

参考 Circadian Rhythmに影響を与える
化学物質のランダムスクリーニングに関する報告
Hirota T et al, PLoS Biol 8: e1000559, 2010

High-Throughput Chemical Screen Identifies a Novel Potent Modulator of Cellular Circadian Rhythms and Reveals CKI α as a Clock Regulatory Kinase

Tsuyoshi Hirota^{1,2}, Jae Wook Lee^{2,3}, Warren G. Lewis^{1,2}, Eric E. Zhang^{1,2}, Ghislain Breton¹, Xianzhong Liu², Michael Garcia², Eric C. Peters², Jean-Pierre Etchegaray⁴, David Traver¹, Peter G. Schultz^{2,3}, Steve A. Kay^{1*}

1 Division of Biological Sciences, University of California San Diego, La Jolla, California, United States of America, **2** Genomics Institute of the Novartis Research Foundation, San Diego, California, United States of America, **3** Department of Chemistry, The Scripps Research Institute, La Jolla, California, United States of America, **4** Department of Neurobiology, University of Massachusetts Medical School, Worcester, Massachusetts, United States of America

Figure 1. Effect of a novel compound on the circadian period in cultured cells and tissues. A: The effect of various concentrations of compound 1 on the period of circadian oscillations in a cell-based reporter system. The heatmap shows bioluminescence over 96 hours. B: The effect of various concentrations of compound 1 on the period of circadian oscillations in a cell-based reporter system. C: The effect of various concentrations of compound 1 on the period of circadian oscillations in adult tail fibroblasts. D: The effect of various concentrations of compound 1 on the period of circadian oscillations in lung. E: The effect of various concentrations of compound 1 on the period of circadian oscillations in the SCN. R=H: Compound 1; R=CF₃: Longdaysin.

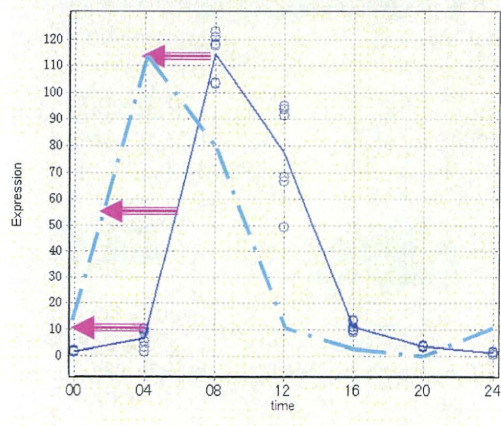
Dbp遺伝子が関わる局所ネットワークを制御する
さらに上流の解析 [肝]

- KLF結合サイトが転写開始点上流に存在

Dbp遺伝子が関わる局所ネットワークを制御する
さらに上流の解析 [肝]

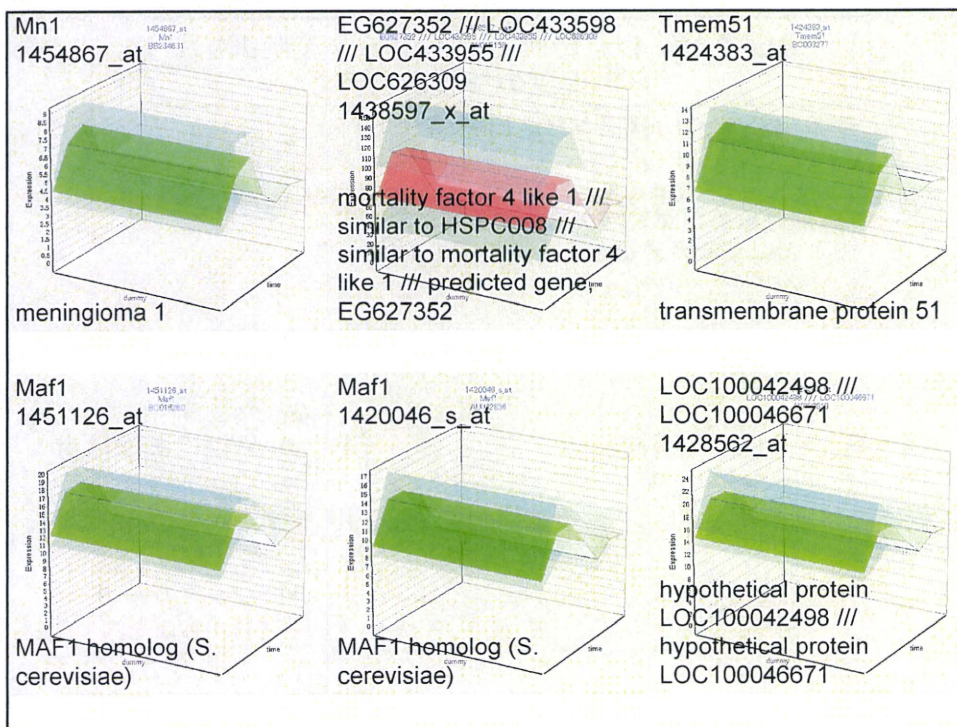
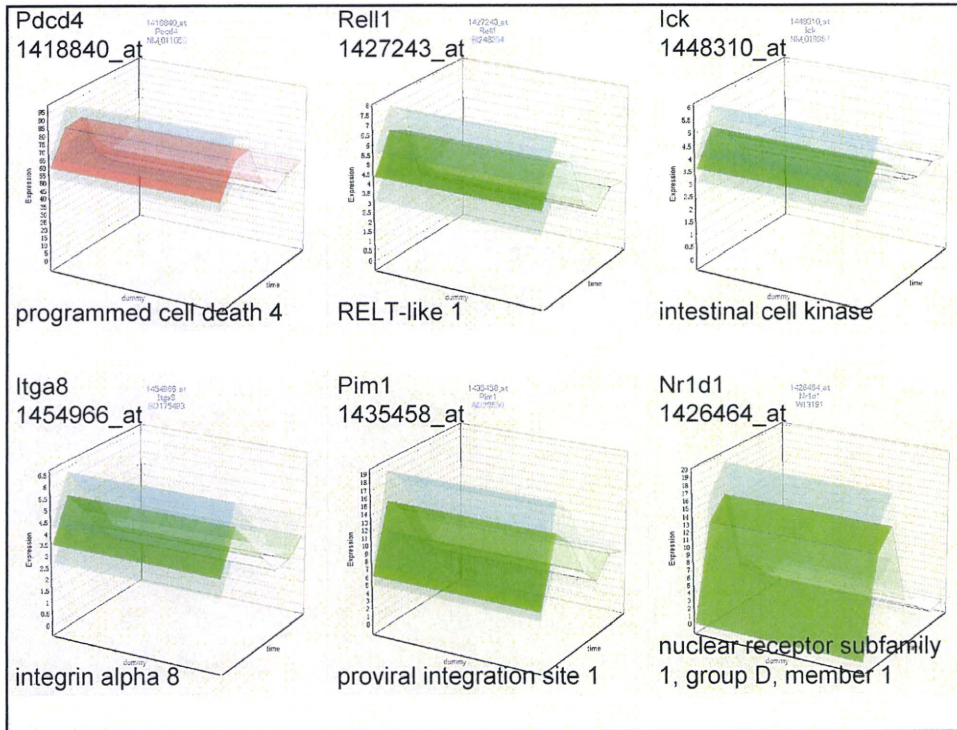
- KLF結合サイトが転写開始点上流に存在

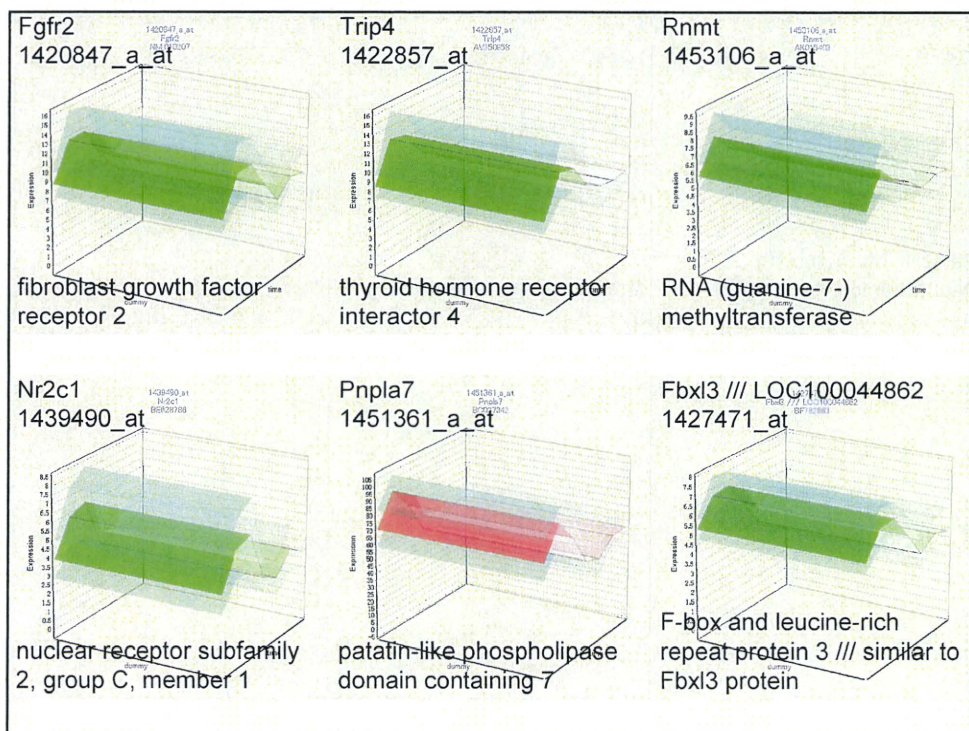
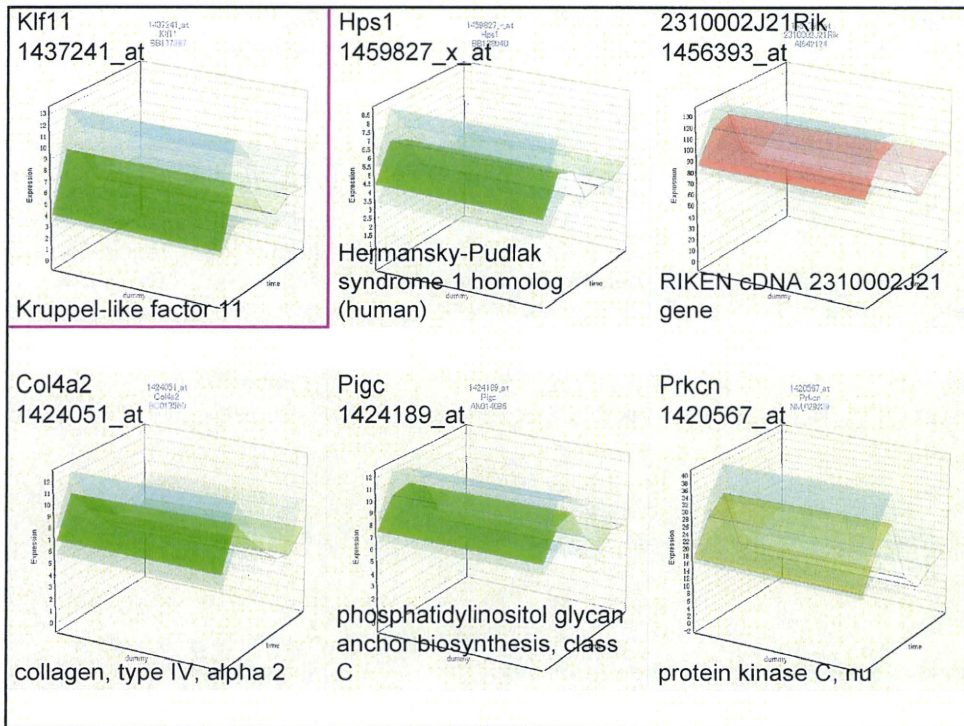
4時間分早めた発現パターンを雛形にして
これと似た経時変化を示す遺伝子の抽出

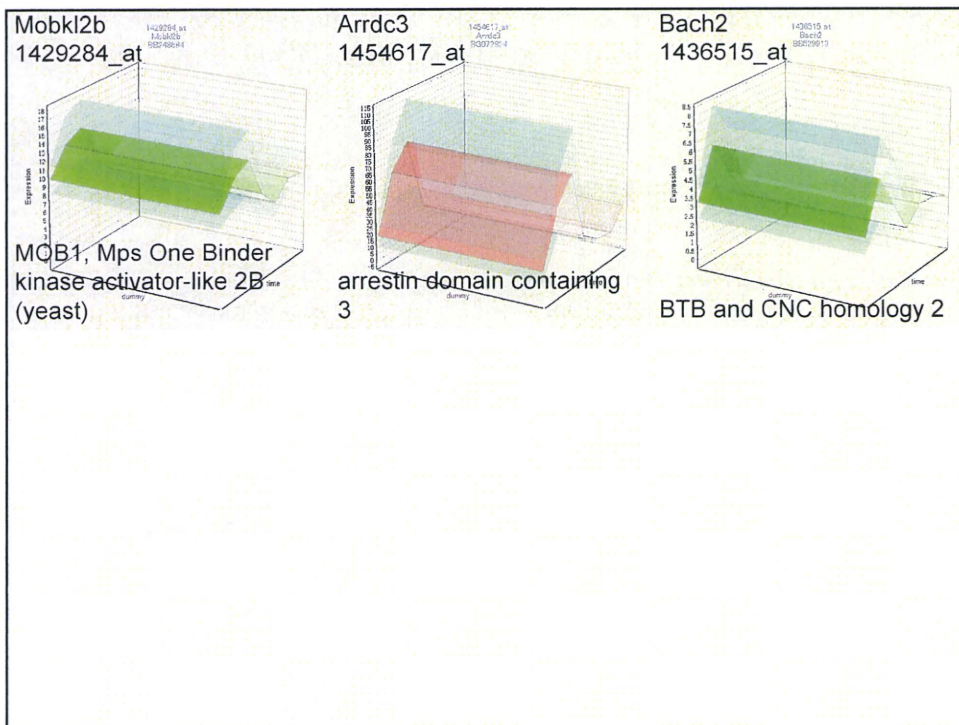
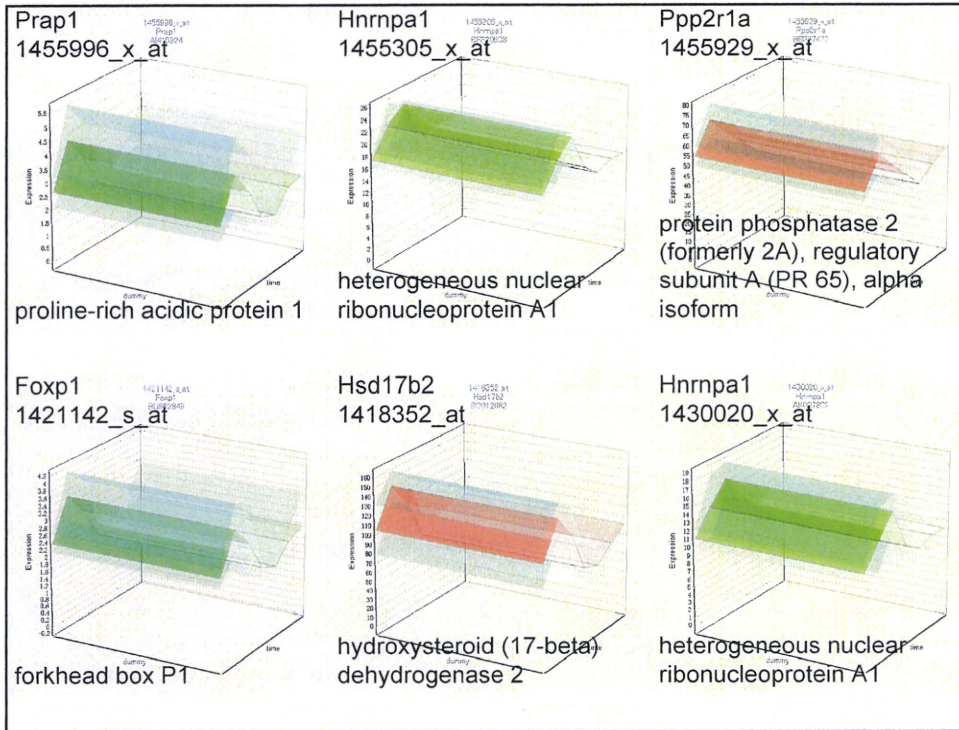


97 ps
⇒ 目視による検討
⇒ 33 ps

KLFファミリーの
Klf11が含まれていた







4時間分早めた発現パターンを雛形にして
これと似た経時変化を示す遺伝子の抽出

97 ps
⇒ 目視による検討
⇒ 33 ps

KLFファミリーの
Klf11が含まれていた

*in silico*でのプロモーター解析

市販のGenomatixを利用

Liver

by Dr Igarashi, K

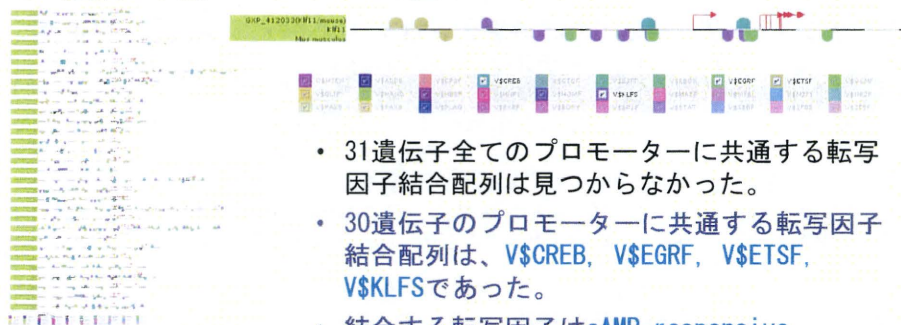
— 4時間分早めた発現パターンを雛形とした際、
似た発現パターンを示す遺伝子について—

Genomatix

33 probe sets → 31遺伝子

30遺伝子に共通

Klf11のみ拡大



- 31遺伝子全てのプロモーターに共通する転写因子結合配列は見つからなかった。
- 30遺伝子のプロモーターに共通する転写因子結合配列は、V\$CREB, V\$EGRF, V\$ETSF, V\$KLFSであった。
- 結合する転写因子はcAMP-responsive element binding proteins, EGR/nerve growth factor induced protein C & related factors, ETS1 factors, Krueppel like transcription factors

4時間分早めた発現パターンを雛形とした際の似た経時変化を示す遺伝子の抽出 & Dbp遺伝子が関わる局所ネットワーク [Liver]

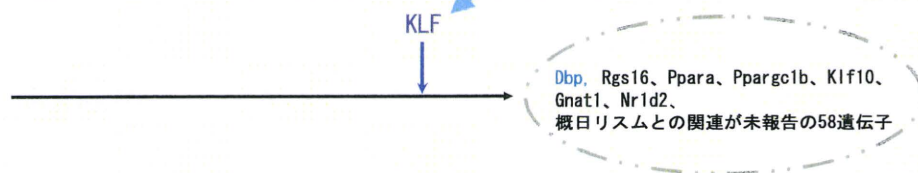
*in silico*でのプロモーター解析のまとめ

転写開始点上流に共通した結合配列

肝 抽出されてきた31遺伝子中30遺伝子



肝 抽出されてきた65遺伝子中64遺伝子



まとめ2

■ Percellome data baseの利用

・肝においてDbp遺伝子の発現を強く抑制する化学物質：
 化審法白物質2-クロロ-4,6-ジメチルアニリン[0, 3, 10, 30 mg/kg]、
 有機リン系殺虫剤アセフェート[0, 7, 20, 70 mg/kg]、
 食品添加物・保存料デヒドロ酢酸ナトリウム[0, 30, 100, 300 mg/kg]

・肝においてDbp遺伝子と似た経時変化を示す「81 ps」の遺伝子の発現変動に対する2-クロロ-4,6-ジメチルアニリンの影響
 ⇒多くの遺伝子の発現について、2-クロロ-4,6-ジメチルアニリンの
 投与影響が認められた = 抽出した遺伝子リストは妥当

■ さらに制御上流のネットワークの探索

4時間分早めた発現パターンを雛形とした際の、似た経時変化を示す遺伝子の抽出
 ピアソン相関係数を利用し (α 値0.9以上) 抽出、さらに目視による検討により生物学的な変化が示唆されたもの：

肝：33 ps (31 遺伝子) KLFファミリーのKlf11を含む

⇒ *in silico*でのプロモーター解析：

CREB、EGR family、ETS1及びKLFの結合配列が見いだされた

⇒これらの制御下でKlf11が発現しDbp遺伝子の関与する局所ネットワークの発現が引き起こされる事が示唆された

厚生労働科学研究費補助金（化学物質リスク研究事業）
化学物質の有害性評価手法の迅速化、高度化に関する研究
—網羅的定量的大規模トキシコゲノミクスデータベースの維持・拡充
と毒性予測評価システムの実用化の為にインフォマティクス技術開発—
(H21-化学—一般-001)

分担研究報告書

【トキシコゲノミクスのためのインフォマティクス開発】

複数の化学物質のクラスタ交差点からの遺伝子発現ネットワーク抽出

研究分担者 相崎 健一

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 毒性部 第二室長

研究要旨

既に 100 余の化学物質による遺伝子発現変動情報を収録し、現在もさらに拡張しつつある Percellome データベースを有効活用して、特に複数の化学物質のクラスタ交差点からの遺伝子発現ネットワーク抽出を実現するための基本技術開発を行う。H22 年度は特徴的な発現変動パターンの自動抽出プログラム RSort を改良し抽出効率・精度を大幅に向上させた。また RSort 解析結果を利用した総当たりの遺伝子発現反応類似度解析を容易に実現するためのプログラム PercellomeExplorer を作成してクラスタ交差点解析に供した。

A. 研究目的

全遺伝子の発現データから生物学的に有意な情報を効率的に抽出し、これを元に遺伝子発現誘導カスケードを描出する技術を生じ、安全性評価におけるトキシコゲノミクスの有機的活用促進の一助とする。特に、複数の化学物質暴露により共通して誘導・抑制される遺伝子群に着目し、毒性に関わる主たる要素を分離すると共に、それらの連関の解明を目指す。

B. 研究方法

我々は遺伝子発現データの絶対量化技術である Percellome 手法を考案し、これに最適化した” Millefeuille”（ミルフィーユ）データ処理システムを開発している。

開発・改良したアルゴリズムは、研究者本人がソフトウェアを作成し、実際に本研究計画で得られた生データをサンプルに用いて、ワークステーション上で性能検証を行う。これは本研究におけるプログラム開

発はアルゴリズムの改良・微調整を兼ねた try and error 方式で進めなければならず、職業プログラマに開発委託する際に必要となる仕様の事前決定が出来ないためである。開発言語には、研究的なプログラミングに適した、つまり生成したプログラムの実行速度が早いだけでなく、実行プログラムの生成（コンパイル速度）も速い Delphi 言語（=object pascal 言語，開発環境としては Borland Delphi7 もしくは Delphi2007, Delphi2010）を使用した。また一般的なグラフ描画には TeeChart コンポーネント（Steema 社）を利用し、100 万件以下の小規模なデータベース操作には組み込み型リレーショナルデータベースコンポーネントである DBISAM（ElevateSoftware 社）を利用して、プログラム開発の効率化を図った。

大規模計算が必要な場合は、専用計算サーバーシステム（MF サーバーシステム）にアプリケーションソフトウェアを移植して、大規模計算用グリッド PC や高速データベースエンジンである Teradata データベース（日本 Teradata 社）を活用して計算処理を行った。

C. 研究結果

本年度新たに測定された遺伝子発現変動データと未処理であった既存の遺伝子発現変動データを 3 次元平面波（Surface グラフ）に変換し、その凹凸の多寡を評価して整ったものから複雑なものへ並び替える RSort アルゴリズム（先行研究で独自に開発）を利用して、3 次元平面の凹凸が少ないパターンを示す遺伝子、つまり、より強力もしくは よりユニークな発現制御を受

ける遺伝子を化学物質および臓器ごとに網羅的に自動抽出した。

本年度は、RSort.exe のソースコードを精査・改良して旧版の判定条件値自動調節ルーチンで発生していた誤差を抑え、また新たな抽出条件として最大ピークだけでなく、第 2、第 3 ピークの参照機能を追加した。この結果、false negative が減り候補遺伝子の自動抽出精度が向上した他、最大ピークが用量相関を示しているも、第 2、3 ピークで解釈不能の挙動を示す発現パターン（複雑な制御を受けているもの、もしくはノイズ）を排除することが出来るようになり、RSort を利用する化学物質暴露による遺伝子誘導パターンの類似度比較（RSortMatching）等の解析精度・効率を大幅に向上させた。

また煩雑な RSortMatching 等のリスト類似度計算を数クリックの操作で簡便に行い、用量・時間情報を加味した比較散布図の描画や主成分解析までを行うプログラム PercellomeExplorer.exe を作成し、誰でも Percellome データベース全体を対象とした膨大かつ複雑な解析処理を容易に実行できるようにした。

PercellomeExplorer の開発とあわせ、RSortMatching 計算の対象範囲を広げ、肝、肺、腎に加えて、心、海馬、大脳皮質、脳幹、小脳、精巣、胸腺等を含めることとし、計 270 の Surface データセットについて最新の RSort.exe を用いて有意な挙動を示す候補遺伝子の自動抽出処理を実施し、延べにして約 20 万件の候補遺伝子情報を得た。

D. 考察

RSort アルゴリズムの実装アプリケーションソフトウェア RSort.exe は一層の改良を進めており、本年度は抽出効率と精度の大幅な向上を実現した。訓練された研究者の目視による判定にはまだ及ばないが、従来発生していた false positive を半減させることに成功した。

本年度の研究では Surface データセットを昨年度比 61 増の 270 用いた。これらから得られた候補遺伝子の延べ数は前述の通り 20 万と、昨年度の 58 万個から半減以下となった。これは RSort.exe の改良により false positive の混入を大幅に減らした成果であり、RSortMatching の精度を格段に改善させるだけでなく、同解析のデータ処理スピードを大幅に向上させた。

候補遺伝子群同士の類似度比較は、原理的に要素数の多いノード (RSortMatching の場合は各 Surface データセットから抽出した候補遺伝子群を要素とする塊) が有利になり、場合によっては解析結果に悪影響を及ぼすことが懸念されていた。

本年度開発した PercellomeExplorer.exe を用いて検証を行ったところ、要素数と正に相関して解析結果へのノードの出現頻度が上昇する傾向があったが、各ノードの要素 (候補遺伝子群) の内容を精査したところ、各ノードが由来する化学物質暴露実験の結果を適切に反映していることが確認された。つまり要素数が多くても、全く性質の異なる遺伝子群の発現を誘導する場合、類似リストの上位には出現せず、妥当な解析結果が得られていた。

今後は、バルプロ酸、クロフィブレート、

DEHP・MEHP など多要素ノード間の比較解析も進め、毒性に関わる遺伝子発現シグナルネットワークの解析 (特に回路の使い分けについて) を進める。

E. 結論

先行研究により世界有数の規模・精度のトキシコゲノミクスデータベースとなった Percellome データベースを用いることで、化学物質暴露時に誘導される遺伝子情報を総合的に捉えることが出来るようになってきた。特に本年度作成した PercellomeExplorer ソフトウェアを用いれば、IT 専門家ではない生物学・毒性学の研究者であっても、化学物質暴露による生体影響を遺伝子発現レベルで検索し、研究・試験に容易に活用出来るようになった。実際に本プログラムの利用を以て従来無関係と思われていた化学物質同士が、生体影響レベルで類似した遺伝子発現誘導能を有する例が多数見いだされ、その情報を毒性試験・研究計画の立案に反映させるなど、既に実用となっている。

この技術の幅広い活用を促進すべく、来年度は同ソフトウェアのオンラインアプリケーション化を含め、活発な情報公開と啓蒙活動を予定し、研究成果の早期還元を目指す。

F. 研究発表

1. 論文発表

Yoshida T, Sekine T, Aisaki K, Mikami T, Kanno J, Okayasu I., CITED2 is activated in ulcerative colitis and induces p53-dependent apoptosis in

response to butyric acid., J Gastroenterol 2011 46(3): 339-49.

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. 2011 84(4):734-42.

2. 学会発表

Natalia Polouliakh, Jun Kanno, Yukiko Matsuoka, Ken-Ichi Aisaki, Richard Nock, Frank Nielsen, Keigo Oka, Satoshi Kitajima and Hiroaki Kitano, Discovery of Gene Network Regulated by the Toxicity Equivalent Factor of 2, 3, 7, 8-Tetrachlorodibenzo-p-Dioxin (TCDD) and 2, 3, 7, 8-Tetrachlorodibenzofuran (TCDF) chemicals. the 11th International Conference on System Biology (2010.10.11) (Edinburgh, UK) Poster

菅野 純、相崎健一、北嶋 聡、Percellome トキシコゲノミクスの抗がん剤研究への応用、第 69 回日本癌学会学術総会 (2010.9.24) (大阪) 口演

菅野 純、北嶋 聡、高橋祐次、五十嵐勝秀、相崎健一 インフォマティクス局面にある Percellome トキシコゲノミクスの食品・食品添加物への適用、第 37 回日本トキシコロジー学会 学術年会

(2010.6.16) (沖縄) 口演

高木篤也、北嶋 聡、五十嵐勝秀、相崎 健一、江馬 眞菅野 純 Percellome 手法を用いた TCDD 投与マウスの胎児口蓋の遺伝子発現解析 (3)、第 37 回日本トキシコロジー学会 学術年会(2010.6.17) (沖縄) 口演

G. 知的財産権の出願・登録状況 (予定を含む。)

1. 特許取得

特許出願 2010 年 12 月、「競合的ハイブリダイゼーションにおける遺伝子データの補正方法及び補正装置」

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金(化学物質リスク研究事業)
化学物質の有害性評価手法の
迅速化、高度化に関する研究
—網羅的定量的大規模トキシゲノミクスデータベースの維持・拡充と
毒性予測評価システムの実用化の為にインフォマティクス技術開発—
(H21-化学-一般-001)

平成23年2月1日
於. 東京国際フォーラム

分担研究課題

複数の化学物質のクラスタ交差点からの 遺伝子発現ネットワーク描出

相 崎 健 一
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
毒性部

1

トキシゲノミクス応用による毒性予測・評価

(完成型=毒性シグナルネットワークの全容解明後)

- ① 候補化学物質の暴露影響を受ける遺伝子群の抽出
- ② 毒性シグナルネットワーク全情報と照合
- ③ 反応シグナルのin silico合成
- ④ 毒性予測・評価

2

トキシコゲノミクス応用による毒性予測・評価

(暫定型＝毒性シグナルネットワークの全容解明前)

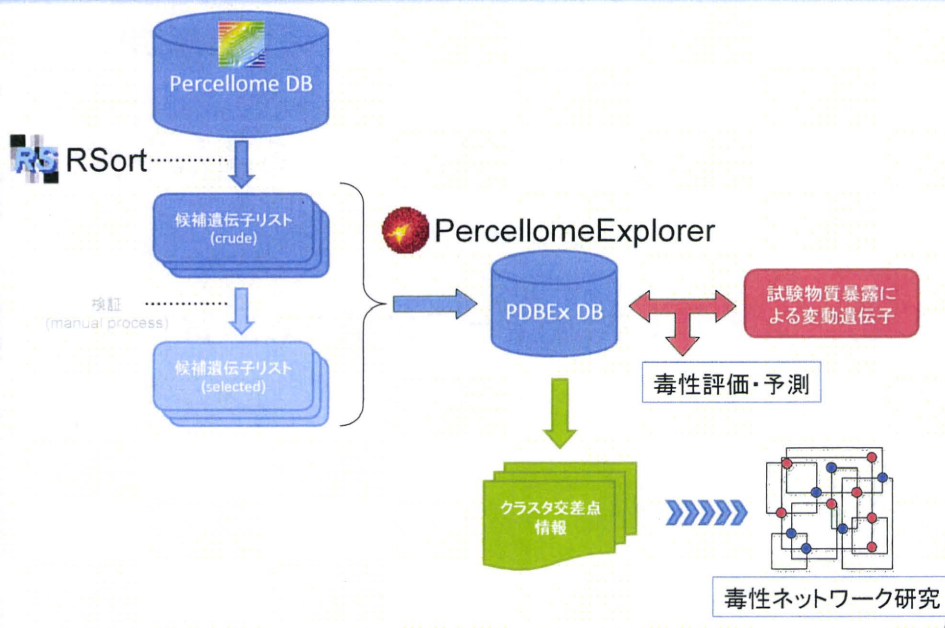
- ① 候補化学物質の暴露影響を受ける遺伝子群の抽出
- ② 基盤DB (PercellomeDB)と照合、類似反応を呈する化学物質を抽出
- ③ 影響を受ける遺伝子や類似反応を呈する化学物質の既知情報より、反応シグナルを推測

④ 毒性評価・推測

>>>> ②, ③から、毒性シグナルネットワークの描出を進める

3

類似反応解析を中心とした研究・試験システム



4

RSort

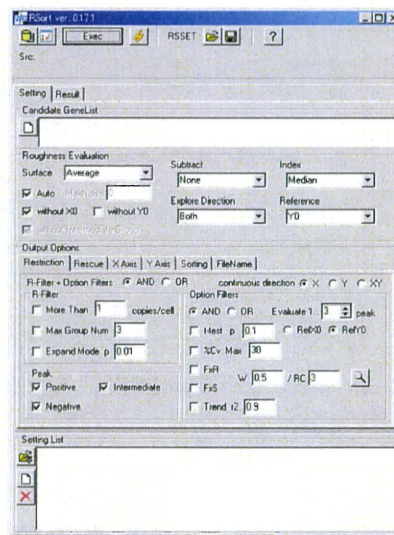
特徴的な遺伝子発現パターンの自動抽出

5

RSort ver.0.17.1

発現パターンの凹凸の程度・特徴点の位置(=実験条件)を評価する。

→ 特定の実験条件において、強い発現誘導パターンを呈する遺伝子を抽出できる。



6

RSort ver.0.17.1

Median Surface対応
(外れ値対応)

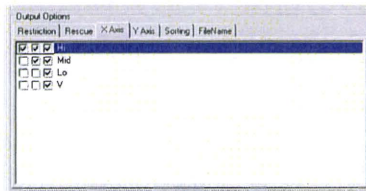
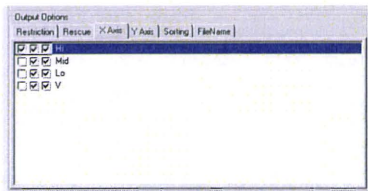
2nd, 3rdピーク対応

RSort ver.0.17.1

ピーク数 ≤ 2 , t-test $p < 0.05$

1stピーク Hiのみ

1stピーク Hiのみ
および2ndピークHiまたはMid



Average Surface: 1605.1 ± 1267.7 ps
Median Surface: 1416.7 ± 1099.9 ps

Average Surface: 869.8 ± 815.8 ps
Median Surface: 745.6 ± 682.1 ps

約50%の絞り込み効果

PercellomeExplorer

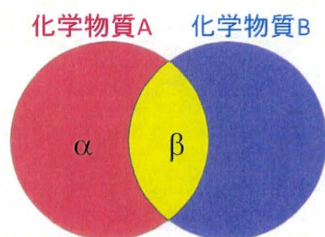
リスト類似度解析用ソフトウェア

9

PercellomeExplorer

類似度の定義

化学物質AのCandidate GeneList
化学物質BのCandidate GeneList



→ 化学物質Aの化学物質Bに対する類似度
 $\beta/(\alpha + \beta)$

PercellomeExplorer ver.0.2.7では、さらに発現誘導の正負や発現コピー数、最大変動時点などにより、化学物質の比較・絞り込みが可能

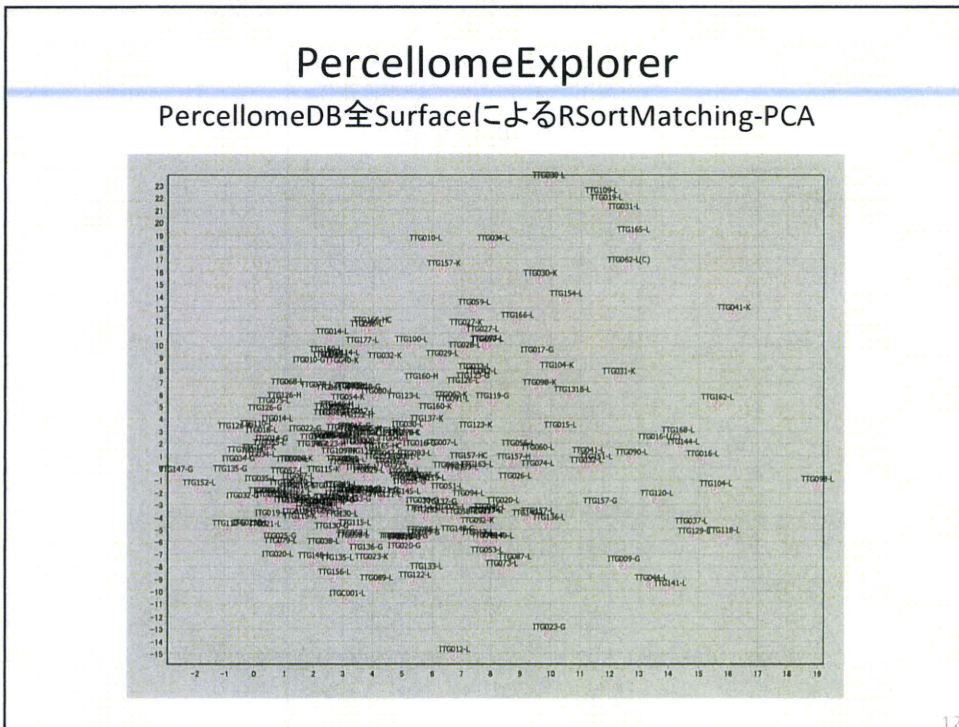
10

PercellomeExplorer

The screenshot displays the Percellome Explorer interface. The main window shows a list of cell lines with columns for PCellID, Name, Condition, CP, GL, Descripto, Surface, and Tissue. A list of conditions is shown below, including DEHP, Tributhin + Clofibrate, Phenobarbital, Clofibrate, Clofibrate + Clofibrate, Valproic Acid, Tributhin + Phenobarbital, OCH + Clofibrate, and Acetaldehyde (2hr x 1day). To the right, a scatter plot matrix compares cell lines for two chemical substances (A and B) at four time points (2hr, 4hr, 8hr, 24hr). A legend indicates that red squares represent positive induction and negative induction, blue squares represent negative induction and negative induction, and grey squares represent other. A gene information window is also visible, showing details for the gene 1418099_at (M60469), including its description as a tumor necrosis factor receptor superfamily member and its biological processes.

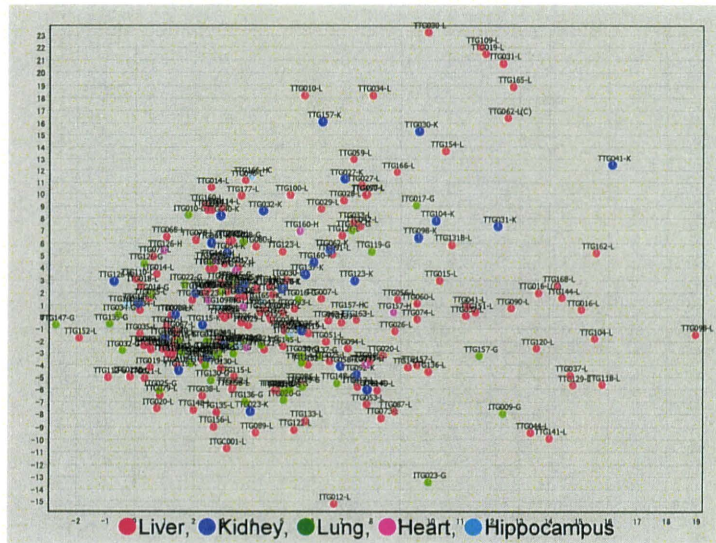
PercellomeExplorer

PercellomeDB全SurfaceによるRSortMatching-PCA



PercellomeExplorer

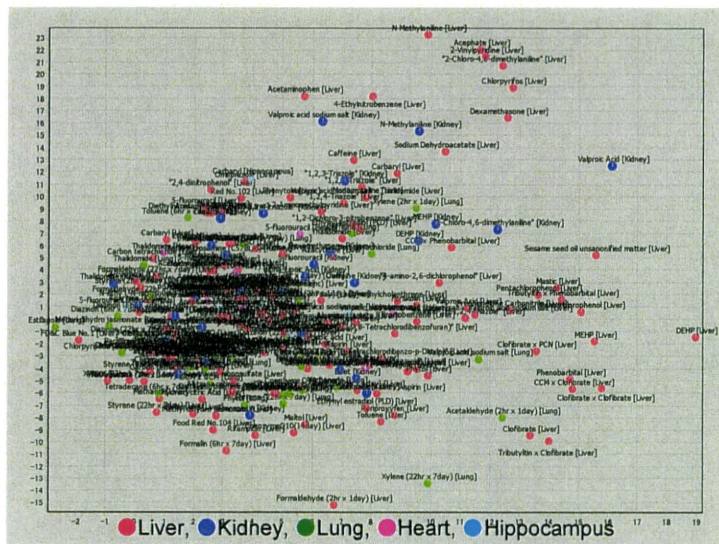
PercellomeDB全SurfaceによるRSortMatching-PCA



13

PercellomeExplorer

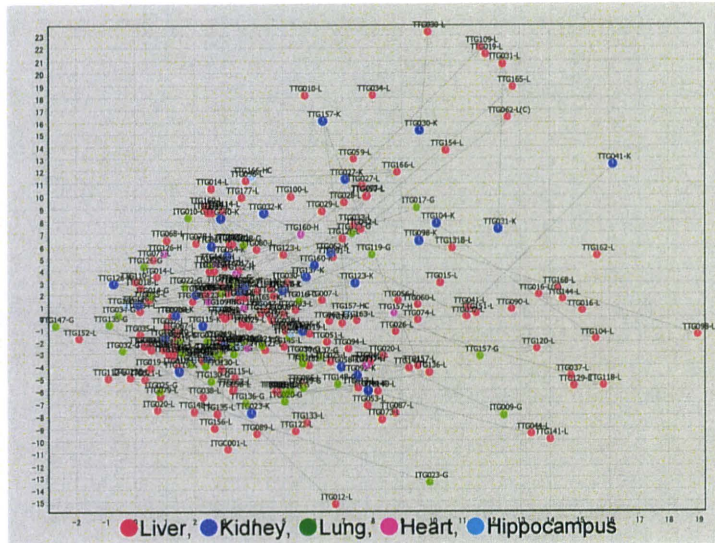
PercellomeDB全SurfaceによるRSortMatching-PCA



14

PercellomeExplorer

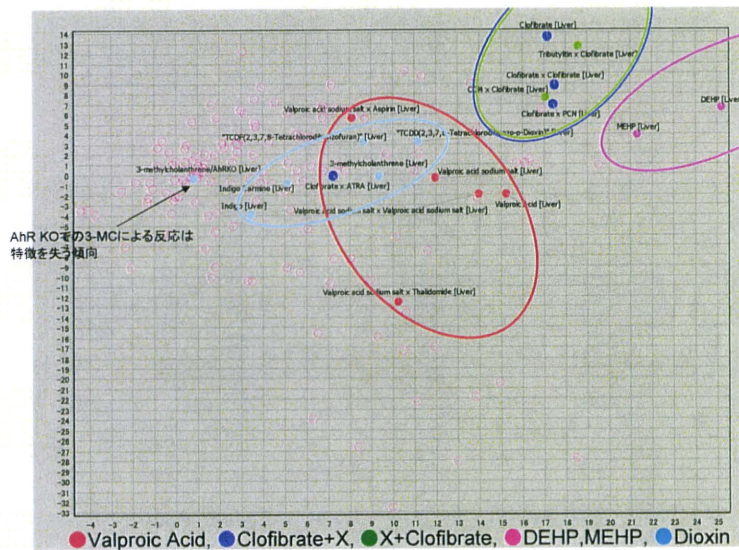
PercellomeDB全SurfaceによるRSortMatching-PCA



15

PercellomeExplorer

経口暴露および吸入暴露実験の肝SurfaceによるRSortMatching-PCA



16