K305 Prediction Model Construction for Cholestasis by Toxicogenomics Approach

Takashi Matsuda[1)], Daisuke Kigami[2)], Koutaro Tamura[2)], Akira Unami[2)], Masato Kobori[1)], Tetsuro Urushidani[3)4)] and Yasuo Ohno[5)]

1) Molecular Medicine Research Labs., Drug Discovery Research, Astellas Pharma Inc., 21, Miyukigaoka, Tsukuba-shi, Ibaraki 305-8585 JAPAN
2) Drug Safety Labs., Drug Discovery Research, Astellas Pharma Inc., 1-6 Kashima 2-chome, Yodogawa-ku, Osaka 532-8514 JAPAN
3) Toxicogenomics Informatics Project, National Institute of Biomedical Innovation, Ibaraki, Osaka, 567-0085, JAPAN
4) Department of Pathophysiology, Doshisha Women's College of Liberal Arts, Kyotanabe, Kyoto 610-0395, JAPAN
5) National Institute of Health Sciences, Setagaya-ku, Tokyo 158-8501, JAPAN
takashi-matsuda@jp.astellas.com

1. Introduction

The Toxicogenomics Project in Japan (TGP) [1] was started in 2002 and finished in 2007 as a collaborative project by the National Institute of Health Science (NIHS) and 15 pharmaceutical companies. In 2007, the Toxicogenomics Informatics Project (TGP2) was started as a subsequent project of TGP. In these projects, about 150 compounds were comprehensively analyzed and large-scale transcriptome database were constructed.

In the present study, we selected cholestasis which is one of serious adverse effects as a target phenotype and constructed classifier for cholestasis based on linear. discriminant analysis (LDA) from microarray profile.

2. Materials and Methods

Animal treatment Male Sprague-Dawley rats were daily administered with each compound (3 dose levels) or vehicle control for 3, 7, 14 or 28 days. The animals were sacrificed 24 hours after the last dosing and the liver in each animal was obtained for microarray analysis. Three animals were analyzed for each group (4 time points and 4 dose levels including vehicle control).

Microarray analysis Microarray analysis was performed with GeneChip® Rat Genome 230 2.0 Arrays (Affymetrix, Santa Clara, CA, USA). The raw signal intensity data were normalized by global mean normalization method.

Compounds and training dataset From our database, we selected 10 cholestasis positive/negative compounds for training dataset. Alpha-naphthylisothiocyanate (1.5, 5, 15 mg/kg), chlorpromazine (4.5, 15, 45 mg/kg), carbamazepine (30, 100, 300 mg/kg), azathioprine (3, 10, 30 mg/kg) and ethinylestradiol (1, 3, 10 mg/kg) were selected as cholestasis positive compounds. Benzbromarone (20, 60, 200 mg/kg), hexachlorobenzene (30, 100, 300 mg/kg), Wy-14,643 (10, 30, 100 mg/kg), adapin (10,

30, 100 mg/kg) and pemoline (7.5, 25, 75 mg/kg) were selected as cholestasis negative compounds.

Statistical analysis We first extracted differentially expressed probe sets of each positive compound by comparing highest dose profile to corresponding control profile for each time point with Welch's t-test ($p<0.2$) and mean fold change ($fc>2.0$). Then we selected probe sets, which were commonly extracted from more than 4 out of 5 positive compounds, for next prediction model construction step.

In prediction model construction step, each profile was converted to fold change compared to control mean and classifier was constructed by LDA. To select effective probe sets for classifier from among probe sets selected previous step, we applied forward stepwise selection based on singular values and selected the smallest set of probe sets which provided the best accuracy estimated by leaving one compound out cross-validation (LOCO CV). Furthermore, stepwise feature elimination were performed while estimated accuracy by LOCO CV was not reduced. Finally, we constructed classifier for cholestasis by LDA with selected probe set.

3. Results and Discussion

The probe sets selected finally for cholestasis classifier are listed in Table 1 and ROC curve with test compounds profile consist of 7 positive and 6 negative compounds is shown in Figure 1. From the figure it is found that for example we can detect 80 % of cholestasis positive compound if we allow 30 % false positive.

Thus the classifier constructed in present study can classify cholestasis positive and negative compound profiles effectively.

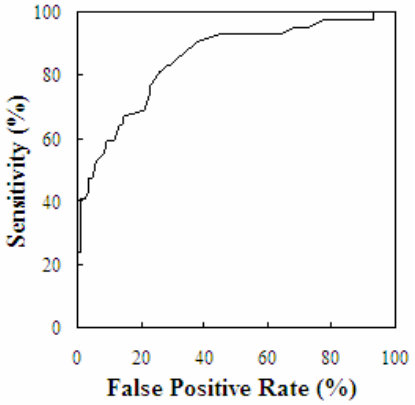

Fig. 1 ROC Curve

Table 1 probe set list

probe set	GeneSymbol	probe set	GeneSymbol	probe set	GeneSymbol
1367802_at	Sgk	1379027_at	RGD1308329_predicted	1387139_at	Hao2
1368147_at	Dusp1	1379513_at	Tmem30b_predicted	1387197_at	Omd
1368171_at	Lox	1380351_at	Sugt1	1387270_at	Hhex
1368272_at	Got1	1381748_at	Raph1_predicted	1387583_at	Cyp26a1
1368607_at	Cyp4a12	1382443_at	RGD1562451_predicted	1387985_a_at	Obp3
1370026_at	Cryab	1382451_at	Hebp2_predicted	1388199_at	Tacstd1
1370384_a_at	Prlr	1383047_at	Gas6	1389221_at	Mmd2_predicted
1371143_at	Serpina7	1383248_at	Fmo5	1389528_s_at	Jun
1373146_at	Ssx2ip	1383486_at	EST	1390860_at	Igf2bp3
1374093_at	EST	1383946_at	Cldn1	1395542_at	EST
1374883_at	Mtmr7_predicted	1385005_at	EST	1397596_at	Trim2
1375647_at	EST	1386922_at	Ca2	1398309_at	Pigl
1376958_at	RGD1562844_predicted	1387022_at	Aldh1a1	1398759_at	Tgfb1i4

Acknowledgement

This study was supported in part by a grant from the Ministry of Health, Labor and Welfare (H14 and H19-Toxico-001).

References

1. Urushidani T, Nagao T, **2005**, Handbook of Toxicogenomics – Strategies ans Applicatioins, Wiley – VCH, 623-631.