Data in Brief 32 (2020) 106136

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Transcriptomic data from the rat liver after five days of exposure to legacy or emerging brominated flame retardants

Keith R. Shockley[a,*], Michelle C. Cora[b], David E. Malarkey[b], Daven Jackson-Humbles[b], Molly Vallant[c], Brad J. Collins[c], Esra Mutlu[c], Veronica G. Robinson[c], Surayma Waidyanatha[c], Amy Zmarowski[d], Nicholas Machesky[d], Jamie Richey[d], Sam Harbo[d], Emily Cheng[d], Kristin Patton[d], Barney Sparrow[d], June K. Dunnick[e]

[a] *Biostatistics & Computational Biology Branch, Division of Intramural Research, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC 27709, United States*
[b] *Cellular & Molecular Pathology Branch, Division of the National Toxicology Program (DNTP), NIEHS, Research Triangle Park, NC 27709, United States*
[c] *Program Operations Branch, DNTP, NIEHS, Research Triangle Park, NC 27709, United States*
[d] *Battelle, Columbus, Ohio 43210, United States*
[e] *Toxicology Branch, DNTP, NIEHS, Research Triangle Park, NC 27709, United States*

ARTICLE INFO

Article history:
Received 15 July 2020
Revised 28 July 2020
Accepted 30 July 2020
Available online 5 August 2020

Keywords:
Toxicity of legacy and emerging brominated flame retardants
5-Day toxicogenomic studies
Liver toxicity
Genomic benchmark dose analysis

ABSTRACT

Large-scale gene expression analysis of legacy* and emerging** brominated flame retardants were conducted in the male Harlan Sprague Dawley rat [1]. Each animal was dosed for 5 days with the chemical at concentrations of 0.1 – 1000 μmol/kg body weight per day. Following the last dose, a specimen of the left liver was removed for RNA extraction. The amplified RNA (aRNA) was fragmented and then hybridized to Affymetrix Rat Genome 230 2.0 Arrays. Each GeneChip® array was scanned using an Affymetrix GeneChip® Scanner 3000 7G to generate raw expression level data (.CEL files). Statistical contrasts were used to find pairwise gene expression differences between the control group and each dose group using the R/maanova package [2]. The transcriptomic data can be used to provide insights into

DOI of original article: 10.1016/j.toxlet.2020.07.016
* Corresponding author.
E-mail address: shockleykr@niehs.nih.gov (K.R. Shockley).

https://doi.org/10.1016/j.dib.2020.106136

the degree of toxicity, toxic mechanisms, disease pathways activated by exposure, and for benchmark dose analysis. The gene expression data for each of the nine flame retardants discussed here accompanies the research article entitled, "Comparative Toxicity and Liver Transcriptomics of Legacy and Emerging Brominated Flame Retardants following 5-Day Exposure in the Rat" [1].

* polybrominated diphenyl ether 47 (PBDE 47), decabromodiphenyl ether (decaBDE), hexabromocyclododecane (HBCD); ** 2-ethylhexyl-2,3,4,5-tetrabromobenzoate (TBB); bis(2-ethylhexyl) tetrabromophthalate (TBPH); tetrabromobisphenol A-bis(2,3-dibromopropyl ether (TBBPA-DBPE); 1,2-bis(tribromophenoxy)ethane (BTBPE); decabromodiphenylethane (DBDPE); hexachlorocyclopentadienyl-dibromocyclooctane (HCDBCO).

Specifications Table

Subject	Biology
Specific subject area	Toxicogenomics
Type of data	Table Affymetrix .CEL file data
How data were acquired	RNA extracted from the left lobe of the liver from control and treated animals after five days of dosing; microarray data generated from Rat Genome 230 version 2 Affymetrix GeneChip® arrays that were read into the R software environment (http://www.R-project.org) directly from .CEL files using the R/affy package [3].
Data format	Raw Analyzed Filtered
Parameters for data collection	Significant genes with a false discovery rate (FDR) were identified using Ingenuity pathway analysis after exposure to each of the 9 flame retardants. Benchmark dose (BMD) values and the lower bound of the 95% confidence interval of the BMD were calculated for gene sets using BMDExpress version 2.0 [4].
Description of data collection	The relative toxicity of three legacy and six emerging brominated flame retardants was studied in the male Harlan Sprague Dawley rat. Five or six animals per group were dosed with control or one of five doses of the chemical by oral gavage (0.1 – 1000 μmol/kg body weight per day). At necropsy the left lobe of the liver was taken for RNA extraction.
Data source location	Institution: NTP City/Town/Region: Durham, NC Country: United States of America
Data accessibility	The toxicogenomics results and benchmark dose analyses are provided with this article. The Affymetrix .CEL files have been deposited in the National Library of Medicine GEO data base under GEO SuperSeries accession number GSE153366 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153366).
Related research article	**The data are related to the following research article:** Shockley KR, Cora MC, Malarkey DE, Jackson-Humbles D, Vallant M, Collins BJ, Mutlu E, Robinson VG, Waidyanatha S, Zmarowski A, Machesky N, Richey J, Harbo S, Cheng E, Patton K, Sparrow B, Dunnick JK, Comparative Toxicity and Liver Transcriptomics of Legacy and Emerging Brominated Flame Retardants following 5-Day Exposure in the Rat. Toxicol. Lett. 332(2020) 222-234 https://doi.org/10.1016/j.toxlet.2020.07.016.

1. Value of the data

- These data provide liver transcriptomic changes that occur after exposure to three legacy and six emerging flame retardants and, thus, make available information for accessing the comparative toxicity across several brominated flame retardants
- These toxicogenomic data can be used by the scientific community, industry, and regulators to prioritize the need for further toxicity studies and to begin to predict long-term toxic outcomes
- These data may be used to understand what disease pathways are disrupted after short-term exposure to flame retardants and to design further studies on toxicity on a molecular level
- These data show that short-term toxicities studies can be used to provide preliminary information to understand disease mechanisms

2. Data description

The transcript data were obtained from RNA extracted from the left lobe of the liver after exposure of the Harlan Sprague Dawley rat to a chemical from one of three legacy or six emerging brominated flame retardants. Five or six rats were dosed by oral gavage once a day for five days at 0, 0.1, 1, 10, 100, or 1000 μmol/kg/day/flame retardant. Supplements 1–9 provide the transcriptomics data for each of the nine flame retardants. Supplement 10 provides benchmark dose analyses for this data.

Fig. 1 presents principal component analyses for the nine flame retardant liver transcriptomic datasets for 5-day exposure in rats. Each plot shows the first two principal components (PCs) from each dataset, with PC1 on the x-axis and PC2 on the y-axis. The percentage of variance explained by each PC is indicated on each axis. Each dataset contains multiple samples from the vehicle control and 5 dose groups (0.1, 1, 10, 100, and 1000 μmol/kg). The magnitudes of the coefficients of each PC refer to the relative importance of each sample and the sign of the coefficients of each PC describe whether there is positive or negative correction between the sample and the PC. In Fig. 1, the coefficients of PCs ranged from approximately −0.4 to 0.4. For PBDE47, PC1 clearly separates the samples exposed to the highest dose (1000 μmol/kg) and the second highest dose (100 μmol/kg) from the rest of the exposure doses. However, the first two PCs do not show a clear separation of samples for any of the other chemicals.

3. Experimental design, materials and methods

3.1. Flame retardant chemicals

This study compared the toxicity of three legacy flame retardants and six emerging flame retardants. The legacy flame retardants included polybrominated diphenyl ether 47 (PBDE 47, CAS RN: 5436–43–1), decabromodiphenyl ether (decaBDE, CAS RN: 1163–19–5), and hexabromocyclododecane (HBCD, CAS RN: 25637–99–4; 3194–55–6). The emerging flame retardants included 2-ethylhexyl-2,3,4,5-tetrabromobenzoate (TBB, CAS RN: 183658–27–7), bis(2-ethylhexyl) tetrabromophthalate (TBPH, CAS RN: 26040–51–7), tetrabromobisphenol A-bis(2,3-dibromopropyl ether (TBBPA-DBPE, CAS RN: 21850–44–2), 1,2-bis(tribromophenoxy)ethane (BTBPE, CAS RN: 37853–59–1), decabromodiphenylethane (DBDPE, CAS RN: 84852–53–9), and hexachlorocyclopentadienyl-dibromocyclooctane (HCDBCO, CAS RN: 51936–55–1). Chemical identities were confirmed, and chemical purities were determined prior to use in the studies [1]. The oral dose formulations were prepared in corn oil vehicle (Spectrum, New Brunswick, New Jersey) and were determined to be within 10% of the target concentration. Prior to study initiation and at the end of the study the formulations were within 10% of the day 0 value.

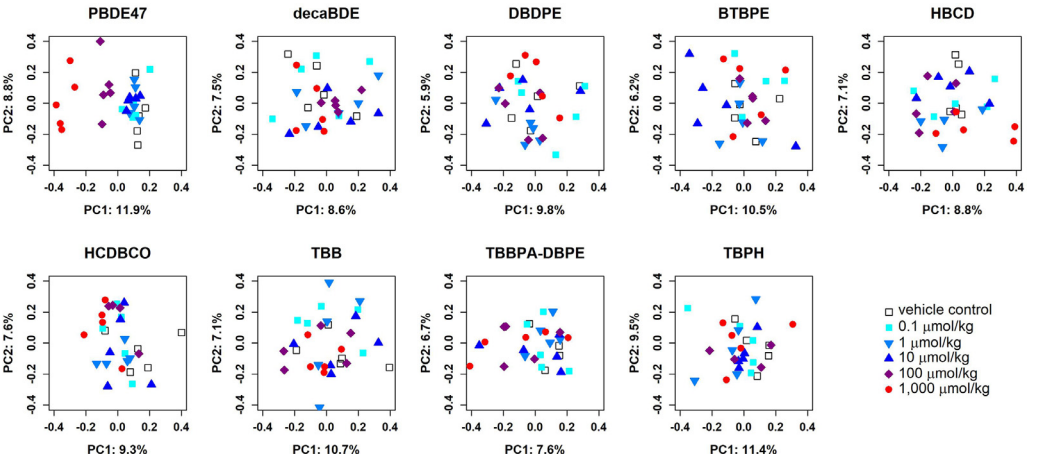

Fig. 1. PCA plots for the nine flame retardants. The percentage of total variance explained by each principal component is shown on each axis.

3.2. Experimental design

Flame retardant chemicals were administered to seven-week-old male Harlan Sprague Dawley rats once daily via oral gavage for 5 consecutive days using 6 rats per dose level per chemical (except for 5 rats/dose in the 100 μmol/kg group for BTBPE and the vehicle control for TBB). The animals were randomized to study group by body weight so that group body weights were similar. The animals were given NTP-2000 diet (Zeigler Brothers, Gardners, PA) and tap water *ad libitum*. The flame retardants were given in corn oil at 5 ml/kg body weight to deliver doses of: 0, 0.1, 1, 10, 100, 1000 μmol/kg body weight per day per flame retardant. The animal studies were conducted according to the guidelines of the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC) and approved by the local Animal Care and Use Committee and according to the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

Necropsy was conducted after five days of dosing (one day after the last dose). Animals were anesthetized with CO_2/O_2 anesthesia (approximately 70/30% mixture). Blood was collected from the vena cava or aorta. Samples for hematology analysis were collected into tubes containing tripotassium ethylenediaminetetraacetic acid (K3 EDTA), and samples for clinical chemistry and thyroid hormone analysis were collected into serum collection tubes without anticoagulant, centrifuged, and the serum harvested.

At necropsy, a portion of the left lobe of the liver for microarray analysis.

3.3. RNA collection

The liver tissue sample was placed in a cryotube containing RNAlater™ and stored at 2–8 °C overnight. After overnight storage the RNAlater™ was removed and the samples stored at −60 °C to −80 °C until processed for RNA isolation for use in microarray analysis. RNA isolation was performed on all liver tissue samples preserved in RNAlater™, using a Qiagen RNeasy Mini Kit (Qiagen Inc., Valencia, CA) with a deoxyribonucleic acid (DNA) digestion step. The concentration and purity of each RNA sample was calculated from UV absorbance readings (A_{260} and A_{280}) obtained using a NanoDrop Spectrophotometer (ND-1000, NanoDrop Products; Thermo Scientific, Wilmington, DE). All samples were also evaluated for RNA integrity using an Agilent RNA 6000 Nano Chip kit with an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA). The RIN was based on the amount of degradation in a sample and was presented as a number between one and ten with one being the most degraded and ten being the most intact RNA.

3.4. Microarray analysis

Microarray expression analysis was performed using total RNA isolated from liver samples. RNA meeting the following acceptance criteria were used for microarray expression analysis: RNA samples with a concentration $\geq$ 35 ng/μL and purity A_{260}/A_{280} ratios ranging between 1.80–2.20 and an RNA integrity number (RIN) $\geq$ 7.0. The RNA was amplified through cDNA synthesis, *in vitro* transcription and biotin labeling using GeneChip® 3′ IVT PLUS Reagent Kit (Affymetrix, Santa Clara, CA). One hundred nanograms of total RNA was amplified as directed in the Affymetrix 3′ IVT Plus kit protocol. Fifteen micrograms of amplified biotin-aRNAs was fragmented, and 12.5 μg was hybridized to each array for 16 h at 45 °C in a rotating hybridization oven using the Affymetrix Eukaryotic Target Hybridization Controls and protocol. The amplified RNA (aRNA) was fragmented then hybridized to Affymetrix Rat Genome 230 2.0 arrays. Arrays were stained and then washed using the GeneChip Hybridization, Wash and Stain Kit according to the user manual. Each GeneChip® array was scanned using an Affymetrix GeneChip® Scanner 3000 7 G to generate microarray image data (.DAT files) and raw expression level data (.CEL files).

All samples within a single flame-retardant group were randomly sorted into a processing order to prevent batch effects. Also, throughout the RNA isolation, *in vitro* transcription, hybridization, fluidics and scanning procedures, a single operator handled all samples within each flame-retardant group.

3.5. Microarray data normalization

Probe intensity data from all Rat Genome 230 version 2 Affymetrix GeneChip® arrays were read into the R software environment (http://www.R-project.org) directly from .CEL files using the R/affy package [3]. Each data set was comprised of six independent samples from six different treatment groups (0, 0.1, 1, 10, 100, or 1000 µmol/kg body weight) for a total of 36 samples in each data set. However, there were only five samples in the 100 µmol/kg body weight dose group for BTBPE and in the vehicle control group for TBB, so that BTBPE and TBB data sets each had a total of 35 samples. Probe-level data quality was assessed using image reconstruction, box plots of raw signal intensity, and histograms of raw signal intensities. Normalization was carried out using the robust multi-array average (RMA) method separately for each data set [5]. Briefly, the RMA method adjusts the background of perfect match (PM) probes, applies a quantile normalization of the corrected PM values, and calculates final expression measures using the Tukey median polish algorithm. RMA scatterplots were used as an additional quality control measure.

To describe the data here, a principal component analysis (PCA) was performed separately for each dataset by mean centering and scaling the data so that each transcript has mean 0 and standard deviation 1 and employing the svd() function in R [6]. Gene expression measures for all probe sets were included for each PCA analysis. Principal components (PCs) represent linear combinations of the original variables (samples) that explain a percentage of the total variance in the dataset, arranged in decreasing order of importance. The coefficients of the PCs describe the relative contributions of each experiment (or sample) to the PC.

3.6. Statistical assessment of differential gene expression

Statistical contrasts were used to find pairwise gene expression differences between the control group and each dose group using the R/maanova package [2]. For each flame retardant, the model

$$Y_i = \mu + \mathrm{DOSE} + \varepsilon_i \quad (1)$$

was used to fit the $\log_2$ transformed gene expression measures Y_i, where µ is the mean for each array, DOSE is a six-level factor representing the dose effect (0, 0.1, 1, 10, 100, or 1000 μmol/kg body weight) and ε_i captures random error for probe set *i*. A total of five different comparisons were tested for each probe set (0 vs 0.1 μmol/kg, 0 vs 1 μmol/kg, 0 vs 10 μmol/kg, 0 vs 100 μmol/kg, and 0 vs 1000 μmol/kg). All statistical tests were performed using F_s, a modified F-statistic incorporating shrinkage estimates of variance components [7]. P-values were calculated by permuting sample labels 1000 times. In order to reduce the number of false positives, p-values were adjusted for multiple hypothesis testing corresponding to all probe sets on the array using the Benjamin-Hochberg false discovery rate (FDR) procedure implemented using the p.adjust() function in R. This correction controls the expected proportion of errors among the significant results [8]. Unless otherwise noted, an FDR threshold of 0.05 was used for statistical significance. Log_2 fold changes were calculated by subtracting the control (0 μmol/kg) and dose treated (0.1 μmol/kg, 1 μmol/kg, 10 μmol/kg, 100 μmol/kg, or 1000 μmol/kg) relative expression values from model (1) above [9].

Over-represented gene sets were determined from the gene list obtained above by testing for association with gene pathway relationships (www.ingenuity.com). Enrichment of pathway members among differentially expressed probe sets were assessed using the one-tailed Fisher exact test for 2×2 contingency tables.

3.7. Benchmark dose analysis

The benchmark dose (BMD) is defined as the dose corresponding to a predetermined change in response referred to as the benchmark response (BMR). Liver transcriptomic data were used to calculate the BMD and the lower bound of the 95% confidence interval of the BMD using BMDExpress version 2.0 [4]. All BMD calculations were performed within the BMDExpress framework separately for each chemical.

The data corresponding to the control-AFFX probe sets were first removed from each data set. Then, a classical one-way ANOVA was used to filter the remaining RMA-normalized probe set intensities to find transcripts that were differentially expressed across dose groups with a P value < 0.05 and |fold-change| ≥ 1.5. In this way, probe sets that did not respond to treatment were removed from the analysis. Next, the Hill, power, linear, second-degree polynomial, third-degree polynomial, and a set of four exponential models were fit to the data for each remaining probe set. The BMR level was set to 1.349 standard deviations above or below the control group, representing a 10% increase over control response rate that is standard in BMD analysis [10]. For the linear, second-degree polynomial, and third-degree polynomial cases, a nested likelihood ratio test was used to select the best model fit. The more complex model was selected if the fit was improved ($P < 0.05$), but the less complex model was selected if the fit was not improved ($P \geq 0.05$). The lowest Akaike information criterion (AIC) was used to select the best fitting model comparing the remaining models with the best nested model. The power parameter was restricted to ≥ 1 for all model fitting to avoid infinite slope at the origin. Hill model fits were not selected if the estimated dose at half maximal response was less than 1/3 of the lowest positive dose, and the next best model was selected instead.

The calculated BMD values was used as input data for Gene Ontology (GO) analyses. When more than one probe set mapped to the same Entrez ID, the BMD values were averaged across probe sets to obtain a single expression value for each Entrez ID. Probe sets that mapped to more than one Entrez ID were removed from the analysis. The resulting Entrez IDs were matched to Biological Process GO terms as a basis for gene set definitions. The output consists of a range of summary exposure levels (mg/kg/day) representing the central tendencies and variability of BMD and BMDL based on the calculated BMD and BMDL values for the genes in a category.

Ethics statement

All authors made substantial contributions to the concept and design of the study, or acquisition of the data, or analysis of the data, or analysis of the data, and drafting of the article. The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

The animal experiments were carried out in accordance with the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised 1978).

Declaration of Competing Interest

All of the authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

This work was supported [in part] by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (ZIA ES103166 and ZIA ES103316). In addition, work

was performed for the National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services, under contracts HHSN273201400020C (MRI Global, Kansas City, MO), HHSN273201400022C (RTI International, RTP, NC), HHSN273201400027C (Battelle, Columbus, OH), HHSN273201400015C (Battelle Columbus) and HHSN316201200054W (ASRC Federal, Research Triangle Park, North Carolina, USA). We thank Dr. A. Merrick and Dr. R. Cannon for the review of the paper. However, the statements, opinions or conclusions contained therein do not necessarily represent the statements, opinions or conclusions of NTP, NIEHS, NIH or the United States government.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.106136. Supplements 1-9 contain liver toxicogenomic data for the nine flame retardants. Supplement 10 contains benchmark dose analyses for the nine flame retardants.

References

[1] K.R. Shockley, M.C. Cora, D.E. Malarkey, D. Jackson-Humbles, M. Vallant, B.J. Collins, E. Mutlu, V.G. Robinson, S. Waidyanatha, A. Zmarowski, N. Machesky, J. Richey, S. Harbo, E. Cheng, K. Patton, B. Sparrow, J.K. Dunnick, Comparative toxicity and liver transcriptomics of legacy and emerging brominated flame retardants following 5-day exposure in the rat, Toxicol. Lett. 332 (2020) 222–234, doi:10.1016/j.toxlet.2020.07.016.

[2] H. Wu, M. Kerr, X. Cui, G. Churchill, MAANOVA: a software package for the analysis fo spotted cDNA mciroarray experiments, in: G.E. Parmgiani G, RA Irizarry, SL Zeger (Eds.), The Analysis of Gene Expression data: an Overview of Methods and Software, Springer, 2003, pp. 313–431.

[3] L. Gautier, L. Cope, B.M. Bolstad, R.A. Irizarry, affy–analysis of affymetrix genechip data at the probe level, Bioinformatics 20 (2004) 307–315, doi:10.1093/bioinformatics/btg405.

[4] J.R. Phillips, D.L. Svoboda, A. Tandon, S. Patel, A. Sedykh, D. Mav, B. Kuo, C.L. Yauk, L. Yang, R.S. Thomas, J.S. Gift, J.A. Davis, L. Olszyk, B.A. Merrick, R.S. Paules, F. Parham, T. Saddler, R.R. Shah, S.S. Auerbach, BMDExpress 2: enhanced transcriptomic dose-response analysis workflow, Bioinformatics 25 (2019) 1780–1782, doi:10.1093/bioinformatics/bty878.

[5] R.A. Irizarry, B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf, T.P. Speed, Exploration, normalization, and summaries of high density oligonucleotide array probe level data, Biostatistics 4 (2003) 249–264, doi:10.1093/biostatistics/4.2.249.

[6] R Core Team, A Language and Environment For Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2019 https://www.r-project.org/.

[7] X. Cui, J.T. Hwang, J. Qiu, N.J. Blades, G.A. Churchill, Improved statistical tests for differential gene expression by shrinking variance components estimates, Biostatistics 6 (2005) 59–75, doi:10.1093/biostatistics/kxh018.

[8] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Statist. Soc. B 57 (1995) 289–300, doi:10.1111/j.2517-6161.1995.tb02031.x.

[9] G.A. Churchill, Using ANOVA to analyze microarray data, BioTechniques 37 (2004) 173–175 177 https://doi.org/10.2144/04372TE01 .

[10] L. Yang, B.C. Allen, R.S. Thomas, BMDExpress: a software tool for the benchmark dose analyses of genomic data, BMC Genomics 8 (2007) 387, doi:10.1186/1471-2164-8-387.