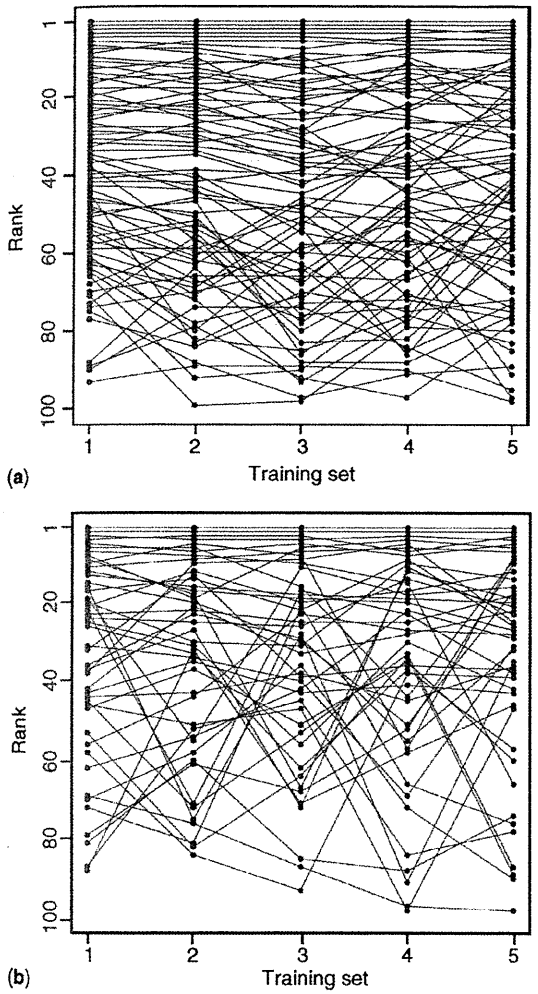

Figure 4. Perturbation of feature gene ranking between different training sets generated during fivefold CV. The horizontal axis indicates the training subsets and the vertical axis indicates the ranking of the feature genes (top 80 probes). Fivefold CV was executed by randomly dividing samples.

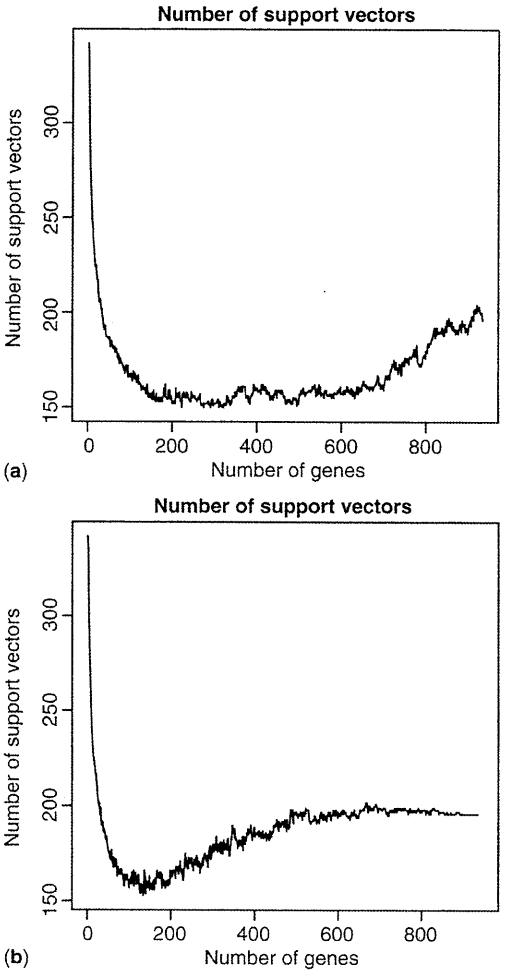

Figure 5. The number of support vectors in the constructed models of SVM-IBMT (a) and SVM-RFE (b), along with the number of selected feature genes.

number of support vectors was adequately low (an excessive number of support vectors could indicate over-fitting of the linear SVM classifier) (Figure 5), and (c) the number of feature genes was substantially lower than the number of samples, to avoid over-fitting.

Also, the selected gene list contained enough key genes to interpret their biological relevance in drug-induced renal tubular injury. The gene list contained well-known biomarkers, such as kidney injury molecule 1, ceruloplasmin, clusterin, and tissue inhibitor of metallopeptidase 1, as well as novel biomarker candidates. Most of the genes involved in tissue remodeling, immune/inflammatory response, cell adhesion/proliferation/migration, and metabolism were predominantly up-regulated. Down-regulated genes participated in cell adhesion/proliferation/migration, membrane transport, and signal transduction. This indicates that the gene list provides us with elaborate knowledge about mechanistic toxicology.

8 CONCLUSIONS

Selecting a small gene set from large microarray data sets is critical from both biological and computational viewpoints, and many algorithms have been developed for feature selection. Most studies on feature selection have found that wrapper methods are superior to filter methods, but many of these studies have over-emphasized prediction accuracy and over-looked the robustness of the selected genes. Prediction reliability assumes the stability of the model. In fact, this study illustrates that IBMT-SVM produces more stable gene lists than RFE-SVM. This finding is adaptable to only this training set. In the case of other training sets, we must evaluate multiple methods and choose the best approach. Therefore, we have to carefully gauge not only prediction performance but also the robustness of gene sets in feature gene selection.

ACKNOWLEDGEMENTS

This work was supported by grants from the Ministry of Health, Labour and Welfare of Japan. This work was also supported in part by the Ministry of Education, Culture, Sports, Science and Technology, Japan; the Japan Society for the Promotion of Science (JSPS); and the New Energy and Industrial Technology Development Organization (NEDO) of Japan.

REFERENCES

Bishop CM. 1995. *Neural Networks for Pattern Recognition*. Oxford University Press: London, UK.

Boverhof DR, Zacharewski TR. 2006. Toxicogenomics in risk assessment: application and needs. *Toxicol. Sci.* **89**: 352–360.

Cunningham ML, Lehman-McKeeman L. 2005. Applying toxicogenomics in mechanistic and predictive toxicology. *Toxicol. Sci.* **83**: 205–206.

Ein-Dor L, Zuk O, Domany E. 2006. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc. Natl Acad. Sci. USA* **103**: 5923–5928.

Hall MA. 2000. Correlation-based feature selection for discrete and numeric class machine learning. *Proceedings of the Seventeenth International Conference on Machine Learning Table of Contents*, San Fransisco, CA: Morgan Kaufmann Publishers Inc. 359–366.

Hayes KR, Vollrath AL, Zastrow GM, McMillan BJ, Craven M, Jovanovich S, Rank DR, Penn S, Walisser JA, Reddy JK, Thomas RS, Bradfield CA. 2005. EDGE: a centralized resource for the comparison, analysis, and distribution of toxicogenomic information. *Mol. Pharmacol.* **67**: 1360–1368.

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.

Kononenko I. 1994. Estimating attributes: analysis and extensions of RELIEF. *European Conference on Machine Learning*, Catania. Italy: Springer Verlag, 171–182.

Langley P. 1994. Selection of relevant features in machine learning. *Proceedings of AAAI Fall Symposium on Relevance*, New Orleans: AAAI Press, 140–144.

Mattingly CJ, Rosenstein MC, Davis AP, Colby GT, Forrest JN, Jr, Boyer JL. 2006. The comparative toxicogenomics database: a cross-species resource for building chemical-gene interaction networks. *Toxicol. Sci.* **92**: 587–595.

Sartor MA, Tomlinson CR, Wesselkamper SC, Sivaganesan S, Leikauf GD, Medvedovic M. 2006. Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments. *BMC Bioinf.* **7**: 538.

Thomas JG, Olson JM, Tapscott SJ, Zhao LP. 2001. An efficient and robust statistical modeling approach to discover differentially expressed genes using genomic expression profiles. *Genome Res.* **11**: 1227–1236.

Tsai C-A, Chen Y-J, Chen JJ. 2003. Testing for differentially expressed genes with microarray data. *Nucl. Acids Res.* **31**: e52.

Urushidani T, Nagao T. 2005. *Toxicogenomics: the Japanese initiative*. In Handbook of Toxicogenomics – Strategies and Applications, Borlak J (ed). Wiley-VCH: Weinheim; 623–631.

Vapnik VN. 1998. *Statistical Learning Theory*. Wiley, New York.

Waters M, Boorman G, Bushel P, Cunningham M, Irwin R, Merrick A, Olden K, Paules R, Selkirk J, Stasiewicz S, Weis B, Van Houten B, Walker N, Tennant R. 2003. Systems toxicology and the chemical effects in biological systems (CEBS) knowledge base. *EHP Toxicogenomics* **111**: 15–28.

別刷り

毒性評価　トキシコゲノミクス

同志社女子大学 薬学部 教授／医薬基盤研究所 基盤的研究部
トキシコゲノミクス・インフォマティクスプロジェクトリーダー　　漆谷　徹郎

『バイオチップ実用化ハンドブック』(2010年4月8日　株式会社エヌ・ティー・エス刊)