

Table 2: Primers for Q-PCR

Gene	Forward	Reverse
AFFX-TrpnX-3_at	TTCTCAGCGTAAAGCAATCCA	GCAAATCCTTTAGTGACCGAATACC
AFFX-DapX-3_at	TCAGCTAACGCTTCCAGACC	GGCCGACAGATTCTGATGACA
AFFX-PheX-3_at	GCCAAATGATATGGCAGCTTCTAC	TGCGGCAGCATGACCATTA
AFFX-LysX-3_at	CCGCTTCATGCCACTGAATAC	CCGGTTCGATCCAAATTTCC
AFFX-ThrX-3_at	CCTGCATGAGGATGACGAGA	GGCATCGGCATATGGAAAC
Ahr_1450695_at	CAGAGACCACTGACGGATGAA	AGCCTCTCCGGTAGCAAACA
Cyp1a1_1422217_a_at	TGCTCTTGCCACCTGCTGA	GGAGCACCTGTTTGTCTATG
Cyp1a2_1450715_at	CCTCACTGAATGGCTTCCAC	CGATGGCCGAGTTGTTATTG
Cyp1b1_1416612_at	GCCTCAGGTGTGTTGATGGA	AGTACAGCCCTGGTGGGAATG
Cyp7a1_1422100_at	TTTACATGCCCTTTGGATCAG	GGACACTTGGTGTGGCTCTC
Hspa1a_1452388_at	ACCATCGAGGAGGTGGATTAGA	AGGACTTGATTGCAGGACAAAC

The "LBM" ("liver-brain mix") standard sample

A pair of samples having dissimilar gene expression profiles was chosen to evaluate the linearity of the platform. The pairs chosen were brain and liver for mouse and rat, two distinct cancer cell lines for humans, and adult liver and embryo for *Xenopus laevis*. The sample pairs were processed as described above including addition of the GSC. Two final homogenates were then blended at ratios of 100:0, 75:25, 50:50, 25:75 and 0:100 (based on cell numbers) to make five samples. These five samples were measured by Q-PCR and/or GeneChips (MG-U74v2A, MEA430A, MEA430B, MG430 2.0 (shown in Figure 1), RAE230A, HG-U95A, HG-U133, and Xenopus array).

Quantitative-PCR

Duplicate homogenate samples were treated with DNaseI (amplification grade, Invitrogen Corp., Carlsbad, CA, USA) for 15 min at room temperature, followed by SuperScript II (Invitrogen) for 50 min at 42°C for reverse transcription. Quantitative real time PCR was performed with an ABI PRISM 7900 HT sequence detection system (Applied Biosystems, Foster City, CA, USA) using SYBR Premix Ex Taq (TAKARA BIO Inc., Japan), with initial denaturation at 95°C for 10 s followed by 45 cycles of 5 s at 95°C and 60 s at 60°C, and Ct values were obtained. Primers for the genes explored in this study were selected from sequences close to the areas of Affymetrix GeneChip probe sets as shown in Table 2.

Affymetrix GeneChip measurement

The sample homogenates with GSC added were processed by the Affymetrix Standard protocol. The GeneChips used were MG-U74v2A for the uterotrophic study and Mouse 430-2 for the TCDD study (singlet measurement). The efficiency of *in vitro* transcription (IVT) was monitored by comparing the values of 5' probe sets and 3' probe sets of the control RNAs (AFFX- probe sets) including the GSC (see Quality Control below). The dose-response linearity of the five GSC spikes was checked and samples showing saturation and/or high background were re-measured

from either backup tissue samples, an aliquot of homogenate, or a hybridization solution, depending on the nature of the anomaly.

Quality control

Any external spiking method, including our Percellome method, is valid for high-quality RNA samples. Therefore, the quality of the sample RNA should be carefully monitored. In addition to a common checkup by RNA electrophoresis (including capillary electrophoresis if necessary), OD ratio, and cRNA yield, we monitor the performance of IVT (*in vitro* translation) or amplification. The 3' and 5' probe set data of the spiked-in RNAs and sample RNAs (actin, GAPD and other AFFX- probe sets) that are prepared in Affymetrix GeneChip are compared to monitor the extension of RNA by the IVT process. When both the spiked-in RNAs and the sample RNAs have similar levels of 5' and 3' signals respectively, it is judged that the IVT extension was normally performed. When both spiked-in and sample RNAs have significantly lower 5' signal than 3' signal, it is judged that the IVT extension was abnormal. When only the sample RNAs showed significantly lower 5' signal than 3' signal, it is judged that the IVT extension was normal but the sample RNAs were degraded. When only the spiked-in RNAs showed significantly lower 5' signal than 3' signal, it is judged that the IVT extension was normal but the spiked-in RNAs were degraded (although we have not encountered this situation). In addition, if the degraded sample was spiked-in by the non-degraded spike RNAs and measured by GeneChip, the position of spiked-in RNAs will be offset toward abnormally higher intensity. Together, this battery of checkups considerably increases the ability to detect abnormal events that will affect the reliability of the Percellome method. When any abnormality was found, each step of sample preparation was reevaluated to regain normal data for Percellome normalization.

The web site for GeneChip data

The GeneChip data are accessible at http://www.nihs.go.jp/tox/TTG_Archive.htm.

Authors' contributions

JK drafted the concept of the Percellome method, led the project at a practical level, and drafted the manuscript. KA developed the algorithm for the Percellome calculation and wrote the calculation/visualization programs. KI developed the laboratory protocols for the Percellome procedures to the level of SOP for technicians. NN developed the Percellome Q-PCR protocol and performed the measurements, and helped in analyzing the Percellome data. AO helped develop the algorithm. YK led the animal studies. TN provided advice and led the toxicogenomics project using the Percellome method, to be approved by the Ministry of Health, Labour and Welfare of Japan.

Additional material**Additional File 1**

Excel spreadsheet file containing 15 Affymetrix Mouse 430-2 GeneChip raw data of five LBM samples in triplicate (cf. Figure 1). The column name LBM-100-0-X_Signal indicates the component percentages, i.e. 100% liver 0% brain, and X = 1,2,3 indicates the triplicates. The LBM-100-0-X_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S1.zip>]

Additional File 2

Excel spreadsheet file containing Percellome data of the same LBM samples, of which raw data is listed in Additional file 1 (cf. Figure 1).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S2.zip>]

Additional File 3

Excel spreadsheet file containing 2 Affymetrix MG-U74v2 raw data of a blank sample with the GSC (horizontal axis of Figure 2a) and blank with the five spike RNAs at a high dosage (vertical axis of Figure 2a).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S3.zip>]

Additional File 4

Excel spreadsheet file containing 2 Affymetrix MG-U74v2 raw data of a liver sample with GSC (horizontal axis of Figure 2b) and without GSC (vertical axis of Figure 2b).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S4.zip>]

Additional File 5

(first quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 2 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S5.zip>]

Additional File 6

(second quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 4 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S6.zip>]

Additional File 7

(third quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 8 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S7.zip>]

Additional File 8

(last quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 24 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S8.zip>]

Additional File 9

(first quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 2 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 5.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S9.zip>]

Additional File 10

(second quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 4 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 6.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S10.zip>]

Additional File 11

(third quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 8 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 7.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S11.zip>]

Additional File 12

(last quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation): an Excel spreadsheet file containing 24 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 8.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S12.zip>]

Additional File 13

Excel spreadsheet file containing 15 Affymetrix MG-U74v2 A GeneChip raw data of the uterotrophic response study (cf. Figure 6). The column name X-Y_Signal indicates the treatment (V = vehicle, Low = low dose, etc) and animal triplicate (Y = 1,2,3). The X-Y_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S13.zip>]

Additional File 14

Excel spreadsheet file containing Percellome data of the same 15 samples of the uterotrophic response study (cf. Figure 6), of which raw data is listed in Additional file 13.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S14.zip>]

Acknowledgements

The authors thank Tomoko Ando, Noriko Moriyama, Yuko Kondo, Yuko Nakamura, Maki Abe, Nae Matsuda, Kenta Yoshiki, Ayako Imai, Koichi Morita, Hisako Aihara and Chiyuri Aoyagi for technical support, and Dr. Bruce Blumberg and Dr. Thomas Knudson for critical reading of the manuscript. This study was supported by Health Sciences Research Grants H13-Seikatsu-012, H13-Seikatsu-013, H14-Toxico-001 and H15-Kagaku-002 from the Ministry of Health, Labour and Welfare, Japan.

References

- Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA: **Dissecting the regulatory circuitry of a eukaryotic genome.** *Cell* 1998, **95**:717-728.
- Hill AA, Brown EL, Whitley MZ, Tucker-Kellogg G, Hunter CP, Sloss DM, Kim DK: **Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls.** *Genome Biol* 2001, **2**: RESEARCH0055
- Lee PD, Sladek R, Greenwood CM, Hudson TJ: **Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies.** *Genome Res* 2002, **12**:292-297.
- van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, Holstege FC: **Monitoring global messenger RNA changes in externally controlled microarray experiments.** *EMBO Rep* 2003, **4**:387-393.
- Yang YH, Dudoit S, Luu P, Lin DM, Peng W, Ngai J, Speed TP: **Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation.** *Nucleic Acids Res* 2002, **30**:e15.
- Hekstra D, Taussig AR, Magnasco M, Naef F: **Absolute mRNA concentrations from sequence-specific calibration of oligonucleotide arrays.** *Nucleic Acids Res* 2003, **31**:1962-1968.
- Sterrenburg E, Turk R, Boer JM, van Ommen GB, den Dunnen JT: **A common reference for cDNA microarray hybridizations.** *Nucleic Acids Res* 2002, **30**:e116.
- Dudley AM, Aach J, Steffen MA, Church GM: **Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range.** *Proc Natl Acad Sci USA* 2002, **99**:7554-7559.
- Talaat AM, Howard ST, Hale W, Lyons R, Gamer H, Johnston ST: **Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis.** *Nucleic Acids Res* 2002, **30**:e104.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**:185-193.
- Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL: **Expression monitoring by hybridization to high-density oligonucleotide arrays.** *Nat-Biotechnol* 1996, **14**:1675-1680.
- Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens W: **The OECD program to validate the rat uterotrophic bioassay. Phase 2: dose-response studies.** *Environ Health Perspect* 2003, **111**:1530-1549.
- Kanno J: **Reverse toxicology as a future predictive toxicology.** In *Toxicogenomics* Edited by: Inoue T, Pennie ED. Tokyo, Springer-Verlag; 2002:213-218.

毒性学

毒性の高精細解析に向けてのトキシコゲノミクス

Toxicogenomics for high-resolution toxicology

毒性学は生体と化学物質との相互作用を研究する分野であり、目的は“ヒトの安全”である。日常遭遇する化学物質(医薬品や食品を含む)が摂取された際の安全性を担保するため(毒性評価)に、実験動物の毒性所見をヒトに外挿することが行われてきた。これは実験動物もヒトも基本的に同等の生体反応を示すという前提に基づいている。そして呼吸鎖阻害、DNA損傷、活性酸素種生成など、非特異的な標的がおもな対象となってきた。これに加え、現在の分子毒性学は、生体反応メカニズムに踏み込み、受容体、転写因子などとの選択的結合によるシグナル伝達障害など標的特異性の高いものや、エピジェネティックな遅発影響などを対象とするようになり、基礎分子生物学と直結する時代に入っている。

医薬品の開発ストラテジーも標的特異性が高いものを追うものとなっており、医薬品審査行政の面からも同じことが毒性学に要求されている。そしてヒトボランティアに対するバイオ医薬品の微量投与が全員をICU送りにした最近の事件は、種差を含む分子毒性評価の重要性とその現状を示していると考えられる。また、インターネット上にはバイオ医薬品紛いの効能を謳いあげるサプリメントやダイエット/健康食品が氾濫している。これらについても場合によっては医薬品と同等の慎重さをもって評価にあたる必要がある。

胎児や子どもといった、ボランティアによる毒性評価がほとんど不可能と思われる対象のみならず、ヒトの安全のための毒性学には今後とも動物実験が必要であると考えられ、その高精度化のひとつの手段としてトキシコゲノミクス

は必須であるとの認識に変わりはない。著者らは定量的PCRやマイクロアレイから細胞1個当りのmRNAコピー数を得るPercellome法¹⁾と、そのデータ解析のためのMillefeuilleシステムを開発し、この3年間でマウス肝を中心とした約90化合物(毒性学的情報の豊富なもの)の単回曝露による初期応答遺伝子データベースを構築するとともに、シックハウス症候群を考慮した低用量域での吸入トキシコゲノミクス、発生毒性についての胎児トキシコゲノミクス、また*in situ* hybridizationによる局在確認系を立ち上げた。今後の3年間で、反復曝露による慢性毒性、および、多臓器間の関連性を検討する研究を行う。

今後の課題

トキシコゲノミクスの有効性がいまだに発揮されていないのかんばしくない評価を耳にすることがある。しかし、著者らの経験から、それには、すくなくとも2つの解決可能な理由があると考えられる。

第1に、mRNAの変動は、思いのほか速いので、いままで以上に実験の管理を厳重にしなければ有意なデータが得られない点である。マウス肝で有意に発現する12,000(いずれかの時点で3コピー/細胞以上)の遺伝子のうち、概日変動を示す遺伝子が3,600以上あり、激しい遺伝子では数十倍以上の発現差がみられる。まず、動物飼育施設の明暗サイクルを2週間以上一定に保ちマウスの概日リズムを安定化する。そのうえで、マウスに検体を投与する時刻、および、サンプルを採取する時刻を±30分以内に限定することで始めて正確なデータが得られ

る。肺や腎も同様の注意が必要である。第2に、蛋白、とくにリン酸化をはじめとする蛋白合成を伴わない早い反応とのリンケージが取れないとの指摘である。この問題点についてもPercellome法を用いた高精度解析によって間接的ながら解決策を提供できることが示されつつある。すなわち、リン酸化により制御される転写因子の働きが、時間差はあるものの、その遺伝子発現の変動として捕捉されている。

おわりに

従来の毒性学に対してのトキシコゲノミクスは、たとえ話としては光学顕微鏡に対しての電子顕微鏡のような立場にある。すなわち、より高精度な情報を得る手段であるため、いままでの実験設計と精度では不十分な場合がある。そして真の実用化には“新しい教科書”を書く時間が必要であることも事実である。一方で、電子顕微鏡と異なるのは網羅的かつ膨大なデジタル情報が得られる点であり、インフォマティクスの助けが必須な理由である²⁾。分子毒性学の実用化のために、トランスクリプトームデータを相互に直接比較できるPercellome法の特徴を生かしたコンソーシアムの構築をめざし、共同研究を含めたさまざまな生体反応研究を進めているところである。臨床の最先端から分子生物学の最先端まで連携をさらに広げる際のひとつのきっかけとなれば幸甚である。

(厚労科研費 H13-生活-012, H13-生活-013, H14-トキシコ-001 および H15-化学-002)

- 1) Kanno, J. et al.: "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays. *BMC Genomics*, 7: 64, 2006.
- 2) Matsumoto, S. et al.: Mass Distributed Clustering: A New Algorithm For Repeated Meas-

が有意に増加していることを報告している³⁾。

今回、著者らは、日本人集団における *HLA*, *KIR* 遺伝子群と MPA の疾患感受性の検討を目的とした多施設共同研究を行った⁴⁾。日本人 MPA 43 例、健常対照者 239 例の末梢血中のゲノム DNA を用いて *HLA-B* と *HLA-C* を決定し、*KIR* の 14 遺伝子座を PCR-SSP 法によりタイピングして、それぞれ単独で疾患との関連を検討するとともに、*KIR* と *HLA* の遺伝子型の組合せと MPA の疾患感受性との関連を検討した。

HLA 単独では MPA と健常群に統計学的に有意な差はみられなかった一方、*KIR* に関しては活性化型受容体遺伝子である *KIR2DS3* の陽性率が健常群の 16.7% に比べ MPA では 4.7% と有

膠原病学

顕微鏡的多発血管炎と *KIR* 多型

Genetic interaction between killer cell immunoglobulin-like receptor (*KIR*) and *HLA* genes with microscopic polyangiitis (MPA)

Killer cell immunoglobulin-like receptor (*KIR*) は NK 細胞や一部の T 細胞に発現する活性化型・抑制型分子群である。活性化型 *KIR* のリガンドは不明であるが、抑制型 *KIR* のいくつかは *HLA*-class I 分子をリガンドとして認識することが知られている。*KIR* 遺伝子ファミリーはヒト染色体 19q13.4 上の leukocyte receptor complex (LRC) 内に位置し、各遺伝子座における塩基多型のみならず、ハプロタイプ上の遺伝子座の数の多型 (copy number polymorphism) が存在する¹⁾。さらに、リガンドである *HLA* も高度に多型的であるが、それぞれが別の染色体上に存在し、独立に遺伝することから、個体における *KIR* と *HLA* の組合せは多様であり、これが免疫応答の個体差に関連する可能性が推測されている。

KIR 多型の関連研究

これまでにいくつかの自己免疫疾患やウイルス感染症の関連研究において、*KIR* 遺伝子多型との関連を示す報告がある²⁾。アメリカ人集団において血管炎を合併した関節リウマチや、乾癆性関節炎などの自己免疫疾患では *KIR2DS1*, *KIR2DS2* など活性化型 *KIR* 遺伝子が患者群において高い頻度で観察されている。また、AIDS の進行が遅いタイプの *KIR* 遺伝子型や、C 型肝炎ウイルスのクリアランスにかかわる *KIR*-*HLA* 遺伝子型の組合せなど、ウイルス感染症の臨床経過との関連も報告されている。

KIR 多型と顕微鏡的多発血管炎

顕微鏡的多発血管炎 (MPA) は急速進行性腎炎や肺出血などの臓器障害を伴う稀少疾患であり、抗好中球細胞質抗体 (ANCA) 産生によって特徴づけられる。患者は高齢者に多く、病因はよく知られていないが、ウイルスや細菌感染との関連を示唆する報告もある。当研究室ではこれまでに日本人 MPA において、*HLA-DRB1**0901-*DQB1**0303 ハプロタイプ

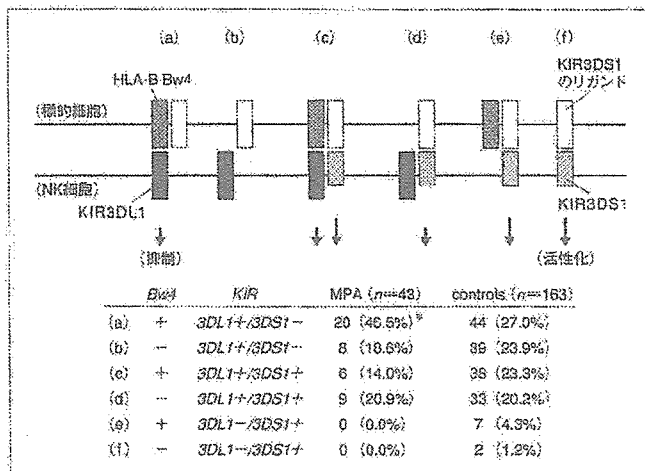


図 1 MPA および健常群における *HLA-B Bw4* と *KIR3DL1/3DS1* の組合せ

KIR3DL1 は *HLA-B* アリルのうち約半数に存在する *Bw4* エピトープを有するアリル産物に反応する。*KIR3DS1* と *HLA-B* との反応は証明されていないため、ここでは未知のリガンドが存在すると仮定して示しているが、*KIR3DS1* は *KIR3DL1* 遺伝子座における活性化型アリルであり、細胞外下メインの相同性から何らかの条件下で *Bw4* と反応する可能性が指摘されている。*HLA-B Bw4* と *KIR3DL1/3DS1* それぞれの遺伝子型の組合せは 6 通りに分類されるが、これらのうちで機能的にもっとも抑制的であると予想される *HLA-B Bw4* 陽性・*KIR3DL1* 陽性・*KIR3DS1* 陰性という群 (a) が MPA において 46.5% と対照群の 27.0% と比較してオッズ比 2.35 で有意に増加していた。
 * $p=0.014$ (χ^2 test), OR=2.35, 95% CI: 1.18-4.70

Special Review

Percellome Projectによる毒性トランスクリプトミクスの新しい試み

Percellome Project as a New Approach to Toxicology Transcriptomics

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

Jun Kanno, Satoshi Kitejima, Ken-ichi Aizaki, Katsuhide Igarashi, Noriyuki Nakatsu, Atsuya Takagi, Yukio Ogawa, Yukio Kodama

身の回りの物質の毒性(有害性)を予測し、その被害を未然に防ぐのが毒性学の役割である。この精度向上を目指したトキシコゲノミクス研究を実施する際に、マイクロアレイなどから細胞1個当たりのmRNAコピー数を得るPercellome法を開発した。90化合物のマウス肝初期応答データを採取し終え、新たな対象(反復投与、胎児毒性、吸入毒性、多臓器連携)を加えたPercellome Projectを展開している。

key words

トキシコゲノミクス、分子毒性学、遺伝子発現カスケード、標準化、Percellome法、3次元多層(Millefeuille)データ

菅野 純 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 E-mail: kanno@rihs.go.jp

1985年東京医科大学大学院医学研究科博士課程修了。人体高理学、実験病理学専攻。国立医薬品食品衛生研究所毒性部室長を経て、2002年より副部長。内分泌かく乱関連などの分子毒性学研究、トキシコゲノミクスプロジェクトなどを厚生労働所業務との有機的連携のもとに推進。

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 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

はじめに

医薬品、食品、化粧品、生活関連用品など、身の回りの物質が我々の身体に取り込まれた際に生じる可能性のある毒性(有害性)を予測し、それらの使用に際しての被害を未然に防ぐのが毒性学の役割である^{注1}(図1)。具体的には、人々の安全を確保するために使用法(用途)や使用量(残留量)を制限したり、場合によっては禁止したりするための科学的根拠を提供するが、その際、人の身代わりとして実験動物を用いる場合が多い。このような毒性学の精度向上の一環として、従来からの毒性研究(毒性症候学、毒性病理学、など)に加えてのトキシコゲノミクス(Toxicogenomics)研究が進められている。

トキシコゲノミクスでは、物質が生体に及ぼす影響をトランスクリプトームとして観測・解析する。その際、①分子毒性学を構築し種差や個体差の問題、複合暴露の問題などを解決するためには、遺伝子発現カスケードの全容解明を目指す必要がある。②形態学的に変化が現れた段階のトランスクリプトームは、遺伝子発現カスケードの最終段階に過ぎない。③形態変化の現れないごく初期段階を含む遺伝子発現カスケードを描出するためにはまとまった量のデータの蓄積が必須である。この観点から、筆者らは、マイクロアレイや定量PCRから細胞1個当たりのmRNAコピー数を得るPercellome手法と、そのデータ解析のための3次元多

層(Millefeuille)システムを開発・実用化した。遺伝子発現量が共通の尺度、すなわち“コピー数/細胞”で表現されることから、検体間、実験間、マイクロアレイのバージョン間、異なったプラットフォーム間、などのデータ比較が直接的に行えるようになり、数年かけて蓄積したデータの有機的活用が可能となった。現在、90種類の化学物質によるマウス肝の初期応答データを採取し終えたところである。新たな対象(反復投与、胎児毒性、吸入毒性、多臓器連携)を加えたPercellome Projectの概要を紹介する。

I. Percellome法:細胞1個当たりのmRNA絶対量を得る方法

原理は単純である。サンプルの細胞数を計測し、外部標準mRNA(スパイクRNA)を細胞1個当たり決まった分子数だけそのサンプルに添加し、そしてRNA抽出、測定に移る。サンプルのRNAの測定値を、スパイクRNAの値を基準に、細胞1個当たりのコピー数に換算する。実際には細胞数を直接計測するのが困難なことが多いため、その代替指標として細胞核内のゲノムDNA量を用いる^{1,2)}。定量性・直線性の検証にはLBM標準サンプル(肝[L]と脳[B]を100:0, 75:25, 50:50, 25:75および0:100に混合した5サンプルから成るセット)を用いる。なお、スパイクRNAは、5種類の枯草菌遺伝子のmRNAを濃度公比3で混合したカクテル(dose-graded spike cocktail; GSC)として用意した。高精度を要求されるDNA定量法は手作業プロトコールおよび自動ロボット(PerkinElmer JANUS)のプロトコールを準備

注1 環境への配慮も含まれる。

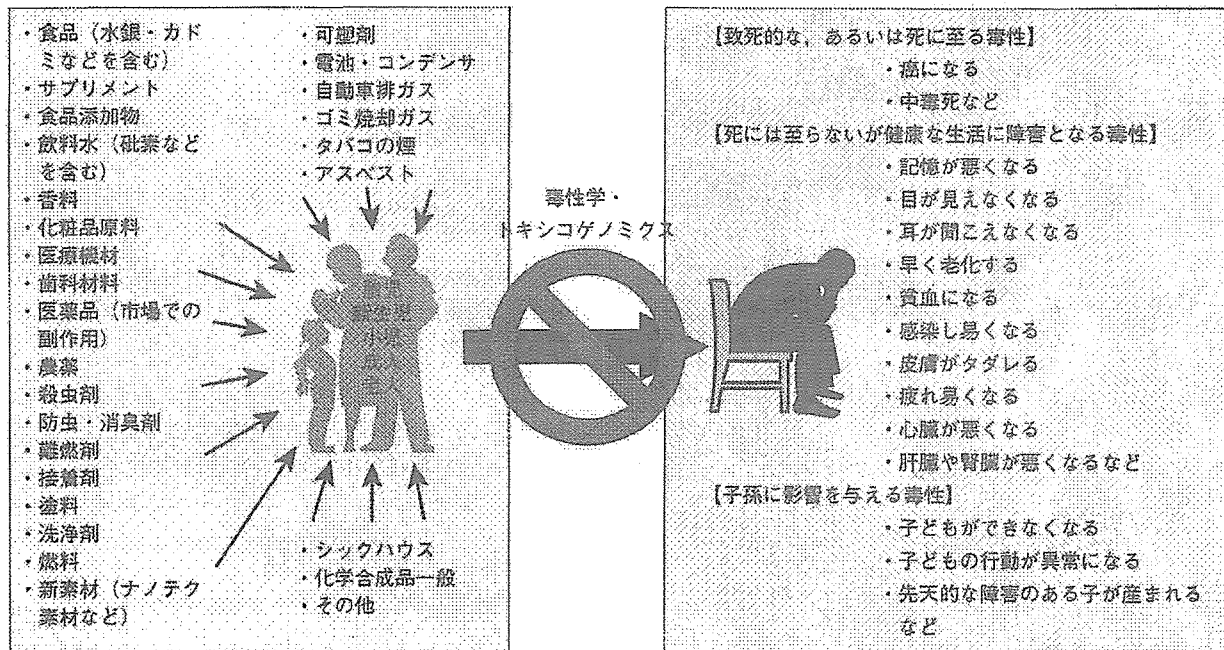


図1. 毒性学の対象

毒性学は、身の回りの物質が引き起こす障害を予測し、その発生を未然に防ぐことを目的としている。トキシコゲノミクス（毒性ゲノミクス）は、最先端の網羅的遺伝子発現解析技術を用いて、従来の毒性学の予測の精度を著しく向上、迅速化させることで、国民の健康安全の確保にさらに貢献することを目指している。

中である。カタテルとも共同研究ベースで供給可能である（連絡先：kanno@nihs.go.jp）。また、ERCC（The External RNA Control Consortium）と連絡をとるとともに、国際的標準化への関与を深めるため平成18年度厚労科研費「医薬品などの有効性・安全性評価に資する遺伝子発現解析の国際的標準化に関わる研究（H18-特別-指定-023）」を立ち上げた。現在、この他にシックハウス症候群を考慮した低用量域での吸入毒性トキシコゲノミクス、1匹のマウスから多臓器を採取しそれらの連携状況をトランスクリプトームから解析する多臓器トキシコゲノミクスを開始し、特徴的な遺伝子について組織内の発現分布を*in situ*ハイブリダイゼーションで確認する作業を並行している。また、下記の3次元データをweb公開するサーバを整備し、一部の化合物から3次元多層（Millefeuille）データを順次閲覧可能とした（<http://toxicomics.nihs.go.jp/db/>）。

II. 3次元多層（Millefeuille）データシステム：生物系研究者に優しいデータ可視化と解析

医薬品を含む毒性既知の90化合物について単回経口投与後のトランスクリプトームデータを取得して、初期応答遺伝子カスケードを解析するための基盤データベースを構築した。現在、第二段階として反復暴露データ集積を開始し

た。データは、用量軸、時間軸、および遺伝子発現軸から成る3次元表示により、遺伝子発現の用量および時間に依存した変化を1枚の曲面として表すことで可視的に変化を判別しやすいように配慮した（図2）。これにより、コンピュータが選び出した遺伝子クラスターの中身を確認する際、特に、mRNAの合成分解のスピードなどの知見から生物学的にありえないパターン（用量軸の方向にも時間軸の方向にもジグザグな変化など）を排除する際に威力を発揮している。

1つの実験から排出されるGeneChip約50枚のデータを一括処理する能力を持ったPercellome自動換算・データ品質管理（QC）に関わるソフトウェアに加えて、3次元多層（Millefeuille）データに最適化した、発現パターン類似性による候補遺伝子検索、およびそれを発展させた教師無しクラスタリング⁹⁾を中心とした解析システム（MF System, MFシリーズ、開発：相崎 健一）を独自に実用化し、開発継続中である（図3）。これらにより、データQCはその日のうちに、基本的な発現情報検索から全遺伝子の教師無しクラスタリングまでを3日間で完遂できるものとなっている。

この基本解析を用いて、発現パターンによって分類された候補遺伝子リストが多数生成される。一部の幸運な例ではただちに新規と思われる毒性関連反応を見いだすことができた。またそうでない場合のための1つの補強手段とし

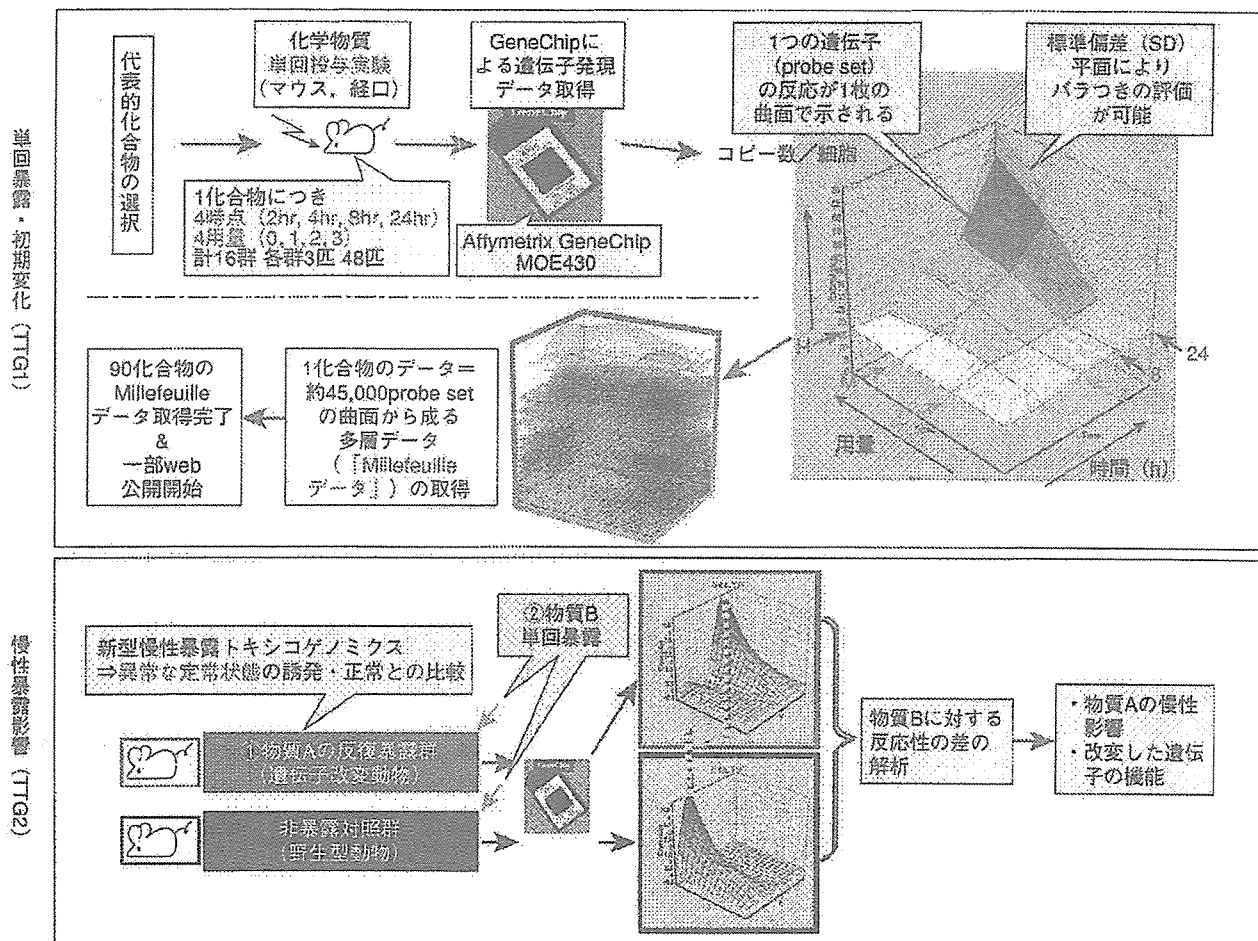


図2. Percellome 法と3次元表示による多層 (Millefeuille) データシステムを用いたプロジェクトの根幹部分の概要
単回投与による遺伝子発現初期変化データを90化合物について取得 (上段)。現在、反復投与の影響を検討中 (下段)。H: 高用量 (high), M: 中用量 (medium), L: 低用量 (low), C: コントロール (control)。

て、Gene Ontologyなどの既存知識を利用して候補遺伝子リストの理解を支援するソフトウェア (MF GoPlot) を用意した。このツールは一種の化合物クラスタリングとしても利用することができる。

さらに候補遺伝子リストを基に複数化合物間比較を行い、複数条件下においても同期して発現する遺伝子群を自動抽出するシステムも開発済みである。本システムで得られた同期遺伝子群はシグナルカスケードの構成単位である可能性があり、データベース化しつつ、その解析を進めている (5TB規模のデータベース部分および、大量計算アルゴリズム実装は (株) NTT コムウェアおよび (株) 日本NCR/Teradataとの共同開発による)。

Ⅲ. Percellome 手法のリアルタイムPCR を含む他のプラットフォームへの適用

Percellome 手法は、GSC の受け入れ条件を整えることに

より、様々なプラットフォームに適用可能である。その1つとして最も定量性が高いとされるリアルタイムPCR (ABI PRISM 7900 HT・96ウェルプレート) への適用例を示す。現行のRT-PCR絶対定量法では、遺伝子ごとに検量線が必要であり、多数のサンプルについて多数の遺伝子を検討するには不向きである。Percellome RT-PCRでは、マイクロアレイと同様の原理を用いる。すなわち、サンプル破碎液に、その細胞数に比例する量のスパイクカクテル (GSC) を添加し、それらのCt値をPCRプレートごとの検量線とすることにより、測定したい遺伝子のCt値を細胞1個当たりのmRNAコピー数に換算する。これにより、GAPDHやActinなどのハウスキーピング遺伝子が変動してしまう際の問題、例えば、少数の遺伝子を検討する際にGlobal normalization法を適用し難い問題などが解決される。共通サンプルを測定しデータを比較することにより、Affymetrix GeneChip のPercellome 結果と9割程度の整合性が確認され、

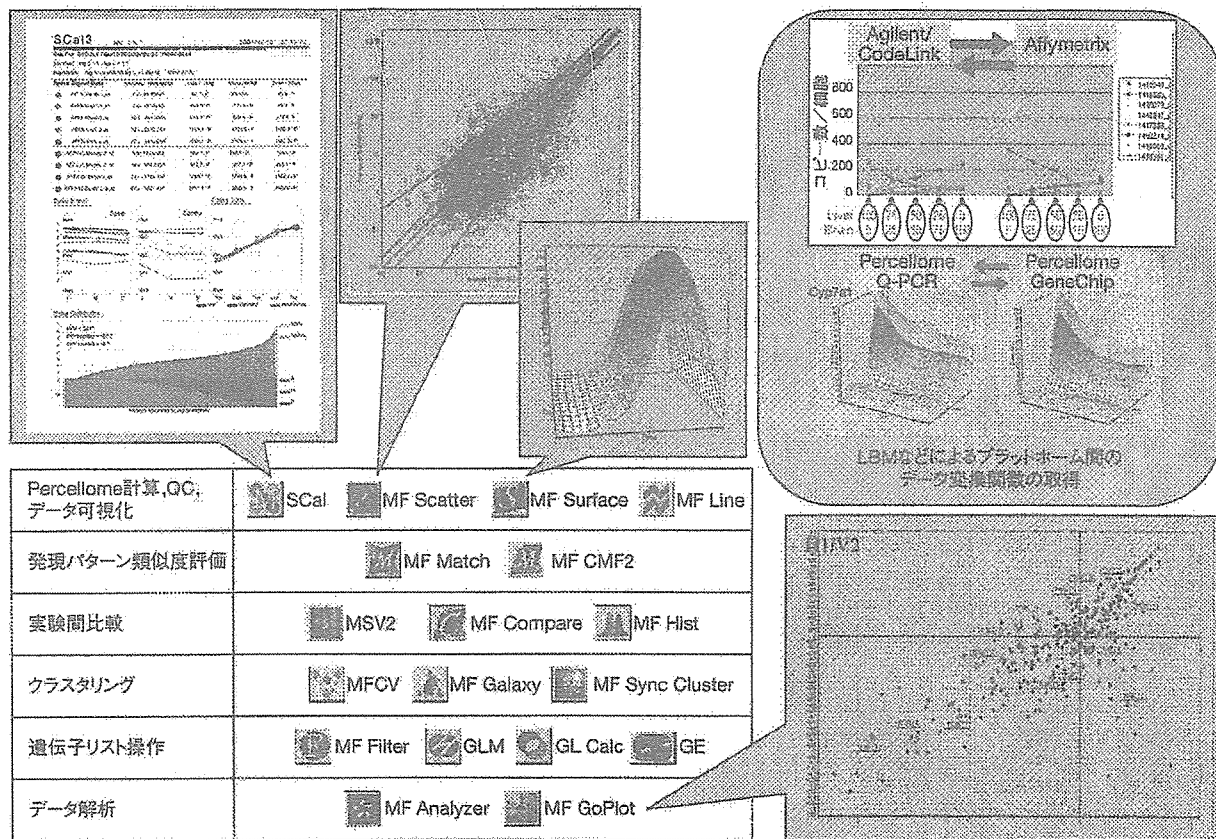


図3. 3次元多層 (Millefeuille) データの解析などに用いる独自開発プログラム群

品質管理とともに Percellome 計算を自動的に実施する S-Cal, Plot ソフトウェア, 3次元曲面の描画ソフト (MF Surface), など, 右上はプラットフォーム間のデータ変換情報の得方を示す。LBM を用いる方法 (上段) と, 実際の実験サンプルを用いる方法 (下段) がある。いずれも, 一度, 両方のプラットフォームでそれらのサンプルを測定する必要がある。

GeneChip と Percellome RT-PCR との間でのコピー数の換算式がいくつかの遺伝子について得られている。この他に, Agilent 社製の単色マイクロアレイと CodeLink アレイに GSC を測定可能なカスタムアレイを用意し終え, LBM サンプルのデータなどをもとに, これらとの間の換算式も得つつある (図3右上)。

Percellome 法は, Affymetrix の新しいエクソンアレイの定量性・直線性の検討にも適応可能である。Affymetrix 社の Human Exon 1.0 ST Array と従来型の発現アレイ Human Genome U133 plus 2 について, 性質の異なるヒト癌細胞株2株から調製した LBM 様標準サンプル (100:0, 75:25, 50:50, 25:75 および 0:100 混合5サンプル) による比較を行い, 両アレイ間の相関性の高い probe set を多数検出することができた。また, 既知のエクソンに対して設計された probe set では発現が見られ, インtron に対して設計された probe set では発現が見られない, あるいは, 既知の splicing variant に対応した probe set の発現が検出された,

などの基本性能が確認された。しかし, Percellome 法を適用して未知の splicing variant の検出力を向上させるためには, 現状では各エクソン間の定量性に問題があることが示唆された。定量値を算出する補正アルゴリズムの開発など, 何らかの対策が必要であることが考えられ, 現在, Affymetrix 社に確認を行っている。

IV. 核内受容体原性毒性の Percellome トキシコゲノミクス解析

受容体原性毒性とは, 化学物質が受容体 (リガンド依存的転写因子を含む) に選択的に結合してシグナルをかく乱し, その結果生じる有害性を指す。代表例としてはダイオキシンが挙げられる。AhR (Arylhydrocarbon receptor) ノックアウトマウスでは, ダイオキシンを大量に投与しても毒性がほとんど観察されない, すなわち, 野生型マウスがダイオキシンで死ぬメカニズムには, AhR が必須であり, AhR からの異常なシグナルがマウスを死に至らせていることに

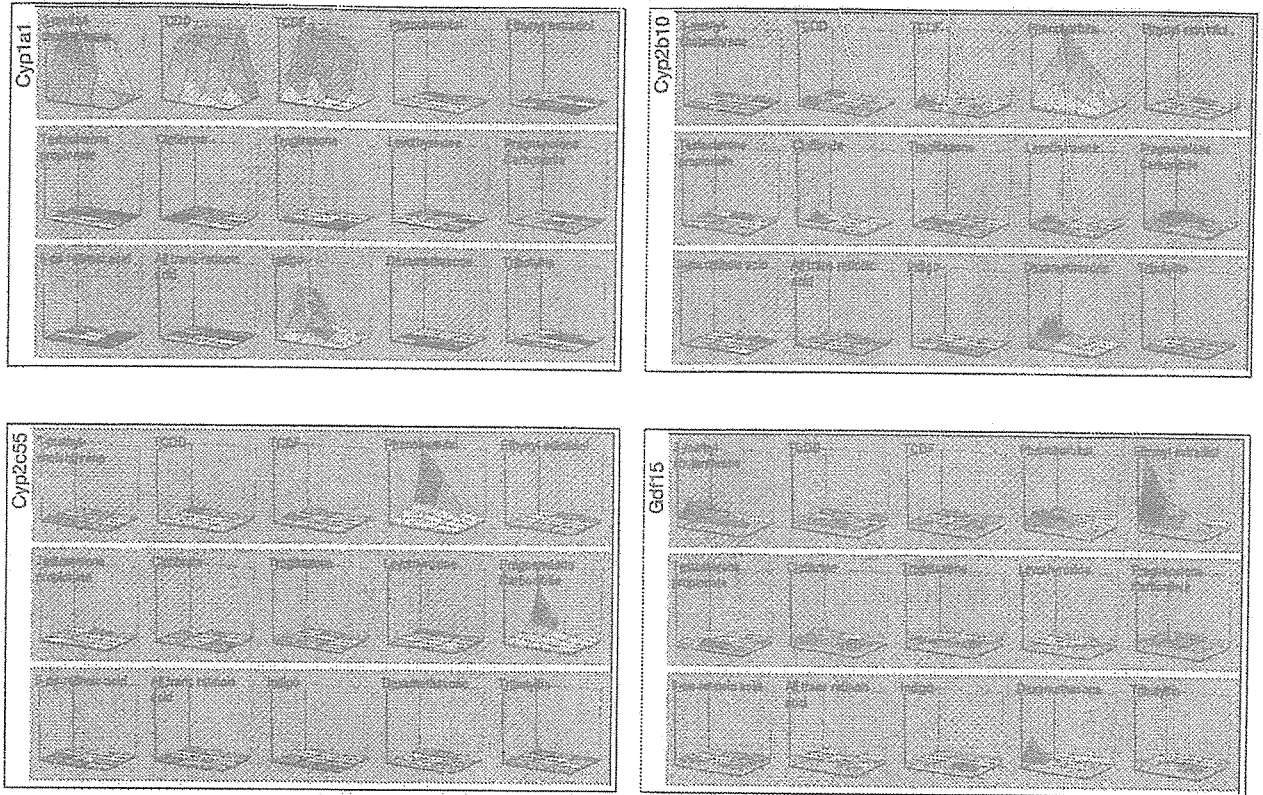


図4. 化合物間の発現比較

15種類の核内受容体リガンド化合物 (各3次元グラフ内に表示) によるCyp11a1 (左上), Cyp2c55 (左下), Cyp2b10 (右上) および, Gdf15 (右下) の遺伝子発現を3次元表示したもの. 各軸は, 図2のとおり, 縦軸のスケールは遺伝子ごとに共通. リガンドに選択的な遺伝子の発現が確認される.

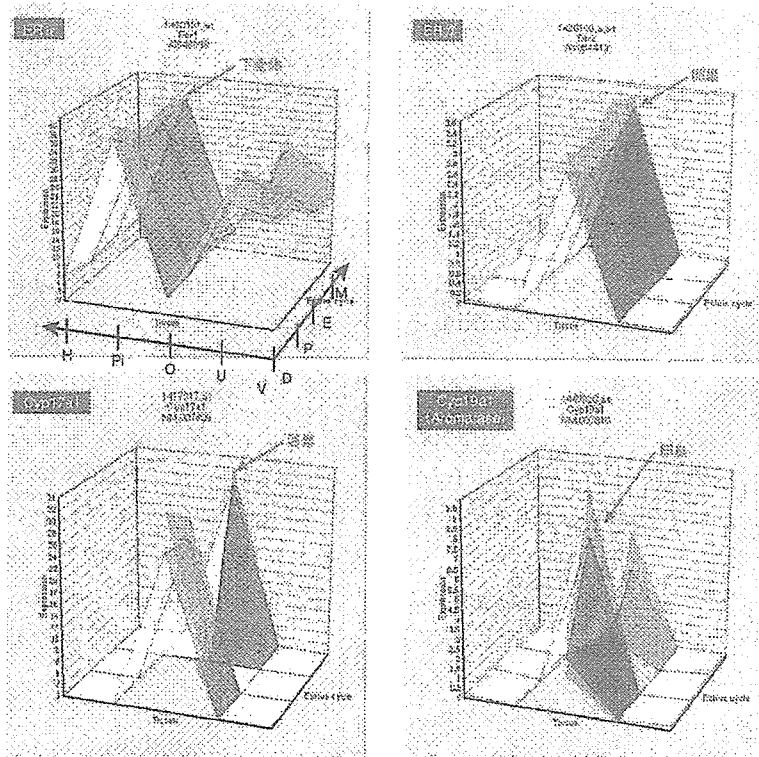


図5. 臓器間の発現比較

マウスの性周期 (Diestrus, Proestrus, Estrus, Metestrus) の4日間で1周期) ごとの視床下部 (H), 下垂体 (PI), 卵巣 (O), 子宮 (U) および陰 (V) における, ER α , ER β , Cyp17a1 (steroid-17 α -hydroxylase), およびCyp19a1 (Aromatase) の遺伝子発現変動を3次元表示したもの. 後二者の酵素は卵巣において周期性を持って発現している.

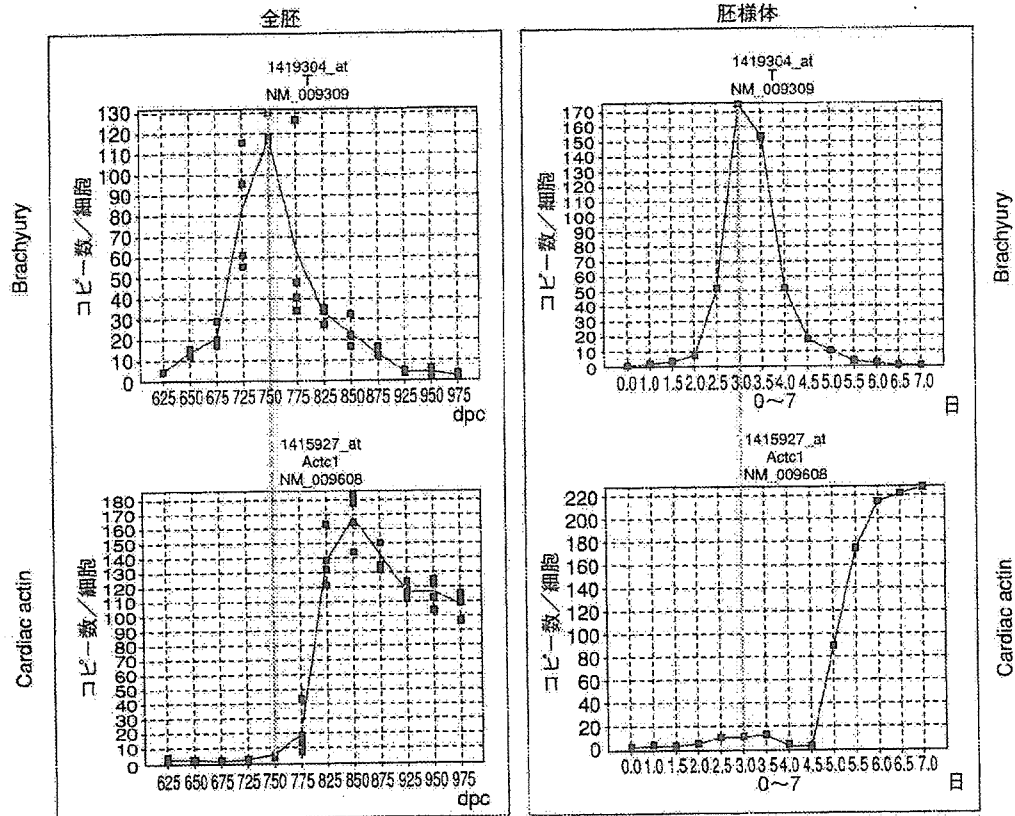


図6. マウス胎児 (全胚) と胚様体の発現比較

マウス全胚の胎生6.25日～9.75日までの遺伝子発現と、胚様体の1日～7日目までの遺伝子発現の網羅的データベースから、初期中胚葉分化マーカーであるBrachyury遺伝子と、Cardiac actin遺伝子の経時変化を示す。

なる。エストロゲン活性化学物質による有害影響(内分泌かく乱化学物質問題)も同様にER (estrogen receptor) を介する受容体原性毒性と考えられ、胎生期にERを発現する組織が、低用量シグナルかく乱影響の重要標的であると考えられている。

ここでは、受容体原性毒性研究の基盤として、Parcellome手法を適用して、①核内受容体作動性物質によるマウス雄肝臓の遺伝子発現変動、②性周期に伴うマウス雌生殖器遺伝子発現変動、③生後の発達過程におけるマウス雌生殖器遺伝子発現変動、の3種類のデータベースを構築した。例えば、①では10種類の核内受容体に作用する典型物質について、単回経口投与後、2, 4, 8, 24時間目の変動を解析し、Ethinyl-estradiolがGDF15, TCDDがCyp1a1, 9-cis Retinoic AcidがCyp26a1, DexamethasoneがCyp2b10, ClofibrateがCyp4a14, PCNがCyp2c55など、各々の受容体に特徴的な遺伝子発現を誘導するところがとらえられている(図4)。②の性周期データベースは視床下部、下垂体、卵巣、子宮、膈を対象としており、性周期との関連が網羅的にとらえら

れている(図5)。これらのデータベースは、今後、各種の候補物質が引き起こす変化を詳細に解析する際の基準として利用される。

V. 発生トキシコゲノミクスへの応用

発生毒性学は、個体発生過程におけるダイナミックな遺伝子発現調節の分子機構を把握することにより、さらに正確なものに補強されると考える。現在、C57BL/6マウス胚の器官形成初期初期にあたる胎生6.5～9.5日(プラグ確認日: 0.5日)の、①全胚の遺伝子発現変動解析、②遺伝子欠失マウス全胚との比較、および③標的が明らかな既知発生毒性物質投与による本データベースの具体的な適用、を実施している。①についてはすでに0.25日間隔(Time point計12点)の遺伝子発現データベースを得て、②遺伝子欠失胚のデータといくつかの注目すべき遺伝子についてはwhole mount ISHを用いた発現の検証を加えた。これと並行して、ES細胞からhanging drop法で得た胚様体の0.5日間隔の遺伝子発現データとの比較を実施している。個体発生に関与

する遺伝子群の多くは経時的に激しく変化しており、既知発生毒性物質投与実験については標的遺伝子シグナルカスケードを解析中である(図6)。

おわりに

ノーザンプロットでは実験サンプルにだけバンドが見られ、対照サンプルには遺伝子発現がないという結果を得ても、細胞1個当たりで定量してみると、対照が10コピーに対して実験サンプルが20コピーである場合がある。“無”が“有”になったのではなく、“10”が“20”になったのである。

さて、筆者らの属する毒性学でも、医学の分野でも、疾患概念や毒性概念が整理され、患者や実験動物を診断する際には、まず、そのどれに当てはまるかを検討する。すなわち、どの“典型”に近い症例であるかを検討することから始まることが多い。

しかし、最近の医学・生物学には多因子疾患・多因子形質発現制御の概念が導入され、今から何年かの後には、“21世紀初頭までは、患者の遺伝子多型を調べずして治療を行っていた時代”として、“血液型を調べずに輸血していた時代”と並び称されるようになる可能性がある。このような多因子概念が定着すると、その多くは、“有(100%)”“無(0%)”の組み合わせではなく、“70%”“50%”“90%”といった半端な数の組み合わせであることが考えられる。すなわち、今までの離散値的な“典型”例を基準とするアプローチから、

連続値的な病態“スペクトラム”を直接扱うアプローチに変革していく可能性が考えられる。その際の網羅的データの解析とその蓄積の必要性を考えると、遺伝子発現データの定量化・標準化という問題は、今まで以上に重みを増すと考えられる。生命現象の網羅的解析にはトランスクリプトームだけでは不十分であることは自明であるが、この定量性を確保することは、これから実現されるであろう網羅的プロテオミクスなどの基盤としても重要ではないかと考える。

マイクロアレイなどから得られるトランスクリプトーム情報が、今後の医薬品審査や化学物質の安全性評価の際に必須なものとなる時代がすぐそこまで来ていることを念頭に、筆者らはPercellome法をさらに展開し、Percellome Projectデータベースを可能な限り高精度に保ちつつ毒性学的な内容を充実させるべく最大限の活動を継続して行く所存であるが、この技術、あるいは研究内容が毒性学以外の研究分野にもお役に立つことができれば幸甚である。

謝辞 本システムの開発とプロジェクトの遂行に当たっては、当毒性部の全メンバー、特に松田菜恵、辻昌貴、森田絢一、今井あや子、安東朋子、安部麻紀、森山紀子、近藤優子、青柳千百合、相原妃佐子、渡辺忍の各氏の卓越した働きに深謝する。本研究は厚生労働科学研究費補助金H13-生活-012、H13-生活-013、H14-トキシコ-001、H15-化学-002、H18-化学-一般-001などによる。

文献

- 1) Kanno J, et al: BMC Genomics (2006) 7: 64
- 2) 菅野 純ら: 細胞工学 (2004) 23: 685-693
- 3) Matsumoto S, et al: Genome Informatics (2005) 16: 183-194