

TGP成果の共有・公開 → 参加企業

TG-GATEs: データベース・解析・予測ツール

TGP2 Toxicogenomics Informatics Project in Japan

- welcome to TG-GATEs
System Ver: 2.0.0 Manual
- TGP-Database
 - Data Search
 - Manager
 - Sample List Manager
 - 1. Search
 - 2. Registration
 - Gene List Manager
 - 1. Search
 - 2. Registration
 - 3. Marker Score
 - Matrix Manager
 - 1. Create Matrix
 - 2. Matrix List
 - 3. Search
 - Preset Manager
 - 1. Delete
 - TGP-Analyzer
 - Base View
 - PCA & Clustering
 - Correlation
 - TGP-Predictor
 - Create Discriminator
 - Discriminator List
 - Prediction
 - Analysis Result
 - Data Management
 - Registration
 - Data Update
 - Log View
 - User Admin

全てのバイオマーカー・予測モデルも配布

TGP成果の共有・公開 → 一般公開

OpenTG-GATEs: Tgxデータ・毒性試験情報の公開

TGP2 Toxicogenomics Informatics Project in Japan

NIBIO 独立行政法人 医薬基盤研究所
National Institute of Biomedical Innovation

<http://toxico.nibio.go.jp>

2011年02月25日

OpenTG-GATEsを公開.

2011年03月22日

CELファイルの一括ダウンロード機能を追加。Open TG-GATEsの全CELファイルをライフサイエンス統合データベースセンターに寄託し、FTPサイトにて公開

2011年04月28日

Pathological Item表示画面で同じレコードが2件表示される不具合を修正。

2011年09月16日

Ciprofloxacin, Rat, in vivo, Kidney, High Dose, 29dayに関する個体データ1件を修正。

2012年1月18日

TGP2で実験された追加データを追加

2012年2月末日

ライフサイエンス統合DBセンターにてTGP2追加実験データ(CELファイル)、毒性データを公開



- このDBについて
- データ検索
- 公開データ一覧
- 利用許諾条件
- お問合せ

CompoundかPathologyのどちらか一つの検索条件を選択し[Search]をクリックしてください。

Compound List

select compound

- meloxicam
- metformin
- methapyllene
- methimazole
- methyldopa
- methyltestosterone
- medfeline
- monocrotaline
- moxitylone
- tophthyl isothiocyanate

上面化合物を一つ選択してください

Search

生物種・実験条件の選択画面に進みます

Pathology List

select organ

- Liver
- Kidney

LiverかKidneyのどちらかを選択してください

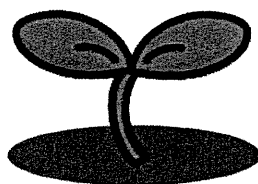
Search

病理所見の選択画面に進みます

個別のバイオマーカーは、論文等で発表

まとめ

- プロジェクト目標の30種以上のバイオマーカー構築を達成した
- 信頼性や再現性について参加企業研究員により評価済み
- トランスクリプトームバイオマーカーは、単なる従来指標の代替としてではなく、毒性メカニズムに関する評価を加味することで、
 - リード選択の最適化
 - 開発後期における安全性問題の回避
 - 既存薬との差別化
 - 医薬品開発に関わるコストや時間の削減
- リファレンスデータ(ゲノクスデータ・毒性データ)及びバイオマーカーの一般公開(論文公開)により、
 - 更なる信頼性の検証
 - バイオマーカーのブラッシュアップ
 - ヒト外挿性の検討



29

Thanks to all member of the TGP2 biomarker Working group

参加企業研究員 (敬称略・50音順)

青木 幹雄	豊柴 博義
赤井 翔	鳥塚 尙樹
芦原 基起	西原 久美子
阿部 香織	新田 浩之
上原 健城	半田 千彰
宇波 明	廣出 充洋
大村 功	福島 亮
奥山 学	堀之内 彰
甲斐 敏裕	本間 俊樹
川上 由里子	松下 智哉
神吉 将之	松田 喬
木野 潤一	南 圭一
木上 大輔	三宅 真波
清澤 直樹	武藤 裕紀
近藤 千晶	森下 克美
清水 俊敦	山田 文博
住田 佳代	山田 徹
棚治 隆史	矢本 敬
富永 幸雄	弓立 恭寛

同志社女子大大学

漆谷 徹郎

TGP2事務局

山田 弘

了戒 由宇子

TGP2研究員

中津 則之

箕輪 洋介

高 衛華

五十嵐 芳暢

森川 裕二

京都大学

奥野 恭史

その他、TGP関係者・テクニカルスタッフ・IT関係者の皆様



30

厚生労働科学研究費補助金（創薬基盤推進研究事業(創薬バイオマーカー探索研究事業)）
分担研究報告書

バイオマーカー候補遺伝子の検証

研究分担者 水川裕美子 同志社女子大学薬学部 特別任用助教

研究要旨

トキシコゲノミクス・インフォマティクスプロジェクト(TGP2)ではバイオマーカーの創出を目指して大規模データベースの解析を行っているが、選び出されたバイオマーカー候補遺伝子が真に創薬に役立つバイオマーカーとして利用可能になるためには各方面からの検証が必要である。そこで、本分担研究で *in vitro* を中心とした小規模実験を行うことによる毒性メカニズムの解析や、公共データベースの外部データを利用した検証を行っている。

最終年度である今年度は *in vitro* のリン脂質症マーカーの抽出と検証、血液凝固不全マーカーの改良と検証および peroxisome proliferator-activated receptor (PPAR) α 活性化マーカーの改良と検証を行った。いずれのマーカーについてもその判定方法にはサポートベクターマシン(SVM)による判別分析を用い、内部データ、外部データともに適用可能なマーカーおよび判別モデルが作成できたと考えられた。

A. 研究目的

トキシコゲノミクス・インフォマティクスプロジェクト(TGP2)では平成 14~18 年度に行われたトキシコゲノミクスプロジェクト(TGP1)で構築した大規模データベース(TG-GATEs)を活用したバイオマーカーの創出を目指して解析を行っている。しかし選び出されたバイオマーカー候補遺伝子が真に創薬に役立つバイオマーカーとして利用可能になるためには各方面からの検証が必要であるため、本研究では *in vitro* を中心とした小規模実験を行い、毒性メカニズムの解析、裏付けの面から検証を行っている。また公共データベースを用いた外部データでの検証も行っている。これまでにグルタチオン枯渇マーカーの抽出と検証、*in vivo* でのリン脂質症マーカーの検証、*in vitro* でのリン脂質症マーカーの抽出などを行って

きたが、最終年度である今年度は *in vitro* のリン脂質症マーカーの抽出と検証を続けた他、血液凝固不全マーカーの検証および PPAR α 活性化マーカーの検証を行った。

リン脂質症は肝臓、肺、腎臓などの細胞内にリン脂質が蓄積する現象であり、薬剤性リン脂質症を惹起する化合物は多数知られている。毒性との関係は明確ではないが、新薬のドロップアウトの原因にもなりかねないため、リン脂質症を惹起するかどうかを新薬開発の早期の段階で判断できることは大変有益である。TGP においてもラット肝臓の遺伝子発現データを用いて抽出したマーカーが複数作成されているが、*in vitro* のマーカーはこれまでなかったため、昨年度、ラット初代培養肝細胞の遺伝子発現データからマーカー抽出を試みた。今年度は引き続きマーカー抽出と判別モデル構築を

行い、一定の有効性が認められるマーカーを作成できた。

また、これまでに TGP1/2 で選抜されたバイオマーカー候補遺伝子セットについて、遺伝子の絞り込みや判別モデルの構築によりマーカーとしての改良を図り、公共データベースに登録された外部データを用いて検証を行った。その一つは肝障害で凝固に必要な因子の産生が低下することなどに起因する血液凝固不全のマーカーであり、もう一つは PPAR α 活性化マーカーである。PPAR α 活性化マーカーについては、これまで選抜されたマーカーはラット *in vivo* の肝臓および初代培養肝細胞の遺伝子発現データについて適用できるものであるが、今回はその改良を行ったことに加え、ヒト初代培養肝細胞のデータに適用できるものを作成した。

B. 研究方法

1. ラット肝細胞初代培養系におけるリン脂質症マーカーの抽出と検証

アミオダロン(AM), イミプラミン(IMI), アミトリプチリン(AMT), ヒドロキシジン(HYZ), クロミプラミン(CPM), プロメタジン(PMZ), パーヘキシリン(PH)の7化合物のラット肝細胞 24 時間曝露遺伝子発現データより、4 化合物以上で共通に同一方向に発現変動した遺伝子を抽出した。これらの化合物はリン脂質症を惹起することが知られており、昨年度検討を行った *in vitro* リン脂質症アッセイでも陽性の結果が出ている。抽出した遺伝子を用いてサポートベクターマシン(SVM)による判別モデルを作成した。サポートベクターの数が最小となる使用遺伝子数を選択し、マーカー遺伝子とした。

またこれを用いて TG-GATEs のラット初代培養肝細胞遺伝子発現データについて判別テストを行った。さらに、外部データを用いて検証するため、米国 NCBI の Gene Expresson Omnibus (GEO)データベースにおいて、リン脂質症を惹起する化合物を曝露したラット初代培養肝細胞の遺伝子発現データを検索し、できるかぎり用量・時点などの条件が TG-GATEs のデータと近いものを選んで検証に用いた。TG-GATEs と同一プラットフォームのデータが見つからなかったため、異なるプラットフォーム(Agilent)のデータを用いた。ノーマライズ済データをダウンロードし、TG-GATEs のデータと同様に対照群に対する log₂ ratio にして判別分析を行った。両プラットフォームに共通に存在する遺伝子のデータのみを用いた。

2. 血液凝固不全マーカーの改良と検証

肝障害により血液凝固に必要な因子の産生が低下するなどし、凝固不全を起こす場合がある。ラット肝臓の遺伝子発現データでそのような血液凝固不全を検出できるとして以前提案されたマーカー(Hirode *et al.*, *J. Toxicol. Sci.*, **34**, 281-293, 2009) について検証を行った。この遺伝子セットは以下のようにして得られたものである。反復投与(3, 7, 14, 28 日間投与 24 時間後にサンプリング)の血液生化学データにおいてプロトロンビン時間(PT)延長, 活性化部分トロンボプラスチン時間(APTT)延長, フィブリノーゲン(Fbg)低下のいずれかがみられた化合物のうち、クロフィブラート(CFB), オメプラゾール(OPZ), エチオニン(ET), チオアセタミド(TAA), ベンズプロマロン(BBr), プロピルチオウラシル(PTU), WY-14643 (WY), AM の 8 化合物を選択し、ラット肝臓遺伝

子発現データにおいて 8 化合物共通にいずれかの時点で変動した遺伝子を抽出後、アノテーション不明のものを除いた。

今回の検証では、これら 344 プローブセットから統計フィルターでさらに遺伝子を絞り、SVM による判別モデルを作成した。Receiver Operating Characteristic (ROC)カーブの曲線下面積(AUC)が最大になる使用遺伝子数を選択し、マーカー遺伝子とした。このモデルを用いて TG-GATEs のラット肝臓遺伝子発現データについて判別を行い、血液生化学データとの比較により検証した。また、NCBI GEO データベースを検索し、ラットに陽性化合物を投与し、24 時間後以降にサンプリングした肝臓を TG-GATEs と同じプラットフォーム(Affymetrix GeneChip Rat Genome 230 2.0 array)で遺伝子発現解析されたものを選択・ダウンロードした。TG-GATEs のデータと同様にデータ処理して判別分析を行うことにより、外部データでの検証を行った。

3. PPAR α 活性化マーカーの改良と検証

in vivo, *in vitro* の両方で使用できる PPAR α アゴニスト検出のバイオマーカー候補として以前提案された遺伝子セット (Tamura *et al.*, *J. Toxicol. Sci.* **31**, 471-490, 2006)について、遺伝子の絞り込みと判別モデルの構築により、マーカー改良を行った。提案されている遺伝子セットは、TG-GATEs のデータの中で代表的な PPAR α アゴニストである CFB, WY, ゲムフィブロジル (GFZ) のうち少なくとも 2 化合物で変動した遺伝子を単回投与 24 時間後のラット肝臓 (*vivo*24h), 28 日反復投与 24 時間後のラット肝臓(*vivo*29d), 曝露開始 24 時間後のラット初代培養肝細胞(*vitro*24h)についてそれぞれ

抽出し, *vivo*24h と *vitro*24h に共通なものとして *vivo*29d と *vitro*24h に共通なものを合わせた 41 プローブセットである。直接 PPAR α の下流の応答を重視するため、まずこれらのうち *vivo*24h と *vitro*24h に共通な 29 プローブセットに絞った。これを用いて SVM による判別モデルを *in vivo*, *in vitro* のデータについてそれぞれ構築し、ROC カーブの AUC が最大となるよう使用遺伝子を選択した。この判別モデルを用いて TG-GATEs のデータについて判別を行ったところ、典型的な PPAR α アゴニストであるフィブラート系薬物以外にも陽性という判別結果が出た化合物がいくつかあったため、無細胞系での PPAR α 活性化アッセイキット(EnBio RCAS for PPAR α ; 藤倉化成)を用いてアッセイを行い、比較することにより検証を行った。また、NCBI GEO データベースから陽性または陰性化合物を投与したラット肝臓およびラット初代培養肝細胞の Affymetrix GeneChip Rat Genome 230 2.0 array による遺伝子発現データをダウンロードし、TG-GATEs のデータと同様にデータ処理して判別分析を行うことにより、外部データでの検証を行った。

さらに、このマーカーがヒト初代培養肝細胞のデータにも使用可能かどうか検討した。まずラットで抽出した 29 プローブセットのオーソログの発現について検討したところ、変動がみられたのはごく一部であったため、新たに発現変動遺伝子を抽出し直した。PPAR α 活性化アッセイで陽性だった 10 化合物のうち 7 化合物以上において、中用量または高用量で曝露 8 時間または 24 時間において同一方向に発現変動した遺伝子を選択した。これを用いて SVM による判別

モデルを作成し、ROC カーブの AUC が最大となるよう使用遺伝子を選択した。このモデルで TG-GATEs のヒト初代培養肝細胞の遺伝子発現データを判別し、文献的な情報や PPAR α 活性化アッセイの結果と比較することによりその内容を検証した。また、NCBI GEO データベースから陽性または陰性化合物を投与したヒト初代培養肝細胞の Affymetrix GeneChip Human Genome U133 2.0 array による遺伝子発現データをダウンロードし、TG-GATEs のデータと同様にデータ処理して判別分析を行うことにより、外部データでの検証を行った。

(倫理面への配慮)

本研究は同志社女子大学動物実験委員会の承認を得て行っている。使用する動物の屠殺にあたっては、麻酔薬の使用で苦痛の少ない方法を用いるといった本学の実験動物取扱規定に従った対応を行っている。

C. 研究結果

1. ラット肝細胞初代培養系におけるリン脂質症マーカーの抽出と検証

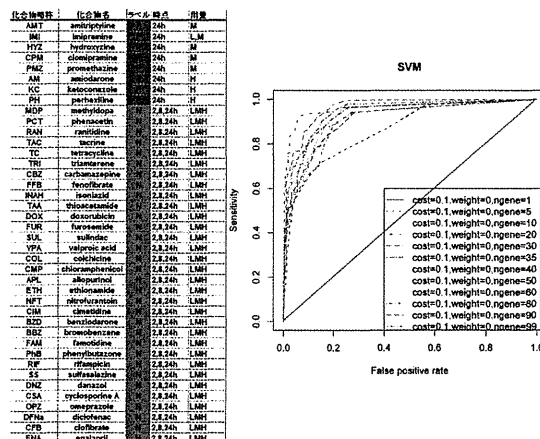
リン脂質症陽性の 7 化合物中 4 化合物以上で同一方向に変動した遺伝子を抽出した結果、1190 プローブセットが抽出された。判別モデルの学習に用いたデータセットと得られた ROC カーブを図 1 に示す。まずまず偽陽性率が低く感度の高いモデルが構築できたといえる。サポートベクター最小となる 71 プローブセット (表 1) が使用遺伝子として選択された。選択された遺伝子にはリン酸と関連するものや膜蛋白質関連のものも多く、内容的に妥当な遺伝子が選択されたと考えられる。このモデルを用いて TG-GATEs の全化合物のデータについて判

別を行った結果、学習セットについてはいずれの化合物も正しく判別され、それ以外については、ジルチアゼムなど一部は *in vivo* と共通に陽性の結果となった。次に GEO データベースから GSE19662 データセットに含まれるデータをダウンロードし、検証を行った。このデータは AM 1 ppm をラット初代培養肝細胞に 24 時間曝露後サンプリングし、Agilent-014879 Whole Rat Genome Microarray 4x44K G4131F でデータ取得したものである。プラットフォーム間で共通に存在する遺伝子のデータだけを用いて同様に判別を行ったところ、判別結果は陰性であった。薬物の濃度は約 1.55 μ M で陽性判定された TG-GATEs の中用量とほぼ同程度であったが、プラットフォームが異なるために正しく判別できなかった可能性が考えられた。

2. 血液凝固不全マーカーの改良と検証

血液凝固不全を引き起こす 8 化合物に共通して変動した 344 プローブセットに統計フィルターを適用後、判別モデルを作成した。主成分分析の結果などから、陽性化合物のうち Fbg 低下が激しいものは PPAR α ア

(図 1) リン脂質症 *in vitro* マーカーの学習セットと ROC カーブ

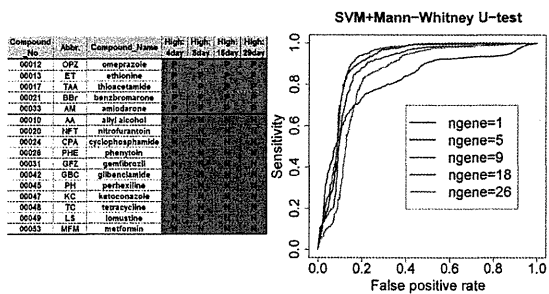


(表 1) リン脂質症 *in vitro* 判別モデルのマーカー遺伝子

Probe set ID	Gene symbol	Gene title
1368491_at	Dnase2b	deoxyribonuclease II beta
1387659_at	Gda	guanine deaminase
1389746_at	RGD1564228	similar to Naglu
1384428_at	Tmem181	transmembrane protein 181
1373939_at	Nagk	N-acetylglucosamine kinase
1370475_at	Cyp2b3	cytochrome P450, family 2, subfamily b, polypeptide 3
1372304_at	Comm4	COMM domain containing 4
1368074_at	Gale	UDP-galactose-4-epimerase
1381982_at	Uap111	UDP-N-acetylglucosamine pyrophosphorylase 1-like 1
1367514_at	LOC361635	similar to RIKEN cDNA 9030624J02
1398254_at	Renbp	renin binding protein
1371871_at	Rab12	RAB12, member RAS oncogene family
1367562_at	Sparc	secreted protein, acidic, cysteine-rich (osteonectin)
1367913_at	Cygb	cytoglobin
1378196_at	Slc43a1	solute carrier family 43, member 1
1398363_at	Wdr54	WD repeat domain 54
1368718_at	Aldh1a7	aldehyde dehydrogenase family 1, subfamily A 7
1372687_at	Croct1	cysteine-rich C-terminal 1
1386967_at	Rhoq	ras homolog gene family, member Q
1368695_at	C4bpb	complement component 4 binding protein, beta
1372675_at	RGD1306954	similar to RIKEN cDNA 1110004E09
1386893_at	Grn	granulin
1387024_at	Dusp6	dual specificity phosphatase 6
1374877_at	LOC687295	similar to translocase of inner mitochondrial membrane 50 homolog
1370789_a_at	Prlr	prolactin receptor
1371021_at	Arsb	arylsulfatase B
1370949_at	---	---
1389106_at	Fbxw9	F-box and WD repeat domain containing 9
1386884_at	Htra1	HtrA serine peptidase 1
1370280_at	Hprt1	hypoxanthine phosphoribosyltransferase 1
1390848_at	Rbml9	RNA binding motif protein 19
1375972_at	LOC100360533 /// RGD1307493	family with sequence similarity 57, member A /// similar to membrane protein expressed in epithelial-like lung adenocarcinoma
1388675_at	Spns1	spinster homolog 1 (Drosophila)
1391435_at	Pltp	phospholipid transfer protein
1372013_at	Ifitm1	interferon induced transmembrane protein 1
1389581_at	Il33	interleukin 33
1367712_at	Timp1	TIMP metalloproteinase inhibitor 1
1373258_at	Ctsf	cathepsin F
1374806_at	Sfn	stratifin
1390860_at	Igf2bp3	insulin-like growth factor 2 mRNA binding protein 3
1375560_at	Dhdps1	dihydrodipicolinate synthase-like, mitochondrial
1376249_at	Fuca2	fucosidase, alpha-L-2, plasma
1371946_at	RGD1303130	kidney predominant protein NCU-G1
1379255_at	Atp6ap2	ATPase, H+ transporting, lysosomal accessory protein 2
1372809_at	LOC290595	hypothetical gene supported by AF152002
1372897_at	Plod2	procollagen lysine, 2-oxoglutarate 5-dioxygenase 2
1370310_at	Hmgcs2	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial)
1368309_at	Txnr2	thioredoxin reductase 2
1372741_at	Scpdp	saccharopine dehydrogenase (putative)
1370155_at	Col1a2	collagen, type I, alpha 2
1371143_at	Serpina7	serine (or cysteine) peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7
1373970_at	Il33	interleukin 33
1371357_at	Igfbp7	insulin-like growth factor binding protein 7
1392547_at	MGC105649	hypothetical LOC302884
1399096_at	Add3	adducin 3 (gamma)
1368191_a_at	Slc22a1	solute carrier family 22 (organic cation transporter), member 1
1392534_at	Pmpa1	prostate transmembrane protein, androgen induced 1
1372806_at	Vps35	vacuolar protein sorting 35 homolog (S. cerevisiae)
1373245_at	Col4a1	collagen, type IV, alpha 1
1389559_at	RGD1564420	similar to Hypothetical protein MGC31278
1377049_at	Pnpla7	patatin-like phospholipase domain containing 7
1370959_at	Col3a1	collagen, type III, alpha 1
1388547_at	Cldn4	claudin 4
1398930_at	Atp6v0b	ATPase, H+ transporting, lysosomal 21kDa, V0 subunit b
1393170_at	RGD1561431	similar to homeobox-containing transcription factor
1389123_at	Ccl6	chemokine (C-C motif) ligand 6
1396279_at	Atp6ap2	ATPase, H+ transporting, lysosomal accessory protein 2
1374396_at	Atp6v1c1	ATPase, H+ transporting, lysosomal V1 subunit C1
1369029_at	Plscr1	phospholipid scramblase 1
1368741_at	C9	complement component 9
1373617_at	Emp2	epithelial membrane protein 2

ゴニストでもある場合が多いことがわかってきたため、その影響を除くため、学習には Fbg 低下の激しい化合物は使用しないこととした。学習に用いたデータセットと得られた

(図2) 血液凝固マーカーの学習セットとROCカーブ



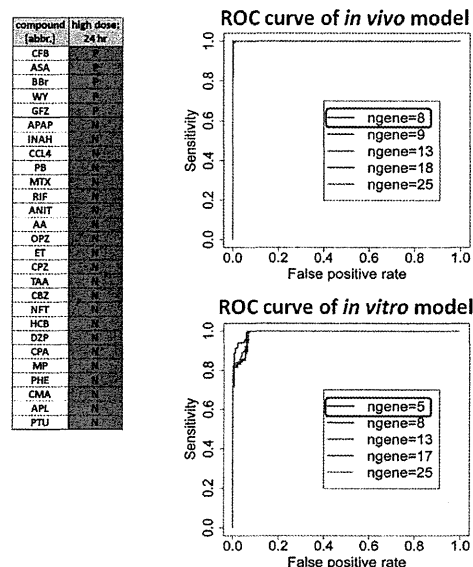
ROCカーブを図2に示すが、まずまずの偽陽性率と感度のモデルが構築できた。学習セットに含まれるデータについては、陰性化合物であるGFZの高用量・単回投与24時間後のデータが陽性と判別された以外は異常な判別結果はなかった。それ以外の化合物について、血液凝固不全を起こすCFB, WY, PTUやアセトアミドフルオレン, エチオナミドなどが陽性と判別されたが、一部、凝固異常はなかったり凝固を促進させたりした化合物で陽性と判別された例もあった。

GEOデータベースから陽性化合物を投与したラット肝臓の遺伝子発現データをダウンロードし判別分析を行ったところ、GSE24363のTAA投与では用量依存的に陽性と判別された。またGSE31307のET投与では陽性との判別結果が出たが、CFBでは陰性だった。このデータセットは用量不明なので用量不足のために判別結果が陰性になったと考えられ、外部データについても適用可能と考えられた。

(表2) 血液凝固不全判別モデルのマーカー遺伝子

Probe set ID	Gene Symbol	Gene Title
1386884_at	Htra1	HtrA serine peptidase 1
1368674_at	Pygl	phosphorylase, glycogen, liver
1368077_at	Fbp1	fructose-1,6- biphosphatase 1
1371789_at	Clpp	ClpP caseinolytic peptidase, ATP-dependent, proteolytic subunit homolog (E. coli)
1371809_at	Mrps18b	mitochondrial ribosomal protein S18B
1386917_at	Pc	pyruvate carboxylase
1389903_at	Pttglip	pituitary tumor-transforming 1 interacting protein
1368651_at	Pklr	pyruvate kinase, liver and RBC
1377033_at	Serpinf2	serine (or cysteine) peptidase inhibitor, clade F, member 2

(図3) ラット *in vivo* および *in vitro* のPPAR α 活性化マーカーの学習セットとROCカーブ



この判別モデルで使用するマーカーとして選択された遺伝子9プローブセット(表2)のうち1つはKEGGのcomplement and coagulation cascadesに含まれている。その他解糖系やピルビン酸代謝に関連する遺伝子が多く含まれていた。

3. PPAR α 活性化マーカーの改良と検証

Vivo24h, vitro24h 共通にCFB, WY, GFZのうち2化合物以上で発現増大した遺伝子29プローブセットから *in vivo*, *in vitro* それぞれSVMによる判別モデルを構築したところ、図3に示すようにいずれも非常に良好なROCカーブを示すモデルが得られた。この判別モデルで使用するマーカーとして選択された遺伝子は表3に示すように *in*

(表 3) ラット *in vivo* および *in vitro* での PPAR α 活性化判別モデルのマーカー遺伝子

in vivo モデル

Probe set ID	Gene Symbol	Gene Title
1388211_s_at	Acot1 /// Acot2	acyl-CoA thioesterase 1 /// acyl-CoA thioesterase 2
1368283_at	Ehhadh	enoyl-Coenzyme A, hydratase/3-hydroxyacyl Coenzyme A dehydrogenase
1389253_at	Vnn1	vanin 1
1377037_at	Acot4	acyl-CoA thioesterase 4
1368934_at	Cyp4a1 /// Cyp4a10	cytochrome P450, family 4, subfamily a, polypeptide 1 /// cytochrome P450, family 4, subfamily a, polypeptide 10
1367659_s_at	Dci	dodecenoyl-Coenzyme A delta isomerase (3,2 trans-enoyl-Coenzyme A isomerase)
1370818_at	Dec2	2,4-dienoyl CoA reductase 2, peroxisomal
1386880_at	Acaa2	acetyl-Coenzyme A acyltransferase 2

in vitro モデル

Probe set ID	Gene Symbol	Gene Title
1368283_at	Ehhadh	enoyl-Coenzyme A, hydratase/3-hydroxyacyl Coenzyme A dehydrogenase
1370397_at	Cyp4a3	cytochrome P450, family 4, subfamily a, polypeptide 3
1387783_a_at	Acaa1 /// RGD1562373	acetyl-Coenzyme A acyltransferase 1 /// similar to 3-ketoacyl-CoA thioesterase B, peroxisomal precursor (Beta-ketothioesterase B)
1367659_s_at	Dci	dodecenoyl-Coenzyme A delta isomerase (3,2 trans-enoyl-Coenzyme A isomerase)
1367777_at	Dec1	2,4-dienoyl CoA reductase 1, mitochondrial

in vivo モデルで 8 プローブセット, *in vitro* モデルで 5 プローブセットであり, 一部は両者に共通していた. また殆どが PPAR との関連が既に知られており, 脂質代謝などに関与する遺伝子である. このことから遺伝子の内容の面からもこの判別モデルの妥当性は支持できた.

TG-GATEs のデータについて判別分析を行ったところ, 学習セットの化合物につい

(表 4) *in vivo*, *in vitro* の予測結果および無細胞系アッセイの結果の比較. ○: 陽性, ×: 陰性, 網掛け: *in vivo* の用量が低かったと考えられるもの, 斜線: 代謝物がアゴニストと考えられるもの, 白抜き: メカニズム不明のもの

compound	prediction <i>in vivo</i> H:24hr	prediction <i>in vitro</i> H:24hr	cell-free assay
CFB	○	○	×
ASA	○	○	×
BBr	○	○	○
WY	○	○	○
GFZ	○	○	○
ANIT	×	M:8,24hr	○
AM	H:8,15,29da	×	×
TBF	○	×	×
VPA	×	○	○
IM	×	○	○
DFNa	×	M:24hr	○
SS	×	○	×
NPX	×	○	○
FFB	○	○	×
MEF	×	○	○
SST	○	○	×
BZD	○	○	○
clofibrlic acid			○
salicylic acid			×

ては *in vitro* の 1 化合物を除き異常な例はなかった. 学習セット以外の化合物についても, PPAR α を活性化するフェノフィブラート (FFB) が陽性となるなど適切な結果が得られた. また *in vitro* ではバルプロ酸 (VPA) や多くの NSAIDs が陽性と判別された. これらフィブラート系薬物以外の化合物が実際に PPAR α アゴニスト活性を持つかどうか検証するため, ヒト PPAR α とコアクチベーターとの結合を検出する無細胞系アッセイを行った. その結果, VPA や多くの NSAIDs に PPAR α アゴニスト活性が検出され, 判別結果が妥当であったことがわかった.

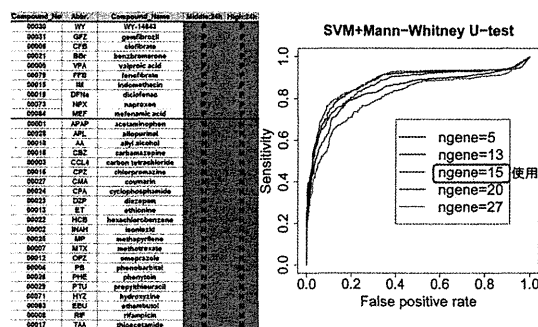
In vivo, *in vitro* の予測結果および無細胞系アッセイの結果の 3 者が一致しなかった化合物もいくつかあったが (表 4), 大部分は説明のつくものだった. 例えば CFB はフィブラート系薬物であり PPAR α を活性化する代表的な薬物といえるが, 無細胞アッセイでは陰性だった. CFB のエステル結合が切断された作用本体であるクロフィブリン酸は陽性だったことから, 細胞内のエステ

ラーゼでの代謝の有無により結果が一致しなかったことがわかった。

外部データで検証を行うため、陽性または陰性化合物を投与した遺伝子発現データをNCBI GEOからダウンロードし、判別を行った。その結果、*in vivo* ではGSE14712のWY投与で陽性、GSE24363のプロモベンゼン(BBZ)、モキシシリト(MCT)、TAA投与でいずれも陰性と判別された。これらはいずれも期待通りの結果である。また*in vitro* では登録されたデータが少なく陽性化合物であるWYのみしか検討できなかったが、正しく陽性と判別された。

次にヒト肝細胞のデータに使用する判別モデルの取得を目指した。まずラットのモデルで出発点とした29プローブセットをオーソログ変換し、発現変動を観察したところ、変動のみられた遺伝子はごくわずかであった。そこで、ヒトで変動する遺伝子群は異なっているかもしれないと考え、全プローブセットから発現変動遺伝子を抽出し直した。フィブラート系薬物および無細胞系PPAR α 活性化アッセイで陽性だった薬物のうち、CFB, WY, GFZ, FFB, VPA, BBr, ナプロキセン(NPX), メフェナム酸

(図4) ヒト初代培養肝細胞用PPAR α 活性化マーカーの学習セットとROCカーブ



(MEF), ジクロフェナク(DFNa), インドメタシン(IM)の10化合物中7化合物で変動した遺伝子を抽出し、判別モデルの構築を行った。学習セットとROCカーブを図4に示す。ROCカーブはラットほどではないが、まずまず良好なモデルが得られたと考えられた。このモデルをTG-GATEsのデータに適用して判別を行ったところ、ラット初代培養肝細胞と非常に近い判別結果であった。またこの判別モデルに用いた遺伝子15プローブセット(表5)のうち、ラットでモデル構築の出発点とした29プローブセットのオーソログが3遺伝子あり、またそれ以外の遺伝子にも脂質代謝に関わりPPARとの関連が報告されているものが多かった。しかしラット*in vivo*, *in vitro*判別モデルに使用した遺伝子とは共通するものはなかった。

(表5) ヒト肝細胞でのPPAR α 活性化判別モデルのマーカー遺伝子。色を付けた遺伝子はラット29プローブセットのオーソログに含まれるもの。

Probe set ID	Gene Symbol	Gene Title
208383_s_at	PCK1	phosphoenolpyruvate carboxykinase 1 (soluble)
205892_s_at	FABP1	fatty acid binding protein 1, liver
225207_at	PKD4	pyruvate dehydrogenase kinase, isozyme 4
219398_at	CIDEC	cell death-inducing DFFA-like effector c
201272_at	AKR1B1	aldo-keto reductase family 1, member B1 (aldose reductase)
238982_at	DENR	density-regulated protein
221009_s_at	ANGPTL4	angiopoietin-like 4
221059_s_at	COTL1	coactosin-like 1 (Dictyostelium)
203633_at	CPT1A	carnitine palmitoyltransferase 1A (liver)
201010_s_at	TXNIP	thioredoxin interacting protein
207256_at	MBL2	mannose-binding lectin (protein C) 2, soluble (opsonin defect)
209122_at	PLIN2	perilipin 2
225354_s_at	SH3BGRL2	SH3 domain binding glutamic acid-rich protein like 2
206292_s_at	SULT2A1	sulfotransferase family, cytosolic, 2A, dehydroepiandrosterone (DHEA)-preferring, member 1
1553181_at	DDX31	DEAD (Asp-Glu-Ala-Asp) box polypeptide 31

外部データで検証するため、NCBI GEO からデータをダウンロードし、この判別モデルを適用した。その結果、WY、リファンピシン(RIF)がそれぞれ陽性、陰性と正しく判別された。TG-GATEs のデータにおいて、シンバスタチン(SST)がメカニズム不明ながらラット *in vivo*, *in vitro*, ヒト肝細胞いずれにおいても陽性と判別されており、プラバスタチンが PPAR α を活性化するという報告(*Mol. Pharmacol.* **78**: 36-45, 2010)もあるので、GEO から同様にスタチン類であるアトルバスタチンとロスバスタチン (各 10 μ M) のデータもダウンロードして判別を行った。しかしこれらについては期待と異なり、判別結果は陰性であった。

D. 考察

リン脂質症マーカーについてはまずまず良い ROC カーブの判別モデルが得られ、学習セット以外のデータについても判別結果は一部 *in vivo* と一致していたこと、遺伝子の内容がリン酸や膜蛋白に関連するものが多く、膜構造にリン脂質が蓄積するリン脂質症のフェノタイプに照らして妥当であることなどから、一定の有用性のあるマーカーが得られたと考えられた。しかし内部データ、外部データ共に判別モデルの正しさを裏付ける検証データが未だ不足しており、信頼性をさらに高めていくためにはさらなる検証が必要である。

血液凝固不全マーカーについてもまずまず良好な ROC カーブを示す判別モデルが得られた。このモデルでは TG-GATEs のデータのみならず外部データについても正しく判別できており、有用性は高いと考えられたが、一部フェノタイプと一致しない部

分もあり、カットオフ値に近い Probability での判別結果に対しては留保してとらえる必要があると考えられた。

PPAR α 活性化マーカーに関しては、ラット *in vivo* モデル, *in vitro* モデルともに ROC カーブは極めて良好であり、内部データ、外部データともに判別結果も大変良好であった。極めて信頼性の高い判別モデルが得られたと考えられる。*In vivo*, *in vitro* の予測結果および無細胞系アッセイの結果の 3 者が一致しなかった化合物もいくつかあったが、*in vivo* では他の毒性のために PPAR α アゴニストとしては用量が不足していた (NSAIDs に多かった)、あるいは代謝されて初めてアゴニストとして働くために無細胞系アッセイでは陰性だった、など大部分は説明のつくものだった(表 4)。ASA, SST の 2 化合物のみ未だ陽性となるメカニズムが不明であるが、条件を変えて判別モデルを構築し直しても常に判別結果は陽性となったため、遺伝子発現の上では PPAR α アゴニストに極めて近いといえる。ASA に関しては主な代謝物であるサリチル酸も無細胞系アッセイで陰性であった。またプラバスタチンが PPAR α を活性化するという報告(*Mol. Pharmacol.* **78**: 36-45, 2010)があるため、SST についても確かに PPAR α を活性化することが予想される。これらのことから、これらの化合物は未知の代謝物が PPAR α アゴニストの活性を持つ、あるいは内因性の PPAR α アゴニストの産生を促進するのではないかと考えており、その検証は今後の課題である。

ヒト肝細胞の PPAR α 活性化判別モデルについても、精度はラットに及ばないものの、まずまず良好な ROC カーブが得られ、

TG-GATEs のデータについても外部データについても基本的に正しい判別結果が得られたと考えられたことから、良い判別モデルが構築できたといえる。スタチン類の外部データでは期待に反して判別結果は陰性であったが、用量が TG-GATEs の SST 高用量に比べて低かったこと、また判別に用いる Probability の数値が 0.2 付近であり、0 よりかなり大きかったこと (Probability > 0.5 を陽性と判定している) から、用量をもっと上げると陽性と判別されることが予想された。

PPAR α 活性化のフェノタイプについてはげっ歯類でのみ発がん性があることは良く知られており、ラットとヒトに種差があることは明らかである。今回構築した判別モデルに使用した遺伝子としてはラットとヒトのモデルに共通するものはなかった。しかしヒト肝細胞モデルのマーカのうち 3 遺伝子はラットモデルの出発点の 29 プロンプセットに含まれ、それ以外のものもラットでも変動のみられる脂質代謝関連遺伝子が大半であって、発現変動する遺伝子の内容自体が異なるということはない。ただしヒトの方が全体的に遺伝子発現変動が弱かった。また化合物の判別結果についてもヒトとラットで結果の一致しないものは特に見当たらなかった。種差の原因がどこにあるかは今後の検討課題といえる。

いずれの判別モデルについても、外部データに対しても有効という意味で一定の頑健性があると考えられたが、プラットフォームやデータ処理の方法については同一でないと正しい判別結果が得られないようであった。これについてはモデル構築に用いる発現データのノーマライズの方法を

RMA など分布を合わせるものに変えると異なるプラットフォームに適用できる可能性があると考えられる。

E. 結論

最終年度である今年度、本研究では *in vitro* のリン脂質症マーカーの抽出と検証、血液凝固不全マーカーの検証および PPAR α 活性化マーカーの検証を行った。方法としては、それぞれについて陽性化合物で変動する遺伝子を出発点とし、SVM による判別モデル構築後、TG-GATEs 内部のデータおよび公共データベースに登録された外部データについて判別を行い、その結果を検証した。その結果、血液凝固不全マーカーおよびラット *in vivo* 肝臓、初代培養肝細胞、ヒト初代培養肝細胞 PPAR α 活性化マーカーについてはいずれも、フェノタイプのアッセイデータで TG-GATEs のデータの判別結果が裏付けられたとともに、外部データでも正しい判別結果が得られ、信頼性・頑健性の高い判別モデルが作成できたと結論付けられる。*In vitro* のリン脂質症マーカーに関しては、その信頼性を裏付けるデータが未だ不足しているものの、一定の有用性のある判別モデルが作成できたと考えられた。さらに検証を進めることによるその信頼性の向上や一部未だ不明の PPAR α 活性化メカニズムの解明、PPAR α 活性化によるフェノタイプの種差の原因解明などは今後の課題である。

F. 研究発表

1. 論文発表
なし。

2. 学会発表

水川裕美子, 森川裕二, 中津則之, 小野敦,
大野泰雄, 山田弘, 漆谷徹郎「ラット肝ト
ランスクリプトームデータベースを用いた
ペルオキシソーム増殖因子活性化受容体
(PPAR) α アゴニストの検出とその検証」第
85回日本薬理学会年会, 2012年3月, 京都.

G. 知的所有権の取得状況

1. 特許取得

なし.

2. 実用新案登録

なし.

3. その他

なし.

2012年2月27日

平成23年度厚生労働科学研究費補助金

創薬基盤推進研究事業:創薬バイオマーカー探索研究事業

「トキシコゲノミクスデータベースを活用した毒性メカニズムに基づく
医薬品安全性評価に関する研究」発表会

バイオマーカー候補遺伝子の検証

同志社女子大学薬学部 病態生理学研究室

水川 裕美子

TGP1 (H14~H18) 大規模データベースの構築



TGP2 (H19~H23) バイオマーカー創出...目標

データベース (TG-GATEs)の解析



バイオマーカー候補遺伝子の抽出



検証

TG-GATEs内での検証

⇒施設間・プラットフォーム間での検証

メカニズムの解析

分担研究として一部を担当

検証を行ったバイオマーカー(グレードIV→III)

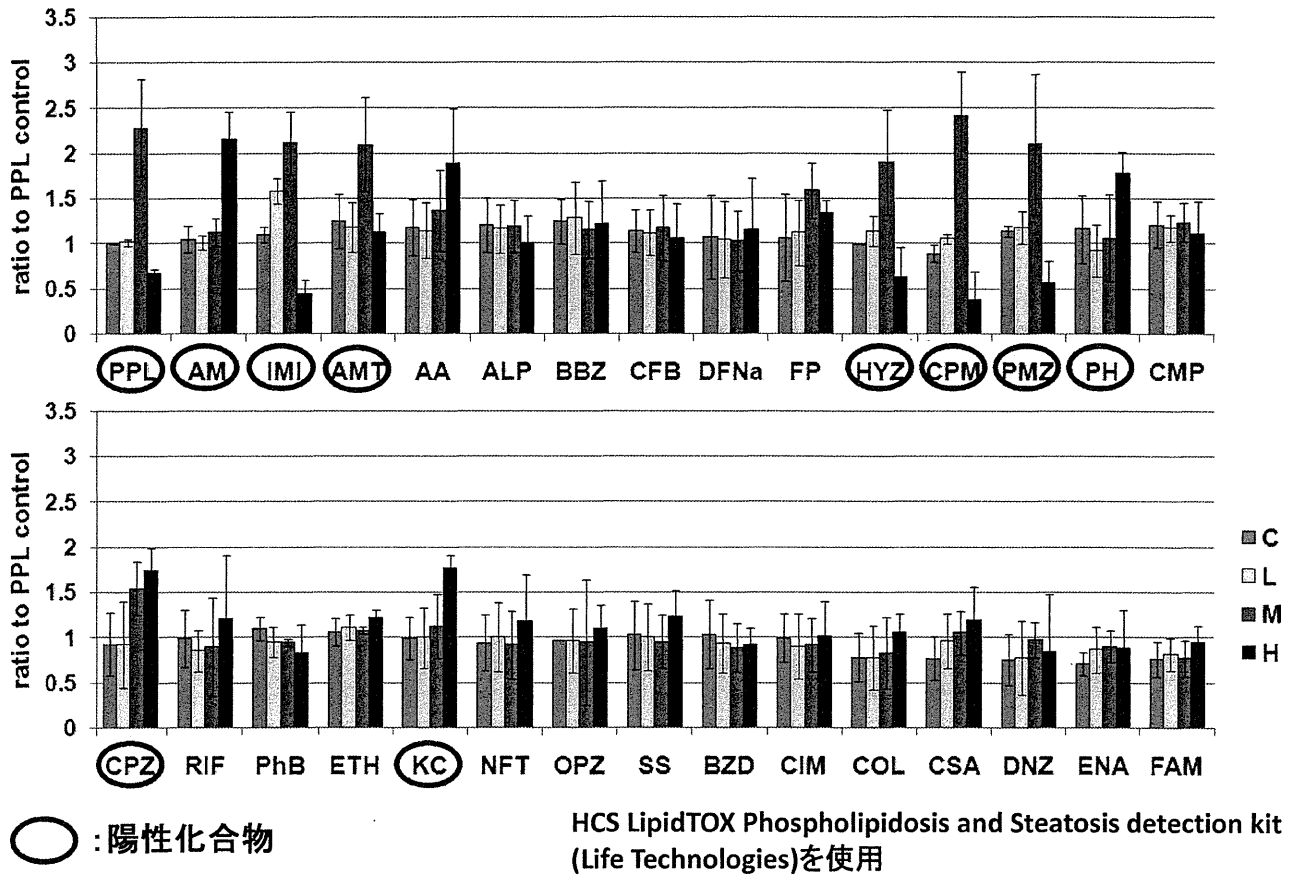
- グルタチオン(GSH)枯渇マーカー(ラット単回投与, 肝臓)
Kiyosawa *et al.*, *J. Toxicol. Sci.* 32, 469-486, 2007
Gao *et al.*, *Toxicol. App. Pharmacol.* 247, 211-221, 2010
- リン脂質症マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *Toxicol. App. Pharmacol.* 229, 290-299, 2008
- リン脂質症マーカー(ラット初代培養肝細胞)
- 血液凝固不全 マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *J. Toxicol. Sci.*, 34, 281-293, 2009
- PPAR α アゴニストマーカー(ラット単回肝臓, ラット肝細胞, ヒト肝細胞)
Tamura *et al.*, *J. Toxicol. Sci.* 31, 471-490, 2006

検証を行ったバイオマーカー(グレードIV→III)

- グルタチオン(GSH)枯渇マーカー(ラット単回投与, 肝臓)
Kiyosawa *et al.*, *J. Toxicol. Sci.* 32, 469-486, 2007
Gao *et al.*, *Toxicol. App. Pharmacol.* 247, 211-221, 2010
- リン脂質症マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *Toxicol. App. Pharmacol.* 229, 290-299, 2008
- リン脂質症マーカー(ラット初代培養肝細胞)
- 血液凝固不全 マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *J. Toxicol. Sci.*, 34, 281-293, 2009
- PPAR α アゴニストマーカー(ラット単回肝臓, ラット肝細胞, ヒト肝細胞)
Tamura *et al.*, *J. Toxicol. Sci.* 31, 471-490, 2006

リン脂質症検出 *in vitro* アッセイの結果

mean ± SD



マーカー抽出の方法

ラット初代培養肝細胞の遺伝子発現データにおいて、
リン脂質症を引き起こす化合物で共通に変動する遺伝子を抽出

使用化合物: amiodarone (AM), imipramine (IMI), amitriptyline (AMT), hydroxyzine (HYZ), clomipramine (CPM), promethazine (PMZ), perhexiline (PH)

化合物処置: 24h曝露, 対照群+3用量

データ取得: Affymetrix Rat Genome 230 2.0 array, MAS 5.0を使用.

データ処理: 全プローブセットのmeanを1とする(global (mean)補正).

max intensity < 0.01 または all Absent のプローブセットを除外.

対照群 vs 各用量で R を用いて rank product を計算し, $p < 0.05$ のプローブセットを抽出.

7化合物中4化合物以上で同方向に共通変動した遺伝子を選択.

up:

463

down:

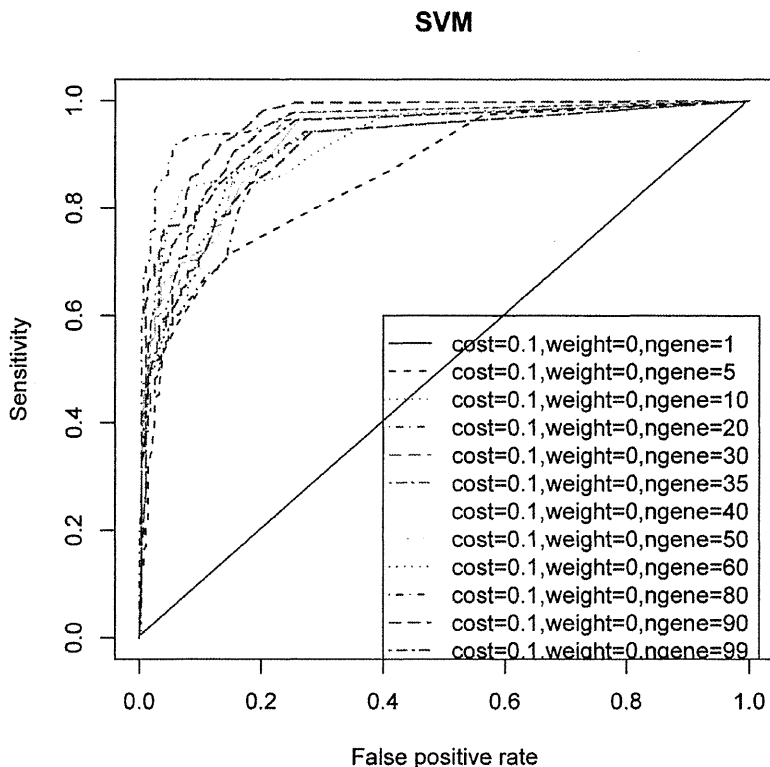
737

Support vector machine (SVM)を用いた判別モデル構築

上記7化合物を陽性, 32化合物を陰性として学習. \log_2 ratio 値 (vs 対照群)を用いた. 判別モデルのランキングに基づき, support vector数が最小となる遺伝子数を選択. Affymetrixのannotation grade Aのもののみ抽出して判別マーカーとする.

判別モデルの学習セットとROCカーブ

化合物略称	化合物名	ラベル	時点	用量
AMT	amitriptyline	N	24h	M
IMI	imipramine	N	24h	L,M
HYZ	hydroxyzine	N	24h	M
CPM	clomipramine	N	24h	M
PMZ	promethazine	N	24h	M
AM	amiodarone	N	24h	H
KC	ketoconazole	N	24h	H
PH	perhexiline	N	24h	H
MDP	methyldopa	N	2,8,24h	LMH
PCT	phenacetin	N	2,8,24h	LMH
RAN	ranitidine	N	2,8,24h	LMH
TAC	tacrine	N	2,8,24h	LMH
TC	tetracycline	N	2,8,24h	LMH
TRI	triamterene	N	2,8,24h	LMH
CBZ	carbamazepine	N	2,8,24h	LMH
FFB	fenofibrate	N	2,8,24h	LMH
INAH	isoniazid	N	2,8,24h	LMH
TAA	thioacetamide	N	2,8,24h	LMH
DOX	doxorubicin	N	2,8,24h	LMH
FUR	furosemide	N	2,8,24h	LMH
SUL	sulindac	N	2,8,24h	LMH
VPA	valproic acid	N	2,8,24h	LMH
COL	colchicine	N	2,8,24h	LMH
CMP	chloramphenicol	N	2,8,24h	LMH
APL	allopurinol	N	2,8,24h	LMH
ETH	ethionamide	N	2,8,24h	LMH
NFT	nitrofurantoin	N	2,8,24h	LMH
CIM	cimetidine	N	2,8,24h	LMH
BZD	benziodarone	N	2,8,24h	LMH
BBZ	bromobenzene	N	2,8,24h	LMH
FAM	famotidine	N	2,8,24h	LMH
PhB	phenylbutazone	N	2,8,24h	LMH
RIF	rifampicin	N	2,8,24h	LMH
SS	sulfasalazine	N	2,8,24h	LMH
DNZ	danazol	N	2,8,24h	LMH
CSA	cyclosporine A	N	2,8,24h	LMH
OPZ	omeprazole	N	2,8,24h	LMH
DFNa	diclofenac	N	2,8,24h	LMH
CFB	clofibrate	N	2,8,24h	LMH
ENA	enalapril	N	2,8,24h	LMH



抽出されたマーカー遺伝子

以下の72プローブセット

1368491_at Dnase2b	1386967_at Rhoq	1367712_at Timp1	1399096_at Add3
1387659_at Gda	1368695_at C4bbp	1373258_at Ctsf	1368191_a_at Slc22a1
1389746_at RGD1564228	1372675_at RGD1306954	1374806_at Sfn	1392534_at Pmepa1
1384428_at Tmem181	1386893_at Grn	1390860_at Igf2bp3	1372806_at Vps35
1373939_at Nagk	1387024_at Dusp6	1375560_at Dhdspl	1373245_at Col4a1
1370475_at Cyp2b3	1374877_at LOC687295	1376249_at Fuca2	1389559_at RGD1564420
1372304_at Commd4	1370789_a_at Prrl	1371946_at RGD1303130	1377049_at Pnpla7
1368074_at Gale	1371021_at Arsb	1379255_at Atp6ap2	1370959_at Col3a1
1381982_at Uap111	1370949_at ---	1372809_at LOC290595	1388547_at Cldn4
1367514_at LOC361635	1389106_at Fbxw9	1372897_at Plod2	1398930_at Atp6v0b
1398254_at Renbp	1386884_at Htra1	1370310_at Hmgcs2	1393170_at RGD1561431
1371871_at Rab12	1370280_at Hprt1	1368309_at Txnrd2	1389123_at Ccl6
1367562_at Sparc	1390848_at Rbm19	1372741_at Sccpdh	1396279_at Atp6ap2
1367913_at Cygb	1375972_at LOC100360533 / RGD1307493	1370155_at Col1a2	1374396_at Atp6v1c1
1378196_at Slc43a1	1388675_at Spns1	1371143_at Serpina7	1369029_at Plscr1
1398363_at Wdr54	1391435_at Pltp	1373970_at Il33	1368741_at C9
1368718_at Aldh1a7	1372013_at Ifitm1	1371357_at Igfbp7	1373617_at Emp2
1372687_at Crct1	1389581_at Il33	1392547_at MGC105649	

リン酸化や細胞内小胞輸送, トランスポーターに関連したものが多い。

TG-GATEsデータの判別結果-1

化合物略称	ラベル	化合物名	Low: 24hr	Middle: 24hr	High: 24hr
AMT		amitriptyline	0.907	0.986	1.000
IMI		imipramine	0.930	0.965	1.000
HYZ		hydroxyzine	0.349	0.912	1.000
CPM		clomipramine	0.461	0.751	1.000
PMZ		promethazine	0.301	0.882	1.000
AM		amidarone	0.861	0.966	0.987
KC		kefconazole	0.020	0.101	0.934
PH		perhexiline	0.081	0.356	0.912
MDP		methylidopa	0.019	0.008	0.019
PCT		phenacetin	0.009	0.003	0.010
RAN		ranitidine	0.001	0.000	0.008
TAC		tacrine	0.006	0.050	0.006
TC		tetracycline	0.003	0.005	0.004
TRI		triamterene	0.000	0.001	0.001
CBZ		carbamazepine	0.000	0.000	0.000
FFB		fenofibrate	0.001	0.000	0.000
INAH		isoniazid	0.000	0.000	0.000
TAA		thioacetamide	0.000	0.000	0.000
DOX		doxorubicin	0.000	0.000	0.000
FUR		furosemide	0.002	0.000	0.000
SUL		sulindac	0.003	0.002	0.000
VPA		valproic acid	0.000	0.000	0.000
COL		colchicine	0.050	0.000	0.000
CMP		chloramphenicol	0.009	0.037	0.036
APL		allopurinol	0.012	0.050	0.025
ETH		ethionamide	0.012	0.023	0.018
NFT		nitrofurantoin	0.006	0.035	0.014
CIM		cimetidine	0.008	0.015	0.012
BZD		benziodarone	0.003	0.001	0.012
BBZ		bromobenzene	0.001	0.004	0.008
FAM		famotidine	0.001	0.002	0.006
PhB		phenylbutazone	0.000	0.000	0.004
RF		rifampicin	0.003	0.001	0.003
SS		sulfasalazine	0.035	0.010	0.001
DNZ		danazol	0.002	0.001	0.001
CSA		cyclosporine A	0.003	0.001	0.000
OPZ		omeprazole	0.000	0.000	0.000
DFNa		diclofenac	0.000	0.000	0.000
CFB		clofibrate	0.000	0.000	0.000
ENA		enalapril	0.000	0.000	0.000

化合物略称	ラベル	化合物名	Low: 24hr	Middle: 24hr	High: 24hr
CCL4		carbon tetrachloride	0.000	0.000	NA
DIL		diltiazem	0.006	0.033	0.966
SLP		sulpiride	0.009	0.016	0.931
EBU		ethambutol	0.002	0.037	0.913
FP		fluphenazine	0.030	0.044	0.842
TMD		trimethadione	0.868	0.815	0.824
ADP		adapin	0.050	0.161	0.819
GF		griseofulvin	0.059	0.628	0.745
CHL		chlorpheniramine	0.001	0.015	0.713
LBT		labetalol	0.025	0.013	0.671
QND		quinidine	0.036	0.858	0.513
K09		K09	0.016	0.025	0.372
GBC		glibenclamide	0.150	0.063	0.356
TMX		tamoxifen	0.053	0.282	0.341
AJM		ajmaline	0.004	0.004	0.293
CPX		ciprofloxacin	0.370	0.111	0.250
K11		K11	0.005	0.008	0.248
TZM		triazolam	0.139	0.089	0.192
K10		K10	0.052	0.091	0.175
K03		K03	0.005	0.005	0.172
BDZ		bendazac	0.391	0.411	0.167
TRZ		thioridazine	0.019	0.006	0.113
HPL		haloperidol	0.004	0.015	0.098
FT		flutamide	0.017	0.011	0.084
DIS		disopyramide	0.000	0.009	0.068
TCP		ticlopidine	0.031	0.029	0.064
BCT		bucefin	0.024	0.042	0.054
MTS		methyltestosterone	0.005	0.025	0.053
NPAA		phenylanthranilic acid	0.014	0.016	0.031
EE		ethinylestradiol	0.114	0.074	0.029
K14		K14	0.045	0.024	0.027
MXS		moxisylyte	0.010	0.013	0.024
K13		K13	0.014	0.016	0.024
CLM		chlormadinone	0.001	0.000	0.024
LS		lomustine	0.019	0.014	0.019
NIM		nimesulide	0.003	0.002	0.018
LNX		lornoxicam	0.014	0.012	0.016
K07		K07	0.013	0.003	0.015
K06		K06	0.001	0.001	0.014
ACZ		acetazolamide	0.008	0.016	0.014
K15		K15	0.001	0.002	0.013

数字はprobability, ピンク: ポジティブ判定(probability>0.5)

TG-GATEsデータの判別結果-2

化合物略称	ラベル	化合物名	Low: 24hr	Middle: 24hr	High: 24hr
WY		WY-14643	0.006	0.003	0.011
PML		pemoline	0.018	0.010	0.009
MFM		metformin	0.001	0.001	0.007
DTL		dantrolene	0.003	0.005	0.007
DSF		disulfiram	0.086	0.162	0.006
DEN		nitrosodiethylamine	0.003	0.001	0.006
ET		ethionine	0.007	0.002	0.006
K05		K05	0.002	0.001	0.005
PHE		phenytoin	0.004	0.001	0.003
GFZ		gemfibrozil	0.013	0.010	0.003
IPA		iproniazid	0.007	0.018	0.003
AAF		acetamidofluorene	0.003	0.003	0.003
AA		allyl alcohol	0.002	0.003	0.003
NFZ		nitrofurazone	0.066	0.080	0.003
K08		K08	0.005	0.004	0.002
K04		K04	0.011	0.052	0.002
K02		K02	0.002	0.001	0.002
MEF		mefenamic acid	0.004	0.002	0.002
MLX		meloxicam	0.000	0.001	0.002
CMN		chlormezanone	0.000	0.000	0.002
BEA		bromoethanamine	0.005	0.008	0.002
CPZ		chlorpromazine	0.000	0.000	0.002
K12		K12	0.000	0.001	0.002
BBF		benzbromarone	0.004	0.001	0.001
VMC		vancomycin	0.002	0.001	0.001
EME		erythromycin ethylsuccinate	0.002	0.002	0.001
K01		K01	0.000	0.000	0.001
NIF		nifedipine	0.002	0.002	0.001
HCB		hexachlorobenzene	0.001	0.001	0.001
K16		K16	0.000	0.000	0.001
TBF		terbinafine	0.001	0.000	0.001
CAF		caffeine	0.031	0.009	0.001
CBP		carboplatin	0.001	0.000	0.001
CMA		coumarin	0.000	0.001	0.001
AZP		azathioprine	0.007	0.012	0.001

化合物略称	ラベル	化合物名	Low: 24hr	Middle: 24hr	High: 24hr
MCT		monocrotaline	0.001	0.001	0.001
ACA		acarbose	0.005	0.003	0.000
ETN		ethanol	0.002	0.001	0.000
CPP		chlorpropamide	0.000	0.000	0.000
CAP		captopril	0.004	0.010	0.000
CPA		cyclophosphamide	0.000	0.000	0.000
ASA		aspirin	0.000	0.000	0.000
MEX		mexiletine	0.000	0.000	0.000
NIC		nicotinic acid	0.000	0.000	0.000
SST		simvastatin	0.001	0.001	0.000
ETP		etoposide	0.000	0.000	0.000
TLB		tolbutamide	0.010	0.001	0.000
TIO		tiopronin	0.000	0.000	0.000
ANIT		naphthyl isothiocyanate	0.001	0.002	0.000
VA		vitamin A	0.000	0.000	0.000
DZP		diazepam	0.001	0.001	0.000
TEO		theophylline	0.010	0.022	0.000
		methapyrilene	0.000	0.000	0.000
IBU		ibuprofen	0.060	0.003	0.000
MTX		methotrexate	0.000	0.000	0.000
PEN		penicillamine	0.000	0.000	0.000
IM		indomethacin	0.000	0.000	0.000
PTU		propylthiouracil	0.001	0.000	0.000
MTZ		methimazole	0.000	0.000	0.000
PAN		puromycin aminonucleoside	0.000	0.000	0.000
PB		phenobarbital	0.000	0.000	0.000
CSP		cisplatin	0.002	0.000	0.000
PAP		papaverine	0.001	0.002	0.000
APAP		acetaminophen	0.000	0.003	0.000
NPX		naproxen	0.002	0.000	0.000
GMC		gentamicin	0.001	0.015	0.000
K17		K17	0.003	0.000	0.000
CLT		cephalothin	0.000	0.000	0.000
TAN		tannic acid	0.002	0.002	0.000

外部データの判別結果

化合物略称	化合物名	ラベル	Low: 24hr	Middle: 24hr	High: 24hr
AMT	amitriptyline		0.978	0.991	1.000
IMI	imipramine		0.915	0.972	1.000
HYZ	hydroxyzine		0.226	0.835	1.000
AM	amiodarone		0.955	0.995	1.000
PMZ	promethazine		0.266	0.819	0.994
CPM	clomipramine		0.279	0.721	0.986
KC	ketoconazole		0.008	0.034	0.886
PH	perhexiline		0.080	0.307	0.835
MDP	methylidopa		0.022	0.004	0.010
TAC	tacrine		0.005	0.019	0.007
RAN	ranitidine		0.000	0.000	0.002
TC	tetracycline		0.001	0.003	0.002
PCT	phenacetin		0.002	0.000	0.001
TRJ	triamterene		0.000	0.000	0.000
CBZ	carbamazepine		0.000	0.000	0.000
DOX	doxorubicin		0.000	0.000	0.000
IHAH	isoniazid		0.000	0.000	0.000
FFB	fenofibrate		0.000	0.000	0.000
TAA	thioacetamide		0.000	0.000	0.000
SUL	sulindac		0.001	0.001	0.000
FUR	furosemide		0.000	0.000	0.000
VPA	valproic acid		0.000	0.000	0.000
COL	colchicine		0.031	0.000	0.000
CMP	chloramphenicol		0.004	0.014	0.024
APL	allopurinol		0.016	0.031	0.018
NFT	nitrofurantoin		0.006	0.031	0.013
BBZ	bromobenzene		0.001	0.003	0.012
BZD	benzodiarone		0.003	0.001	0.009
ETH	ethionamide		0.001	0.005	0.006
CIM	cimebidine		0.003	0.004	0.004
FAM	famotidine		0.000	0.001	0.002
RIF	rifampicin		0.000	0.000	0.001
PhB	phenylbutazone		0.000	0.000	0.001
SS	sulfasalazine		0.023	0.005	0.001
DNZ	danazol		0.001	0.001	0.000
CSA	cyclosporine A		0.001	0.000	0.000
DFNa	diclofenac		0.000	0.000	0.000
OPZ	omeprazole		0.000	0.000	0.000
CFB	clofibrate		0.000	0.000	0.000
ENA	enalapril		0.000	0.000	0.000
CCL4	carbon tetrachloride		0.000	0.000	NA
AM	amiodarone	na	0.000	na	na

外部データ:

NCBI GEO Accession No. GSE19662

AM 1 ppm (約1.55 μM)

ラット肝細胞に24h曝露

Agilent-014879 Whole Rat Genome

Microarray 4x44K G4131Fのデータ.

ノーマライズされたデータをダウンロードし、

対照群に対するlog₂ ratioにして解析

※ AM TGPでの用量: 0.28, 1.4, 7 μM

黄色が外部データ.

数字はprobability.

ピンク: ポジティブ判定(probability>0.5)

← 陰性と判定された

小括

- *In vitro*のラット肝細胞のデータからリン脂質症を判別する判別モデルをSVMを用いて作成した.
- アッセイ結果の判別の性能はますます良好だが、フェノタイプが弱い場合は判別が難しい.
- 外部データでの検証はプラットフォームが同一のデータが見つからず、うまくいかなかった.
- 抽出された遺伝子にはリン酸化や細胞内小胞輸送, トランスポーターに関連したものが多く, 生物学的にも妥当なものであった.

検証を行ったバイオマーカー(グレードIV→III)

- グルタチオン(GSH)枯渇マーカー(ラット単回投与, 肝臓)
Kiyosawa *et al.*, *J. Toxicol. Sci.* 32, 469-486, 2007
Gao *et al.*, *Toxicol. App. Pharmacol.* 247, 211-221, 2010
- リン脂質症マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *Toxicol. App. Pharmacol.* 229, 290-299, 2008
- リン脂質症マーカー(ラット初代培養肝細胞)
- 血液凝固不全マーカー(ラット反復投与, 肝臓)
Hirode *et al.*, *J. Toxicol. Sci.*, 34, 281-293, 2009
- PPAR α アゴニストマーカー(ラット単回肝臓, ラット肝細胞, ヒト肝細胞)
Tamura *et al.*, *J. Toxicol. Sci.* 31, 471-490, 2006

グレードIVマーカー抽出, 評価の方法

反復投与のラット肝臓の遺伝子発現データにおいて, 血液凝固不全を引き起こす化合物で共通に変動する遺伝子を抽出

使用化合物: 血液生化学検査でPT延長, APTT延長, Fbg低下のいずれかがみられた化合物

clofibrate (CFB), omeprazole (OPZ), ethionine (ET), thioacetamide (TAA), benzbromarone (BBR), propylthiouracil (PTU), WY-14643 (WY), amiodarone (AM)

化合物処置: 3, 7, 14, 28日間反復経口投与, 対照群+3用量

サンプリング: 最終投与24時間後に肝臓抽出

データ取得: Affymetrix Rat Genome 230 2.0 array, MAS 5.0を使用.

データ処理: global (mean)補正, all Absentのプローブセットを除外.

各化合物について, 時点ごとにANOVAで $p < 0.05$ の遺伝子を抽出して全時点union. 8化合物共通変動遺伝子のうち, Affymetrixのannotation grade Aのものに絞った.

↓
主成分分析(PCA)で評価

上記8化合物の抽出した遺伝子(344プローブセット)のデータをZ-score化してPCAを行った.