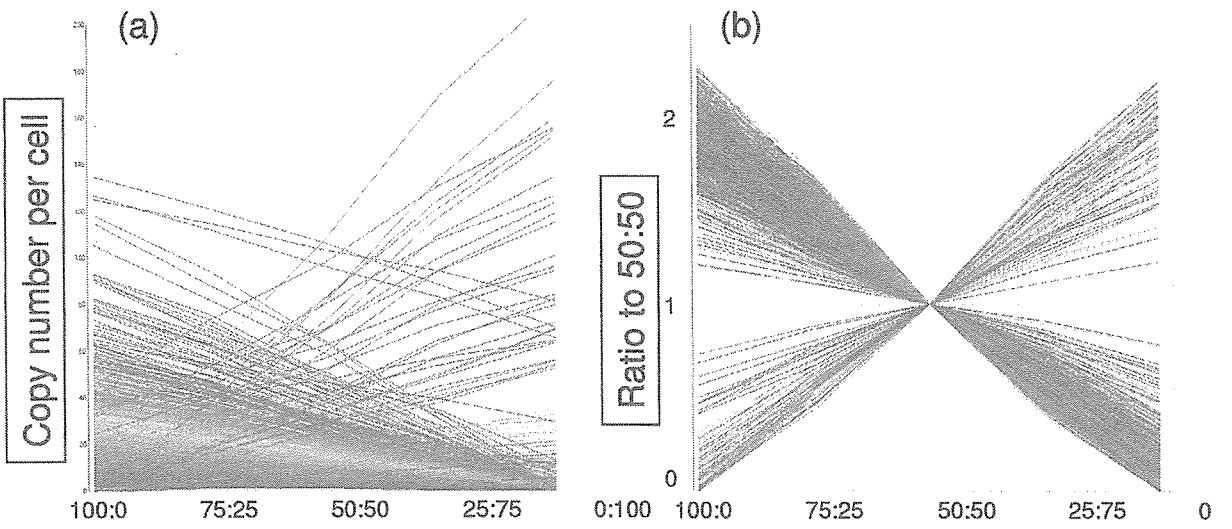


Figure 1
Dose-response linearity check by LBM. Dose-response linearity of the Affymetrix GeneChip by the LBM (liver-brain mix) sample set. Five samples, i.e. mixtures of mouse liver and brain at ratios of 100:0, 75:25, 50:50, 25:75 and 0:100, were spiked with GSC and measured by Affymetrix GeneChips Mouse430-2. Signals were normalized by the Percellome method as described in the text. Line graphs are in (a) copy numbers and (b) ratio to 50:50 sample for the top 1,000 probe sets with coefficient of correlation (R^2) closest to 1 among those having 1 copy or more per cell in the 50:50 sample (19,979 probe sets out of 45,101). The number of probe sets with $R^2 > 0.950$ was 8,655, and $R^2 > 0.900$ was 11,719.

regions using multiple platforms, is central to the establishment of a reliable reference database for toxicogenomics and pharmacogenomics. Transforming expression data into a "per cell" database is an effective way of normalizing expression data across samples and platforms. However, transcriptome data from the quantitative PCR (Q-PCR) and DNA microarray analyses currently deposited in the database are related to a fixed amount of RNA collected per sample. Variations in RNA yield across samples in an experimental series compromise accurate determination of the absolute level of each mRNA species per cell in any sample. Normalization against housekeeping genes for PCRs, and global normalization of ratiometric data for microarrays, is typically performed to account for this informational loss. Additional methods, such as the use of external mRNA spikes, reportedly improve the quality of data from microarray systems. For example, Holstege et al. [1] described a spike method against total RNA, based on their finding that the yields of total RNA from wild type and mutant cells were very similar. Hill et al. [2] reported a spike method against total RNA for normalizing hybridization data such that the sensitivities of individual arrays could be compared. Lee et al. [3] demonstrated that "housekeeping genes" cannot be used as a reference control, and van de Peppel et al. [4] described a normalization method of mRNA against total RNA using an external spike mixture. To achieve satisfactory performance they used multiple graded doses of external spikes, covering a wide range of expression, in order to align the ratiometric data by Lowess normalization [5]. Hekstra et al. [6] presented a method for calculating the final cRNA concentration in a hybridization solution. Sterrenburg et al. [7] and Dudley et al. [8] reported the use of common reference control samples for two-color microarray analyses of the human and yeast genomes, respectively. These are pools of antisense oligo sequences against all sense oligos present on the microarray. Instead of antisense oligos, Talaat et al. [9] used genomic DNA as a common reference control in studies of *E. coli*. Statistical approaches have been proposed for ratiometric data to improve inter-microarray variations, especially of non-linear relations [10]. However, because control samples may differ among studies, ratiometric data cannot easily be compared across multiple studies unless a common reference, such as a mixture of all antisense counterparts of spotted sense sequences is used [7-9]. Nevertheless, as long as the normalization is calibrated to total RNA, variations in total RNA profile cannot be effectively cancelled out. Although

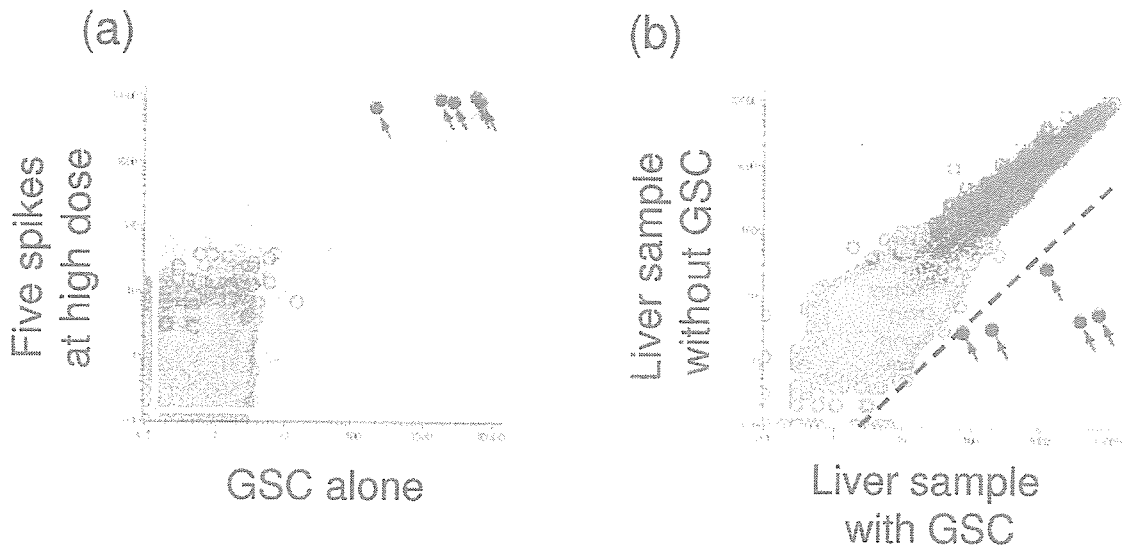


Figure 2
Cross-hybridization of GSC. Cross-hybridization of the GSC spike mRNAs to Affymetrix GeneChip. (a) A scatter plot of a blank sample with the GSC (horizontal axis) and a blank with the five spike RNAs at a high dosage (vertical axis) measured by MG-U74v2A GeneChips (raw values generated by Affymetrix MAS 5.0 software). The five spikes are indicated by black dots with arrows. Signals of the murine probe sets were below 20 on the horizontal axis, indicating negligible cross-hybridization of GSC spike mRNAs to the murine probe sets. (b) A scatter plot of a liver sample with GSC (horizontal axis) and without GSC (vertical axis) measured by MG-U74v2A GeneChips. The five spikes are again indicated by black dots with arrows. The dotted line is the 1/25 fold (4%) line. Cross-hybridization of mouse liver mRNAs to the GSC signals was considered negligible (less than 4%).

some of these reports share the idea that "absolute expression" and "transcripts per cell" should entail robust normalization, further practical development to enable universal application has been awaited.

Here, we report a method for normalizing expression data across samples and methods to the cell number of each sample, using the DNA content as indicator. This normalization method is independent of the gene expression profile of the sample, and may contribute to transcriptome studies as a common standard for data comparison and interchange.

Results

Dose-response linearity of the measurement system as a basis for the Percellome method

The fidelity of transcript detection is the key to this "per cell" based normalization method, which generates transcriptome data in "mRNA copy numbers per cell". The Q-PCR system was tested by serially diluting samples to confirm the linear relationship between Ct values and the log of sample mRNA concentration (data not shown). High density oligonucleotide microarrays from Affymetrix [11] were used in our experiments. We tested the linearity of the Affymetrix GeneChips using a set of five samples made of mixtures of liver and brain in ratios of 100:0, 75:25, 50:50, 25:75, and 0:100 (designated "LBM" for liver-brain mix). The results showed a linear relationship ($R^2 > 0.90$) between fluorescence intensity and input for a sufficient proportion of probe sets, i.e. about 37% of the probe sets in the older MG-U74v2 and 70% in the newest Mouse Genome 430 2.0 GeneChip were above the detection level (approximately one copy per cell) in the 50:50 sample (Figure 1) [see Additional files 1 and 2].

Dose-response linearity alone is not sufficient to generate true mRNA copy numbers. An important additional requirement is that the ratio of signal intensity to mRNA copy number should be equal among all GeneChip probe sets of mRNAs and PCR primers. The Q-PCR primer sets were designed to perform at similar amplification rates to minimize differences between amplicons. The melting

Table 1: The spike factors for various organs/tissues

Species	Organ/Tissue (adult, unless otherwise noted)	Spike Factor	total RNA/genomic DNA	SD
Mouse	Liver	0.2	211	46
Mouse	Lung	0.02	22	4
Mouse	Heart	0.05	-	-
Mouse	Thymus	0.01	8	2
Mouse	Colon Epitherium	0.05	105	30
Mouse	Kidney	0.1	-	-
Mouse	Brain	0.1	-	-
Mouse	Suprachiasmatic nucleus (SCN)	0.1	-	-
Mouse	Hypothalamus	0.1	63	4
Mouse	Pituitary	0.1	52	8
Mouse	Ovary	0.02	35	4
Mouse	Uterus	0.02	42	12
Mouse	Vagina	0.02	81	38
Mouse	Testis	0.15	56	7
Mouse	Epididymis	0.07	53	16
Mouse	Bone marrow	0.02	14	3
Mouse	Spleen	0.02	-	-
Mouse	Whole Embryo	0.15	97	36
Mouse	Fetal Telencephalon E10.5–16.5	0.1	48	9
Mouse	Neurosphere (E11.5–14.5)	0.03	42	10
Mouse	E9.5 embryo heart	0.15	58	15
Mouse	cell lines	0.2	-	-
Rat	Liver	0.2	-	-
Rat	Kidney	0.2	-	-
Rat	Uterus	0.04	56	5
Rat	Ovary	0.04	56	9
Human	Cancer Cell Lines	0.2	116	26
Xenopus	liver	0.03	-	-
Xenopus	embryo	0.15	-	-

temperature was set between 60° and 65°C with a product size of approximately 100 base pairs using an algorithm (nearest neighbor method, TAKARA BIO Inc., Japan), and the amplification co-efficiency (E) was set within the range 0.9 ± 0.1 (E = 2^{-(1/slope)}-1 on a plot of log2 (template) against Ct value). For the GeneChip system, the signal/copy performance of each probe set depended on the strategy of designing the probes to keep the hybridization constant/melting temperature within a narrow range, ensuring that the dose-response performances of the probe sets were similar (cf. http://www.affymetrix.com/technology/design/index.affx). Failing this, any differences should at least be kept constant within the same make/version of the GeneChip. Taking into consideration the biases that lead to imperfections in estimating absolute copy numbers in each gene/probe set, we developed normalization methods to set up a common scale for Q-PCR and Affymetrix GeneChip systems.

The grade-dosed spike cocktail (GSC) and the "spike factor" for the Percellome method

A set of external spike mRNAs was used to transfer the measurement of cell number in the sample (as reflected by its DNA content) to transcriptome analysis. For the spikes, we utilized five *Bacillus subtilis* mRNAs that were left open for users in the Affymetrix GeneChip series. The extent to which the *Bacillus* RNAs cross-hybridized with other probe sets was checked for the Affymetrix GeneChip system. The GSC was applied to Murine Genome U74Av2 Array (MG-U74v2) GeneChips with or without a liver sample. As shown in Figure 2, cross-hybridization between *Bacillus* RNAs and the murine gene probe sets was negligible [see Additional files 3 and 4]. Mouse Genome 430 2.0 Array (Mouse430-2), Mouse Expression Arrays 430A (MOE430A) and B (MOE430B), Rat Expression Array 230A (RAE230A), *Xenopus laevis* Genome Array and Human Genome U95Av2 (HG-U95Av2) and U133A (HG-U133A) Arrays sharing the same probe sets for these spike mRNAs showed no sign of cross-hybridization with the *Bacillus* probes (data not shown).

We prepared a cocktail containing in vitro transcribed Bacillus mRNAs in threefold concentration steps, i.e. 777.6 pM (for AFFX-ThrX-3_at), 259.4 pM (for AFFX-LysX-3_at), 86.4 pM (for AFFX-PheX-3_at), 28.8 pM (for AFFX-DapX-3_at) and 9.6 pM (for AFFX-TrpnX-3_at). By referring to the amount of DNA in a diploid cell and employing a "spike factor" determined by the ratio of

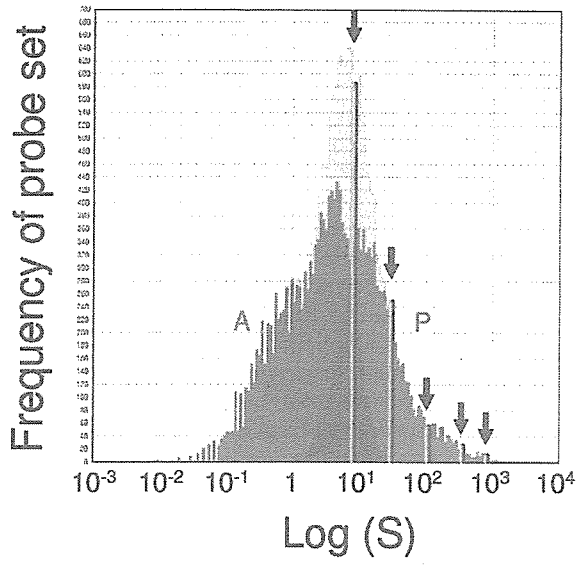


Figure 3
Positioning of GSC spike mRNAs in Affymetrix GeneChip dose-response range. A frequency histogram of the probe sets of Affymetrix GeneChip Mouse430-2 is shown. The histogram for all probe sets (gray) shows near-normal distribution. Blue columns are the "present" calls (P), red columns "absent" calls (A) and green "marginal" calls. The five yellow lines indicate the positions of the GSC spike mRNAs that are chosen to cover the "present" call range by a proper "spike factor".

total RNA to genomic DNA in a tissue type (Table 1), the spike mRNAs were calculated to correspond to 468.1, 156.0, 52.0, 17.3 and 5.8 copies per cell (diploid), respectively, for the mouse liver samples (spike factor = 0.2). The ratio of mRNAs in the cocktail is empirically chosen depending on the linear range of the measurement system and the available number of spikes. Here, we set the ratio to three to cover the "present" call probe sets of the Affymetrix GeneChip system (Figure 3).

We tested this grade-dosed spike cocktail (GSC) by Q-PCR and confirmed that the Ct values of the spike mRNAs were linearly related to the log concentrations (cf. Figure 4a), i.e. could be expressed as

$$Ct = \alpha \log C + \beta \quad \{1\}$$

The GSC was also tested by the GeneChip system and it was confirmed that the log of the spike mRNA signal intensities was linearly related to the log of their concentrations (cf. Figure 4b),

$$\log S = \gamma \log C + \delta \quad \{2\}$$

The linear relationship between the Ct values (Ct) and the log of RNA concentration (log C) was reasonable given the definition of Ct values (derived from the number of PCR cycles, i.e. doubling processes). The linear relationship between the log of GeneChip signal intensity (log S) and the log of RNA concentration (log C) was rationalized by the near-normal distribution of log S over all transcripts (cf. Figure 3).

Calculation of copy numbers of all genes/probe sets per cell

As described above, using a combination of DNA content and the spike factor of the sample, the GSC spike mRNAs become direct indicators of the copy numbers (C') per cell. When the samples were measured by Q-PCR or GeneChip analysis, the five GSC spike signals in each sample should obey function {1} for Q-PCR and function {2} for GeneChip with a good linearity. If the observed linearity was poor, a series of quality controls was performed and the measurement repeated. The coefficients of the functions were determined for each sample by the least squares method. Under the assumption that all genes/probe sets share the same signal/copy relationship, signal data for all genes/probe sets were fitted to the functions {1'} or {2'}, which are the individualized functions of {1} and {2} for each sample measurement (i).

$$Ct = \alpha i \log(C') + \beta i \quad \{1'\}$$

$$\text{Log } (S) = \gamma i \log(C') + \delta i \quad \{2'\}$$

(i = sample measurement no.)

The Q-PCR Ct values (Ct) and microarray signal values (S) of all mRNA species in the sample (i) are converted to copy numbers per cell (C') by the inverses of functions {1'} and {2'}, i.e. {3} and {4} below:

$$C' = B\text{^}((Ct-\beta i)/\alpha i) \quad \{3\}$$

for Q-PCR (Figure 4a).;

$$C' = B\text{^}((\log S-\gamma i)/\delta i) \quad \{4\}$$

for GeneChips (Figure 4b),

where B is the logarithmic base used in {1} and {2} (see Materials and Methods for details).

Real world performance of the Percellome method

The correspondence between Q-PCR and GeneChip was tested using a sample set from 2,3,7,8-tetrachlorodibenzodioxin (TCDD)-treated mice. Sixty male C57BL/6 mice

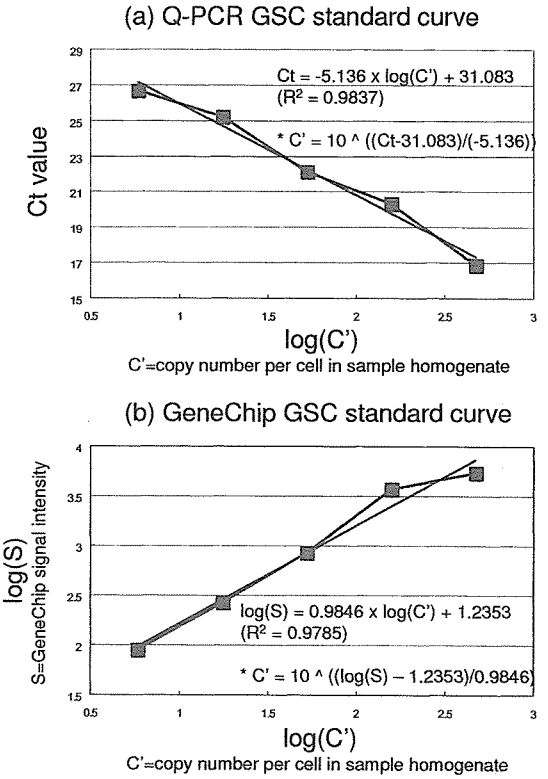


Figure 4
The dose-response linearity of the GSC spikes in Q-PCR and the Affymetrix GeneChip array system. Linear relationships are shown between (a) the Q-PCR Ct values and log of copy number (log (C')), and (b) the GeneChip log signal intensity (log(S)) and log of copy number (log (C')) of the GSC mRNAs. The regression functions were obtained by the least squares method. The inverse functions (*) were further used to generate the copy numbers of all other genes/probe sets for Percellome normalization.

were divided into 20 groups of 3 mice each. TCDD was administered once orally at doses of 0, 1, 3, 10 and 30 μg/kg, and the livers were sampled 2, 4, 8 and 24 h after administration. Nineteen primer pairs were prepared for Q-PCR and the Ct values of the liver transcriptome were measured. The same 60 liver samples were measured using the Affymetrix Mouse430-2 GeneChip [see Additional files 5 through 8 and 9 through 12]. Q-PCR and GeneChip data were normalized against cell number by functions {3} and {4}, respectively. The averages and standard deviations (sd) of each group (n = 3) were calculated and plotted as three layers of isoborograms on to 5 × 4 matrix three-dimensional graphs (Figure 5). Together with another sample set (data not shown), a total of thirty-six primer pairs were compared, and there was a correlation of up to 90% between the Q-PCR and GeneChip surfaces. It is notable that not only the average surfaces but also the +1sd and -1sd surfaces corresponded closely in shape and size. We infer that the differences resulted mainly from biological variations among the three animals in each experimental group rather than from measurement error (cf. Figure 7).

An important feature of Percellome normalization is its independence from the overall expression profile of the sample. When gene expression profiles differ among samples, Percellome normalization produces a robust transcriptome that is different from total-RNA dependent global normalization. As an example, Figure 6 shows the results of an experiment on the uterotrophic response of ovariectomized mice to estrogen treatment [12] [see Additional files 13 and 14]. The uteri of the vehicle control are atrophic because the ovaries, the source of intrinsic estrogens, are absent. The uteri of the treated groups are hypertrophic owing to estrogenic stimulus from the test compound administered. Global normalization (90 percentile) between the vehicle control group and the high-dose (1,000 mg/kg) group indicated that 4,600 of 12,000 probe sets showed 2-fold or greater increase, 470 were reduced by 0.5 or less, and 7,400 remained between these extremes. In contrast, analysis of Percellome-normalized data revealed that almost all the 12,000 probe sets showed a 2-fold or greater increase, including actin, GAPDH and other housekeeping genes. The hypertrophic tissues, consisting of cells with abundant cytoplasm, provide convincing evidence for the increases in various cellular components including housekeeping gene products.

Another important feature of Percellome normalization is the commonality of the expression scale across platforms. Batch conversion can be performed between results obtained from different platforms when the data are generated by the Percellome method. A practical strategy for such normalization is to prepare a set of samples from a target organ of interest with differences in gene expression, and measure them once by each platform. Data conversion functions with good linear dose-response relationships can be obtained individually for those genes/probe sets that are measured by both platforms (Figure 7).

Discussion

We have developed a novel method for normalizing mRNA expression values to sample cell numbers by adding external spike mRNAs to the sample in proportion to the genomic DNA concentration. For non-diploid or aneuploid samples, an average DNA content per cell should be determined beforehand for accurate adjustment. When there is significant DNA synthesis, a similar adjustment should be considered.

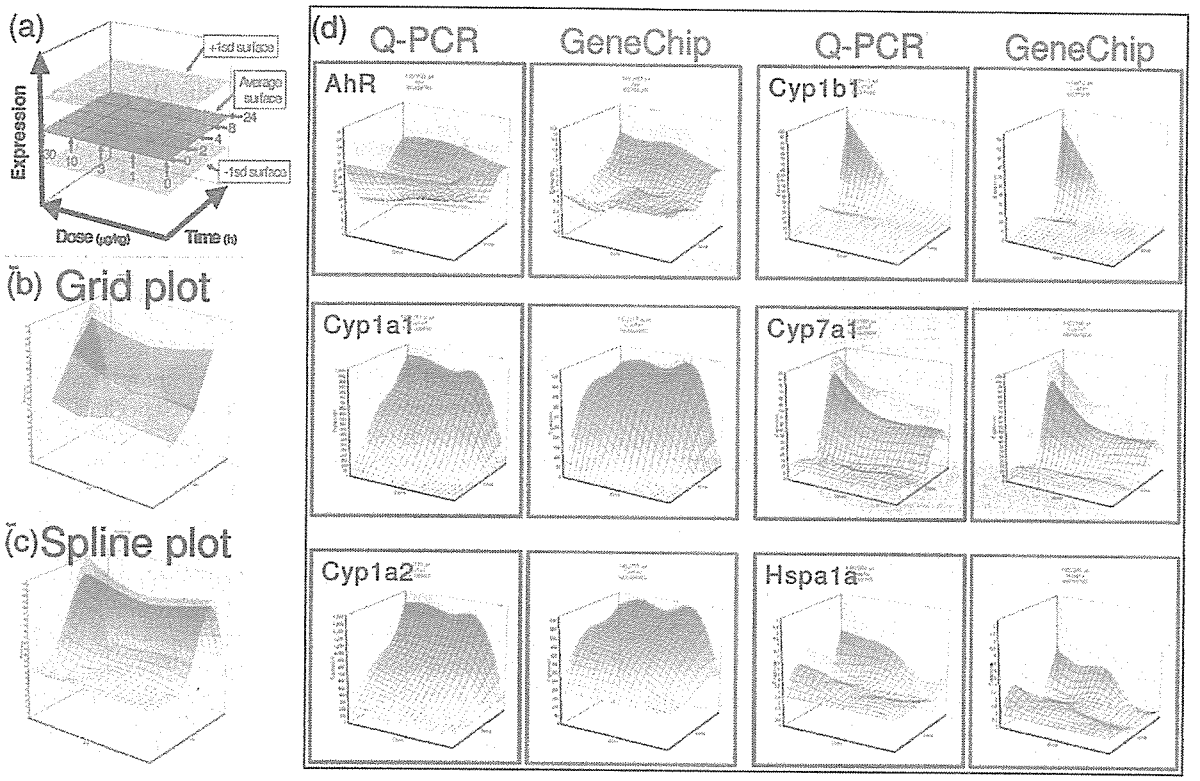


Figure 5

Correspondence between Q-PCR and GeneChip data. Sixty male C57BL/6 mice were divided into 20 groups of 3 mice each. 2,3,7,8-tetrachlorodibenzodioxin (TCDD) was administered once orally at doses of 0, 1, 3, 10 and 30μg/kg, and the liver was sampled 2, 4, 8 and 24 h after administration. The liver transcriptome was measured by the Affymetrix Mouse430-2 GeneChip. For Q-PCR, nineteen primary pairs were prepared and the Ct values of the same 60 liver samples were measured (19 genes and 5 spikes in duplicate, using a 96-well plate for 2 samples, total 30 plates). The Percellome data were plotted on to 3-dimensional graphs for average, +1sd, and – 1sd surfaces as shown in (a). The scale of expression (vertical axis) is the copy number per cell. The 0 h data (*) are copied from the 2 h/dose 0 point for better visualization of the changes after 2 h. The surfaces are demonstrated as a grid plot (b) where the grid points indicate one treatment group (n = 3), and a smoothened spline surface plot (c) for easier 3D recognition ((b), (c): Gys2 (glycogen synthase 2, 1424815_at) showing a typical circadian pattern. (d) the smoothened plots of 6 representative genes/ probe sets generated by Q-PCR (red) and GeneChip (blue). AhR (arylhydrocarbon receptor, 1450695_at) showed imperfect correspondence. Cyp1a1 (cytochrome P450, family 1, subfamily a, polypeptide 1, 1422217_a_at) and Cyp1a2 (1450715_at) showed good correlations between Q-PCR and GeneChip except for the saturation in GeneChips above c. 400 copies per cell. Cyp1b1 (1416612_at) and Cyp7a1 (1422100_at) showed good correspondence. Hspa1a (heat shock protein 1A, 1452888_at) showed fair correspondence despite low copy numbers, near the nominal detection limit of the Affymetrix GeneChip system.

The smallest sample to which we have successfully applied the direct DNA quantification method with sufficient reproducibility is the 6.75 dpc (days post coitus) mouse embryo which consists of approximately 5,000 cells. This sample size is also approximately the lower limit for double amplification protocol to obtain sufficient amount of RNA for Affymetrix GeneChip measurement (cf. http://www.affymetrix.com/Auth/support/downloads/manuals/expression_print_manual.zip.)

High-resolution technology such as laser-capture microdissection (LCM) has become popular and the average sample size analyzed is getting smaller. An alternative method for LCM samples is to count the cell number in the course of microdissection. Although we have not yet applied Percellome method to LCM samples, we have applied the alternative method to cell culture samples to gain Percellome data. Stereological and statistical calculations should become available to correct the number of partially sectioned cells in the LCM samples. Another issue for small samples is the yield of RNA. Approximately

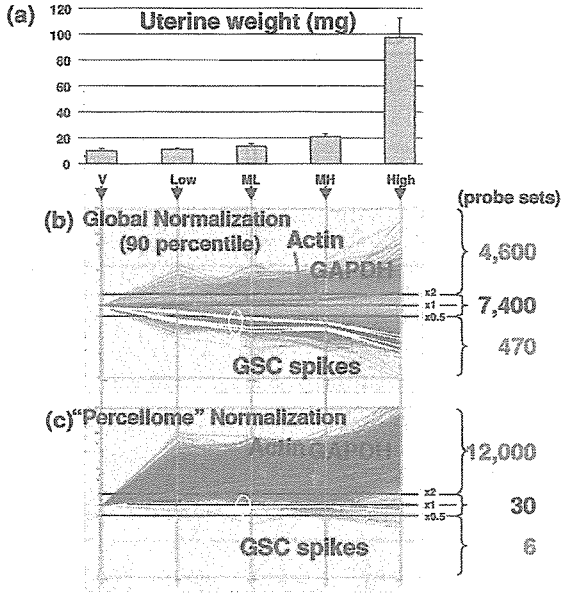


Figure 6
Uterotrophic response of ovariectomized female mice by an estrogenic test compound.(a) Shows the uterine weight, which increases in a dose-dependent manner; V, vehicle control; Low, low dose; ML, medium-low dose; MH, medium-high dose; High, high dose group. (b) Shows the line display of uterine gene expression (Affymetrix MG-U74v2 A GeneChips) normalized by global normalization (90 percentile), and (c) by the Percellome normalization. Averages of three samples per group were visualized (by K. A.). The five white lines are the GSC mRNAs. The green and blue lines are actin (AFFX-b-ActinMur/M12481_3_at) and GAPDH (glyceraldehyde-3-phosphate dehydrogenase, AFFX-GapdhMur/M32599_3_at), respectively. By global normalization, 7,400 probe sets remained unchanged and 4,600 probe sets increased more than two-fold in the H group compared to the V group, whereas almost all probe sets measured had increased. It is noted that housekeeping genes such as actin and GAPDH are significantly induced on a per cell basis.

30 ng of total RNA is retrieved from a single 6.75 dpc mouse embryo. This amount is sufficient for a double amplification protocol (DA) to prepare enough RNA for an Affymetrix GeneChip measurement. An inherent problem with the DA data is that the gene expression profile differs from that of the default single amplification protocol (SA). Consequently the DA percellome data differ from that of SA as if they were produced by a different platform. To bridge the difference, we applied the procedure that was used for data conversion between Q-PCR and GeneChip (cf. Figure 7). A set of spiked-in standard samples including the LBM sample set (of sufficient concentration) were measured by the SA protocol and diluted versions to the limit measured by the DA protocol. These data provided us with information about whether DA was successful as a whole (by comparing 5' signal to 3' signals of selected probe sets) and which probe sets were properly amplified by DA (by checking the linearity of the diluted LBM data). For those probe sets that proved to be linearly amplified, conversion functions between DA and SA were generated. These details, along with embryo expression data will be published elsewhere.

Figures 5 and 7 indicate a close correspondence between the data generated by Q-PCR and GeneChip analyses. Since each of the 60 samples was normalized individually against each GSC signal, the high similarity between the two platforms indicates the robustness and stability of this spike system (cf. Figure 7, Cyp7a1 data). Although more spikes could potentially increase the accuracy of normalization, our experience is that five spikes are practically sufficient for covering the detection range of GeneChip microarrays and Q-PCR, as long as they are used in combination with the "spike factor". The overall benefits of using a minimum number of external spikes include lower probability of cross-hybridization, a reduced number of wells and spots occupied by the spikes in the Q-PCR plates and small scale microarrays, and less effort in preparation, QC and supply.

The Percellome data can be truly absolute when all mRNA measurements including GSC spikes are strictly proportional to the original copy numbers in the sample homogenate. As noted earlier, this condition is not guaranteed by any platform despite linearity of response. Therefore, the Percellome-normalized values have some biases for each primer pair/probe set, depending on the steepness of the dose-response curves. An advantage of Percellome normalization is that, as long as such biases are consistently reproduced within a platform, the data can be compared directly among samples/studies on a common scale. Consequently, when a true value is obtained by any other measure, all the data obtained in the past can be simultaneously batch-converted to the true values.

This batch-conversion strategy can be extended to data conversion between different versions and different platforms, as long as the data are generated in copy numbers "per cell". We have shown an example between Affymetrix GeneChip and Q-PCR for limited numbers of probe sets (cf. Figure 7). Custom microarrays that accept our GSC for Percellome normalization are in preparation by Agilent Technologies (single color) and GE Healthcare (CodeLink Bioarray).

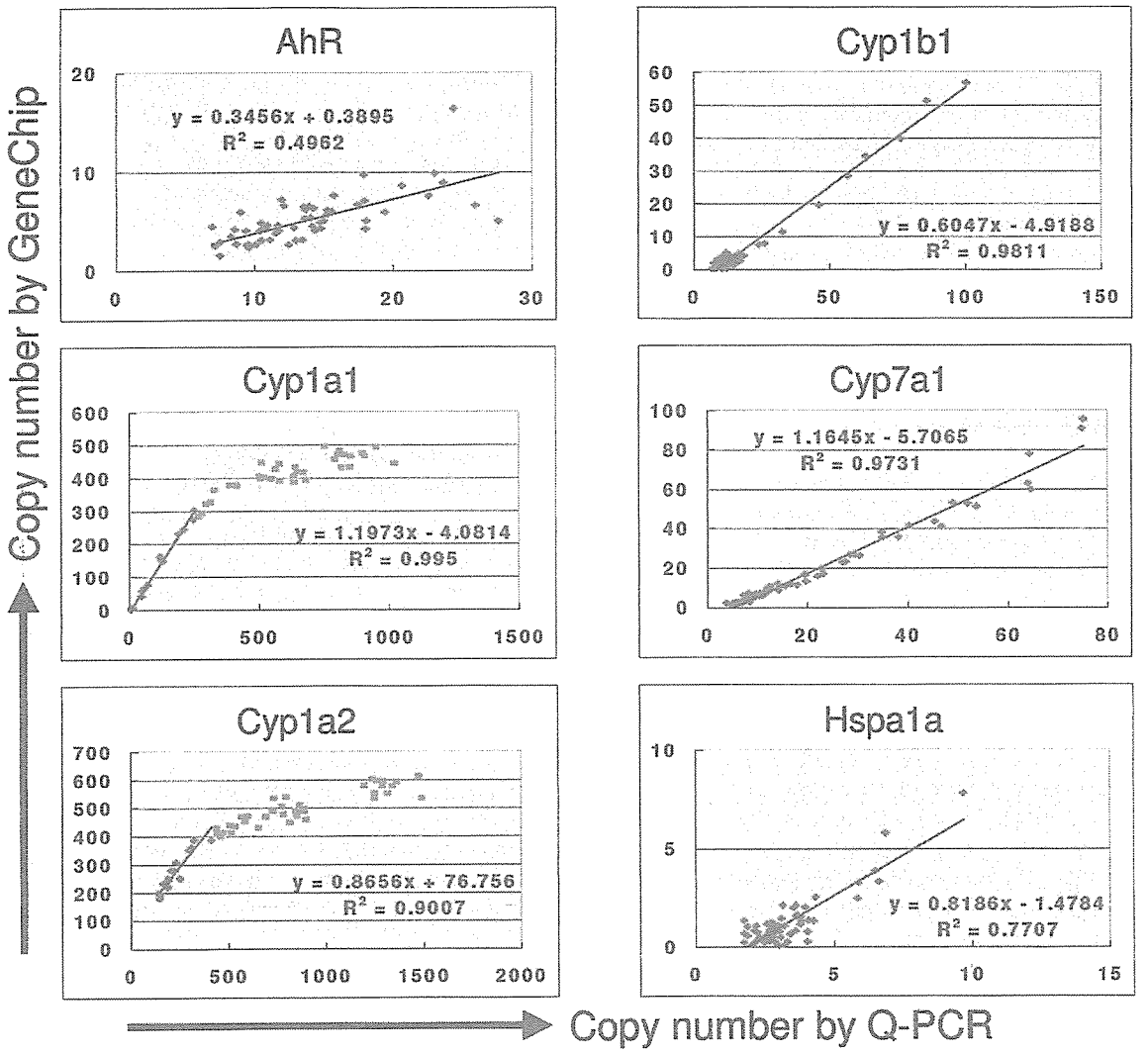


Figure 7
Conversion functions between Q-PCR and GeneChip. The data shown in Figure 5 as 3D surfaces are shown as a scatter plot (60 plots). The regression function can be used to convert Q-PCR to GeneChip and vice versa, with a level of certainty indicated by coefficient of correlation. It is noted that Cyp1a1 and Cyp1a2 became saturated above about 400 copies per cell in GeneChip system (indicated in pink plots). Cyp7a1 showed high linearity, indicating that the variation shown by the split +1sd and -1sd surfaces in Figure 5 reflected biological (animal) variation, not measurement errors.

Another important contribution of Percellome analysis is in the area of archived data in private and public domains. Firstly, Percellome data are the result of a simple linear transformation of the raw microarray data; preserving the distribution and order of the probe set data. Therefore, parametric or non-parametric methods should be able to align the data distribution and generate estimates of mRNA copy number of the non-spiked archival samples. Any archival samples that are re-measurable by Percellome method will greatly increase the accuracy of estimation. Secondly, percellome can provide appropriate bridging information between old and new versions of Affymetrix GeneChips, such as human HU-95 and HU-133, murine MU-74v2 and MOE430 series. This should also facilitate comparisons between newly generated and archived data.

The Percellome method was developed for a large-scale toxicogenomics project [13] using the Affymetrix GeneChip system. It was intended to compile a very large-scale database of comprehensive gene expression profiles in response to various chemicals from a series of experiments conducted over an extended time period. However, the method also proved to be useful for small-scale platforms such as 96 well plate-based Q-PCRs as shown above, and probably for small-scale targeted microarrays. In both cases, highly inducible or highly transcribed genes are likely to be selected. Therefore, the expression profiles may differ significantly among samples such that profile-dependent normalization (e.g. global normalization) may not be applicable. In such cases, the profile-independent nature of the Percellome method provides a robust normalization.

To demonstrate the profile-independence of the Percellome method, we chose an extreme case – the uterotrophic response assay (cf. Figure 6). The treated uteri were composed of hypertrophic cells with abundant cytoplasm whereas the untreated uteri were composed of hypoplastic cells with scant cytoplasm. This indicates that the uteri of untreated ovariectomized mice were quiescent, and that a majority of the inducible genes were probably transcriptionally inactive. Therefore, the identification of most genes as being induced by 2-fold or greater is reasonable and expected. In most *in vivo* experiments, the gene profiles of the samples are much more similar. However, there is always a set of genes that is found to be "increased" when analyzed on a "per one cell" basis that are declared to be "decreased" by global type normalization, or vice versa. Such increase/decrease calls made by the global type normalization can differ according to the normalization parameters. In both cases, the Percellome method can inform the researcher how much the expression profiles are distorted by the treatment, such as in the case of the uterotrophic assay. We also note that *in vitro* experiments such as cell-based studies tend to generate data similar to that of uterotrophic experiment.

Conclusion

Percellome data can be compared directly among samples and among different studies, and between different platforms, without further normalization. Therefore, "percellome" normalization can serve as a standard method for exchanging and comparing data across different platforms and among different laboratories. We hope that the Percellome method will contribute to transcriptome-based studies by facilitating data exchanges among laboratories.

Methods

Animal experiments

C57BL/6 Cr Slc (SLC, Hamamatsu, Japan) mice maintained in a barrier system with a 12 h photoperiod were used in this study. For the liver transcriptome experiments, twelve week-old male mice were given a single dose of the test compound by oral gavage, and the liver was sampled at 2, 4, 8 and 24 h post-gavage. For the uterotrophic experiment, 6 week old female mice were ovariectomized 14 days prior to the 7 day repeated subcutaneous injection of a test compound [12]. Animals were euthanized by exsanguination under ether anesthesia and the target organs were excised into ice-cooled plastic dishes. Tissue blocks weighing 30 to 60 mg were placed in an RNase-free 2 ml plastic tube (Eppendorf GmbH., Germany) and soaked in RNAlater (Ambion Inc., TX) within 3 min of the beginning of anesthesia. Three animals per treatment group were used and individually subjected to transcriptome measurement.

Sample homogenate preparation

The tissue blocks soaked in RNAlater were kept overnight at 4°C or until use. RNAlater was replaced in the 2 ml plastic tube with 1.0 ml of RLT buffer (Qiagen GmbH., Germany), and the tissue was homogenized by adding a 5 mm diameter Zirconium bead (Funakoshi, Japan) and shaking with a MixerMill 300 (Qiagen GmbH., Germany) at a speed of 20 Hz for 5 min (only the outermost row of the shaker box was used).

Direct DNA quantitation

Three separate 10 μl aliquots were taken from each sample homogenate to another tube and mixed thoroughly. A final 10 μl aliquot therefrom was treated with DNAse-free RNase A (Nippon Gene Inc., Japan) for 30 min at 37°C, followed by Proteinase K (Roche Diagnostics GmbH., Germany) for 3 h at 55°C in 1.5 ml capped tubes. The aliquot was transferred to a 96-well black plate. PicoGreen fluorescent dye (Molecular Probes Inc., USA) was added to each well, shaken for 10 seconds four times and then incubated for 2 min at 30°C. The DNA concentration was measured using a 96 well fluorescence plate reader with excitation at 485 nm and emission at 538 nm. λ phage DNA (PicoGreen Kit, Molecular Probes Inc., USA) was used as standard. Measurement by this PicoGreen method and the standard phenol extraction method correlated well (coefficient of correlation = 0.97, data not shown). The smallest sample size for reproducible and reliable DNA quantitation is about 5,000 cells that corresponds to a 6.75 dpc mouse embryo.

The grade-dosed spike cocktail (GSC)

The following five *Bacillus subtilis* RNA sequences were selected from the gene list of Affymetrix GeneChip arrays (AFFX-ThrX-3_at, AFFX-LysX-3_at, AFFX-PheX-3_at, AFFX-DapX-3_at, and AFFX-TrpnX-3_at) present in the MG-U74v2, RG-U34, HG-U95, HG-U133, RAE230 and MOE430 arrays: thrC, thrB genes corresponding to nucleotides 248–2229 of X04603; lys gene for diami-

nopimelate decarboxylase corresponding to nucleotides 350–1345 of X17013; pheB, pheA genes corresponding to nucleotides 2017–3334 of M24537, dapB, jojF, jojG genes corresponding to nucleotides 1358–3197 of L38424; TrpE protein, TrpD protein, TrpC protein corresponding to nucleotides 1883–4400 of K01391. The corresponding cDNAs were purchased from ATCC, incorporated into expression vectors, amplified in *E. coli* and transcribed using the MEGAscript kit (Ambion Inc., TX). The mRNA was purified using a MACS mRNA isolation kit (Miltenyi Biotec GmbH., Germany). The concentrations of spike RNAs in the GSC were in threefold steps, from 777.6 pM for AFFX-ThrX-3_at, 259.4 pM for AFFX-LysX-3_at, 86.4 pM for AFFX-PheX-3_at, 28.8 pM for AFFX-DapX-3_at, to 9.6 pM for AFFX-TrpnX-3_at. In general, the ratio depends on the linear range of the measurement system and the available number of spikes.

Setting of the "spike factor" and addition of GSC to a sample homogenate according to its DNA concentration

The GSC was added to the sample homogenates in proportion to their DNA concentrations, assuming that all cells contain a fixed amount of genomic DNA (g/cell) across samples. The amount of GSC added to each sample G (l) was given as

$$G = C * v * f \quad (l),$$

where C is the DNA concentration (g/l), v(l) is the volume of homogenate further used for RNA extraction and f (l/g) is the "spike factor", which is an adjustment factor to ensure that the sample is properly spiked by the GSC (cf. Figure 3). Spike factors have been pre-determined for various organs/tissues to reflect differences in their total RNA/genomic DNA ratios (cf. Table 1). In this way, five spike mRNA signals can properly cover the linear dose-response range of the platform. In practice, for the Affymetrix GeneChips, the spike factor is set so that the five GSC spikes cover the range of "Present" calls given by the Affymetrix system, which corresponds to approximately 80 to 7000 in raw readouts given by the Affymetrix MAS5.0 software. A raw readout of 10 by the current Affymetrix GeneChip system corresponds to approximately one copy per cell in mouse liver (spike factor = 0.2), whereas in mouse thymus (spike factor = 0.01) it corresponds to approximately 0.05 copy per cell. For Q-PCR, the same spike factor corresponds to Ct values ranging approximately from 17 to 27, which is well within the linear range of Q-PCR (data not shown).

"Per cell" normalization (Percellome normalization)

Since murine haploid genomic DNA is made of 2.5×10^9 base pairs and one base pair is approximately 600 Daltons (Da), the haploid genomic DNA weighs 1.5×10^{12} Da, corresponding to

$d = 5 \times 10^{-12}$ (g DNA per diploid cell).

Therefore, the cell number per liter of the sample homogenate (N) is given as

$N = C/d$ (cells/l)

where C is the DNA concentration (g/l).

On the other hand, the copy numbers of GSC RNAs in the homogenate are given as follows:

if Sj (mole/l) (j = 1,2,3,4,5) is the mole concentration of one of the five spike RNAs in the GSC solution and G(l) is the amount of GSC added to each homogenate, the mole concentrations of the spike RNAs in the homogenate (CSj) are given as,

$CSj = Sj*C*f$ (mole/l).

The GSC RNAs in moles per cell (MSj) are given as

$MSj = CSj/N$

$= Sj*C*f/(C/d)$

$= Sj*f*d$ (mole/cell)

The copy numbers of the GSC RNAs per cell (NSj) are given as

$NSj = MSj*A$

$= Sj*f*d*A$ (copies per diploid cell)

where A is Avogadro's number.

As a result, the GSC spikes AFFX-TrpnX-3_at, AFFX-DapX-3_at, AFFX-PheX-3_at, AFFX-LysX-3_at and AFFX-ThrX-3_at correspond approximately to 5.8, 17.3, 52.0, 156.0 and 468.1 copies per cell (per diploid DNA template) for mouse liver sample homogenates, where the spike factor = 0.2. It is our observation that the RNA/DNA ratios are virtually constant across polyploid hepatocytes (data not shown).

For each Q-PCR plate or GeneChip, the coefficients, α, β, γ and δ of functions {1} or {2} are determined from the GSC values using the least-square method. The signal values or Ct values of all the other mRNAs measured are then converted to copy numbers per cell by {3} or {4}, i.e. the inverses of functions {1} or {2}.

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	GCCAATGATATGGCAGCTTCTAC	TGCGGCAGCATGACCATTA
AFFX-LysX-3_at	CCGCTTCATGCCACTGAATAC	CCGGTTCGATCCAAATTTCC
AFFX-ThrX-3_at	CCTGCATGAGGATGACGAGA	GGCATCGGCATATGGAAAC
Ahr_1450695_at	CAGAGACCACTGACGGATGAA	AGCCTCTCCGGTAGCAAACA
Cyp1a1_1422217_a_at	TGCTCTTGCCACCTGCTGA	GGAGCACCCTGTTTGTTTCTATG
Cyp1a2_1450715_at	CCTCACTGAATGGCTTCCAC	CGATGGCCGAGTTGTTATTG
Cyp1b1_1416612_at	GCCTCAGGTGTGTTTGATGGA	AGTACAGCCCTGGTGGGAATG
Cyp7a1_1422100_at	TTCTACATGCCCTTTGGATCAG	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 Super-Script 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

1. 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.
2. Hill AA, Brown EL, Whitley MZ, Tucker-Kellogg G, Hunter CP, Slonim DK: **Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls.** *Genome Biol* 2001, **2**:. RESEARCH0055
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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と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はじ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-

urements in Gene Expression Data. *Genome Informatics*, **16** (2)：183-194, 2005.

菅野　純／Jun KANNO
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

腎原病学

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

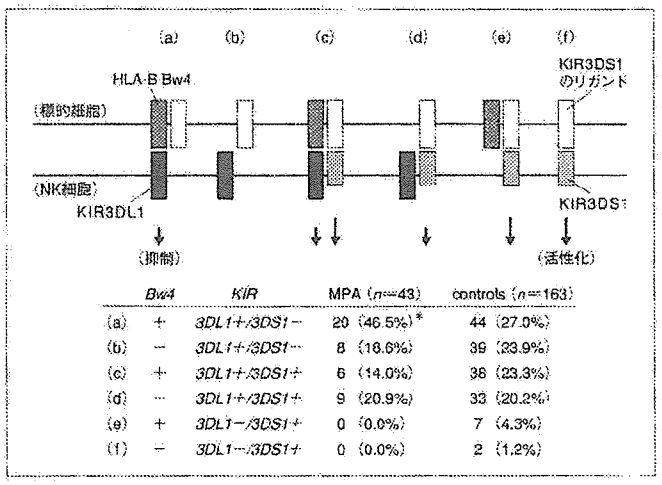
これまでにいくつかの自己免疫疾患やウイルス感染症の関連研究において，*KIR* 遺伝子多型との関連を示す報告がある[2]．アメリカ人集団において血管炎を合併した関節リウマチや，乾癬性関節炎などの自己免疫疾患では *KIR2DS1*，*KIR2DS2* など活性化型 *KIR* 遺伝子が患者群において高い頻度で観察されている．また，AIDS の進行が遅いタイプの *KIR* 遺伝子型や，C 型肝炎ウイルスのクリアランスにかかわる *KIR-HLA* 遺伝子型の組合せなど，ウイルス感染症の臨床経過との関連も報告されている．

*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



	Bw4	*KIR*	MPA (*n*=43)	controls (*n*=163)
(a)	+	*3DL1+/3DS1−*	20 (46.5%)*	44 (27.0%)
(b)	−	*3DL1+/3DS1−*	8 (18.6%)	39 (23.9%)
(c)	+	*3DL1+/3DS1+*	6 (14.0%)	38 (23.3%)
(d)	−	*3DL1+/3DS1+*	9 (20.9%)	33 (20.2%)
(e)	+	*3DL1−/3DS1+*	0 (0.0%)	7 (4.3%)
(f)	−	*3DL1−/3DS1+*	0 (0.0%)	2 (1.2%)

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あるが，*KIR3DS1* は *KIR3DL1* 遺伝子座における活性化型アリルであり，細胞外ド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 Kitajima, Ken-ichi Aisaki, 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@nihs.go.jp
1985年東京医科歯科大学大学院医学研究科博士課程修了．人体病理学，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年より同部長．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掌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，との観点から，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，異なったプラットホーム間，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．

注1　環境への配慮も含まれる．

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

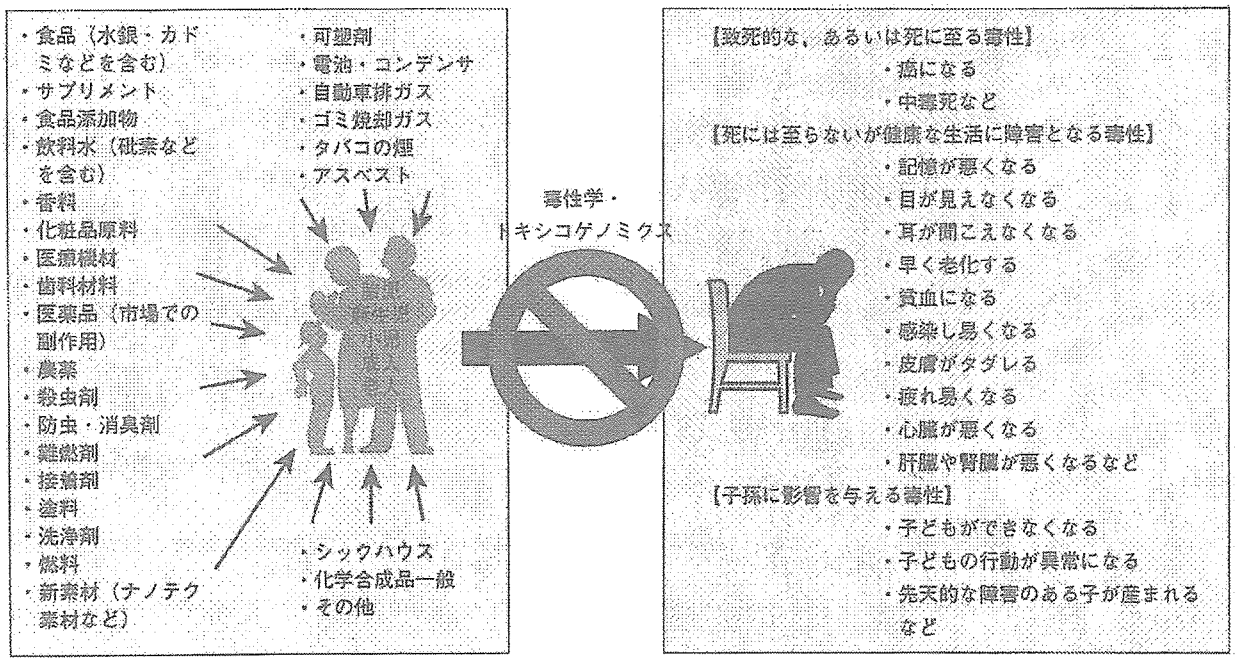


図1. 毒性学の対象

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．カク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）に関わるソフトウエアに加えて，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[3]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

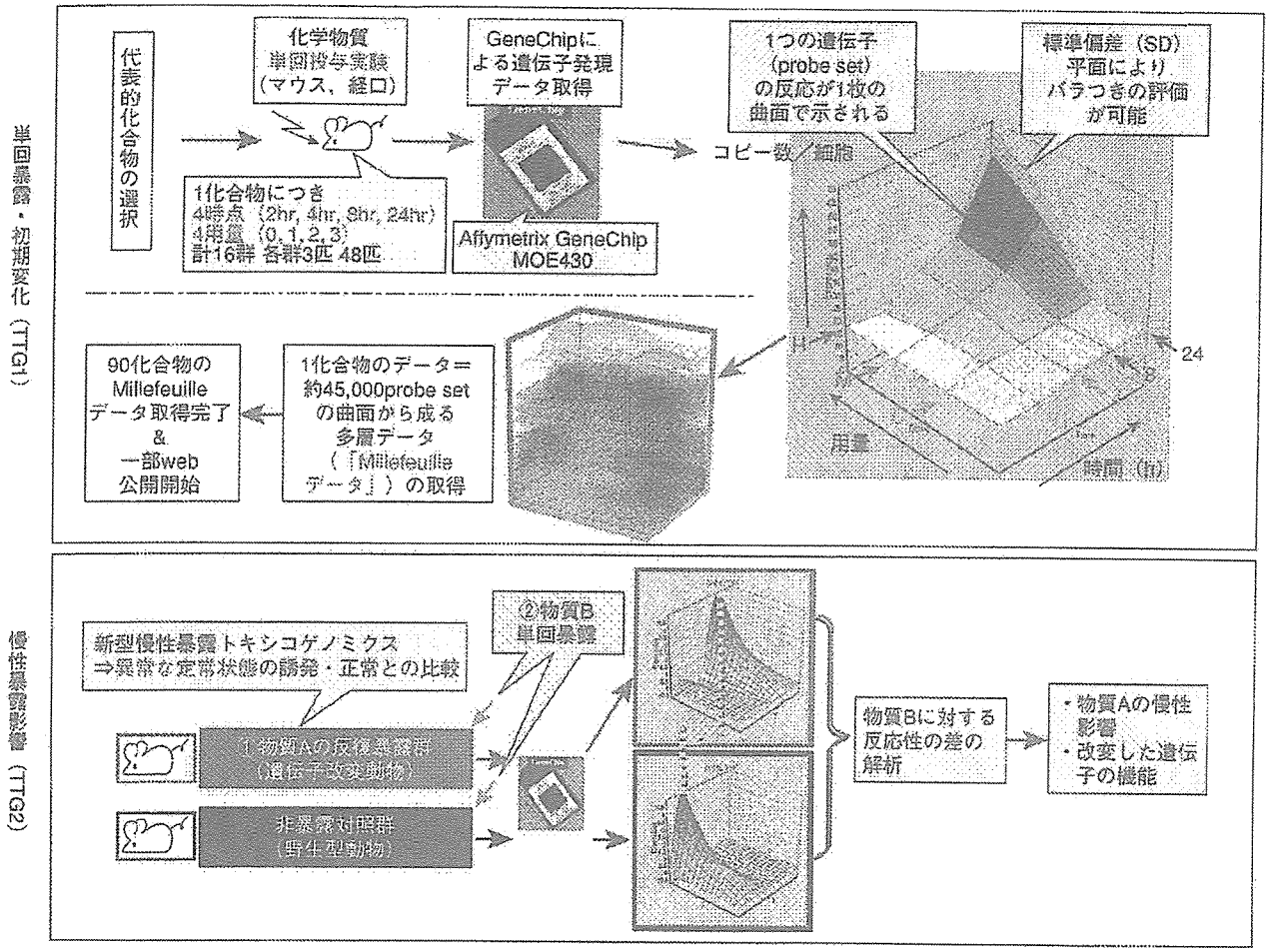


図2. Percellome法と3次元表示による多層（Millefeuille）データシステムを用いたプロジェクトの根幹部分の概要
単回投与による遺伝子発現初期変化データを90化合物について取得（上段），現在，反復投与の影響を検討中（下段）．H；高用量（high），M；中用量（medium），L；低用量（low），C；コントロール（control）．

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エア（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ホームへの適用

Percellome手法は，GSCの受け入れ条件を整えることにより，様々なプラットホ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，サンプル破砕液に，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，

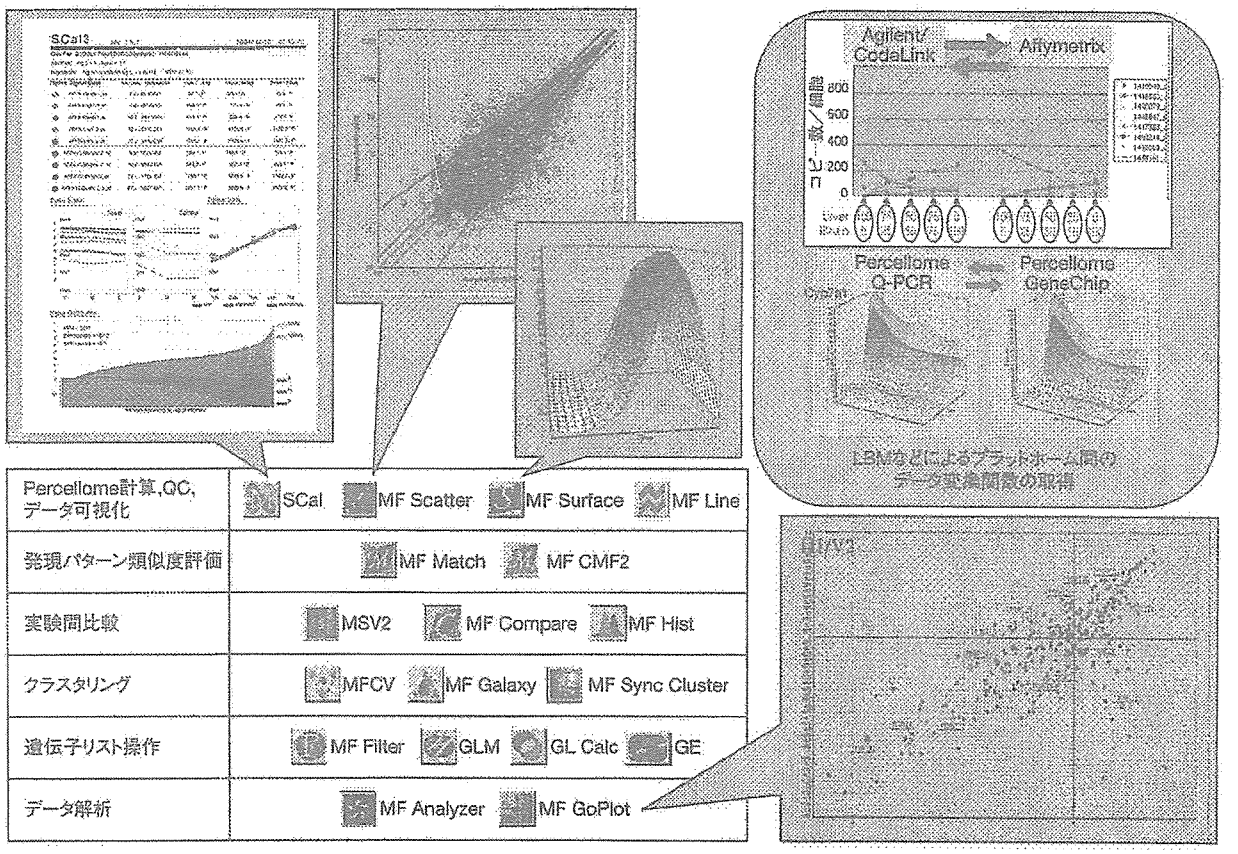


図3．3次元多層（Millefeuille）データの解析などに用いる独自開発プログラム群

品質管理とともにPrecellome計算を自動的に実施するSCal，Plotソフトウエア，3次元曲面の描画ソフト（MF Surface），など．右上はプラットホ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，両方のプラットホ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の新しいエキ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，既知のエキソンに対して設計されたprobe setでは発現が見られ，イントロン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，現状では各エキ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

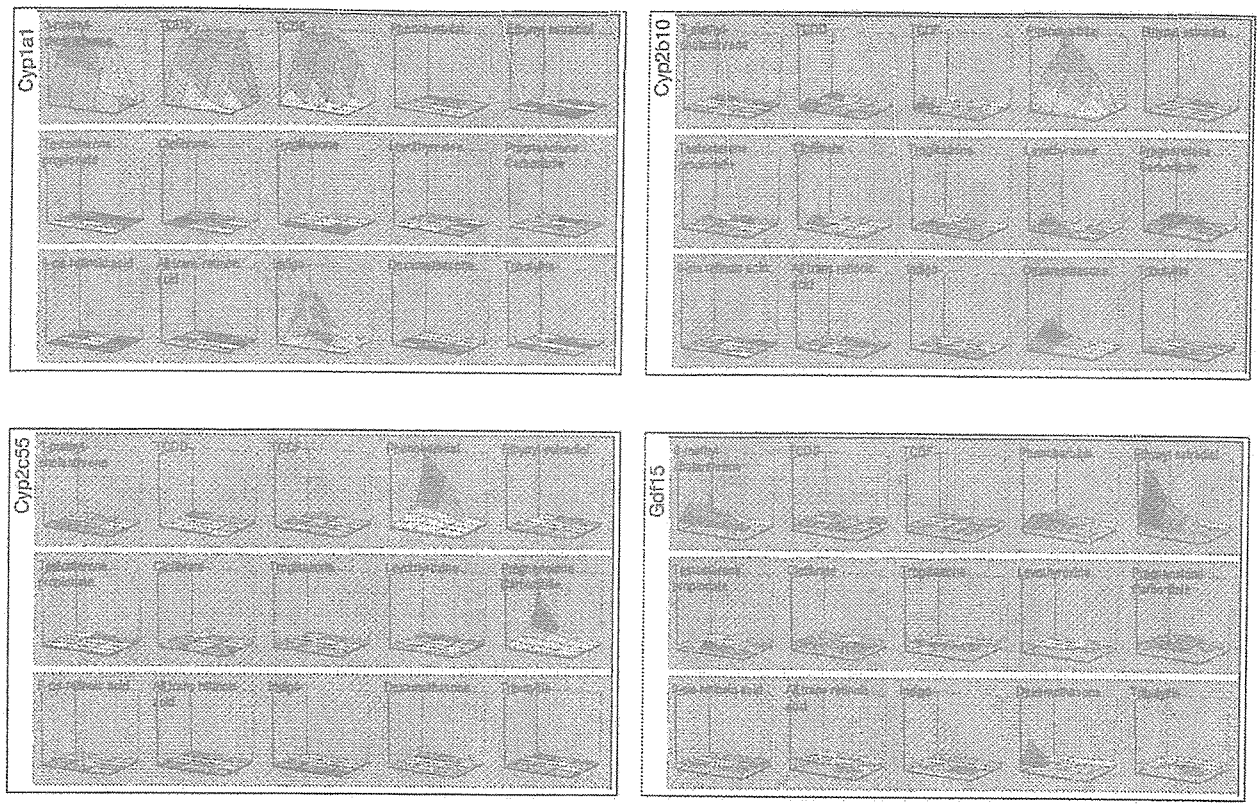


図4. 化合物間の発現比較

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Cyp1a1（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．

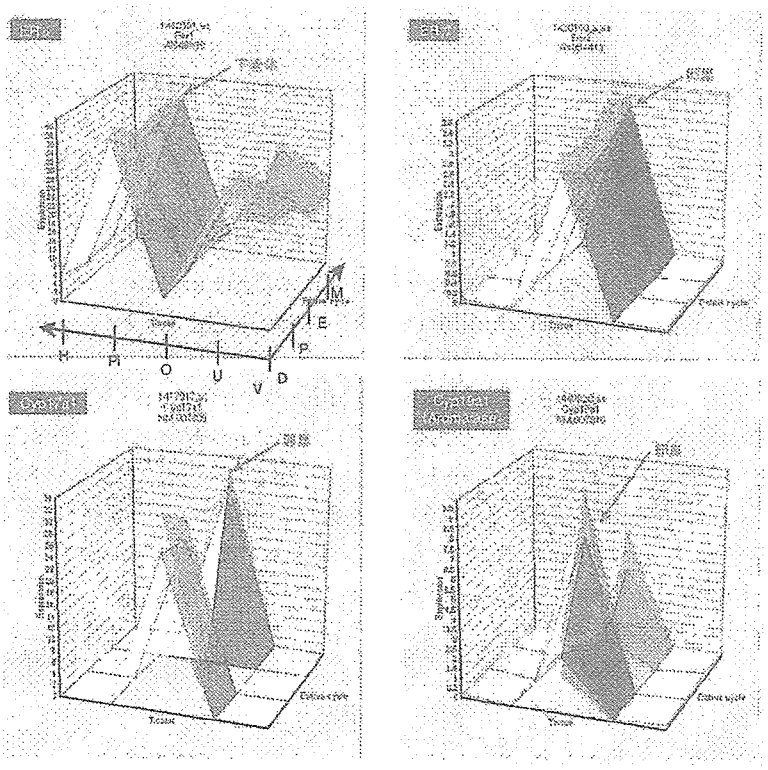


図5. 臓器間の発現比較

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）および膣（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．