Data Article

Dataset of transcriptomic changes that occur in human preadipocytes over a 3-day course of exposure to 3,3′,4,4′,5-Pentachlorobiphenyl (PCB126)

Francoise A. Gourronc\textsuperscript{a}, Brynn K. Helm\textsuperscript{b}, Larry W. Robertson\textsuperscript{c}, Michael S. Chimenti\textsuperscript{d}, Hans-Joachim Lehmler\textsuperscript{c}, James A. Ankrum\textsuperscript{e,f,*}, Aloysius J. Klingelhutz\textsuperscript{a,f,*}

\textsuperscript{a} Department of Microbiology and Immunology, University of Iowa, 3-612 BSB, 51 Newton Road, Iowa City, IA 52242, United States
\textsuperscript{b} Program in Molecular Medicine, University of Iowa, United States
\textsuperscript{c} Department of Occupational and Environmental Health, University of Iowa, United States
\textsuperscript{d} Iowa Institute of Human Genetics, Bioinformatics Division, University of Iowa, United States
\textsuperscript{e} Roy J. Carver Department of Biomedical Engineering, University of Iowa, United States
\textsuperscript{f} Fraternal Order of Eagles Diabetes Research Center, United States

\textbf{ABSTRACT}

Exposure to polychlorinated biphenyls (PCBs) has been associated with the development of metabolic syndrome, a cluster of diseases that includes obesity, diabetes, liver steatosis, and cardiovascular problems. PCBs accumulate and fat and are known to act on adipocytes and their precursors, termed preadipocytes. The PCB congener, PCB126, has been shown to activate the aryl hydrocarbon receptor (AhR) as well as proinflammatory genes. Here, we used RNAseq to assess gene transcript changes that occur in PCB126-exposed human preadipocytes over a time course. RNA was collected from 4 replicates of PCB126-exposed and control-treated preadipocytes at 9 h, 24 h, and 72 h post-exposure. RNA was processed for RNAseq analysis using a NovaSeq 6000 with an obtained minimum of 25 million paired-end 50 bp reads.
per sample. Reads were aligned using the salmon aligner and transcript expression values were summarized to the gene level using tximport. Gene transcript level counts comparing treated- versus control-treated cells were used for differential expression analysis using DESeq2. Differential expression Excel tables (one for each time point) were generated displaying average differential expression (log2 fold change) of the 4 replicates of treated versus control samples with cutoffs of 0.3 log2 fold change (increase or decrease) and p-values of less than 0.05. FastQ, raw, and differential expression tables were uploaded to GEO. A heat map of genes that were changed in common across all time points was generated using GraphPrism. The data generated from this analysis provides a full transcriptional profile of changes that occur over time in preadipocytes that have been exposed to PCB126. The rich datasets can be mined by other researchers to understand how PCB126 and other dioxin-like compounds, including other PCB congeners such as PCB77 and PCB118, affect biological pathways in preadipocytes and other cell types to cause disease.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

### Specifications Table

| Subject                      | Health, Toxicology and Mutagenesis |
|------------------------------|------------------------------------|
| Specific subject area        | Temporal gene expression changes in preadipocytes caused by exposure to PCB126 |
| Type of data                 | Table                              |
| How the data were acquired   | Data was acquired by performing RNA-sequencing using an Illumina NovaSeq 6000 genome sequencer. |
| Data format                  | Raw -Fastq                         |
|                             | Analyzed – “Raw Counts after alignment” |
|                             | Filtered Differential Gene Expression |
| Description of data collection | Immortalized normal human preadipocytes (NPAD) from a non-diabetic female donor were plated and cultured until 90% confluent. The cells were then treated with either 10 μM PCB126 or DMSO as a vehicle control. After 9, 24, and 72 h, cells were harvested for RNA-sequencing analysis. 4-replicates of each condition were collected. |
| Data source location         | • Institution: University of Iowa  |
|                             | • City/Town/Region: Iowa City, Iowa |
|                             | • Country: USA                     |
| Data accessibility           | Repository name: Gene Expression Omnibus (GEO) |
|                             | Data identification number: GSE193578 |
|                             | Direct URL to data:                |
|                             | https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE193578 |
|                             | Code to reproduce the DEG analysis: |
|                             | https://github.com/mchimenti/klingelhutz_rnaseq_july2020_pcb126 |
| Related research article     | F.A. Gourronc, B.K. Helm, L.W. Robertson, M.S. Chimenti, H.J. Lehmler, J.A. Ankrum, A.J. Klingelhutz, Transcriptome sequencing of 3,3′,4,4′,5-Pentachlorobiphenyl (PCB126)-treated human preadipocytes demonstrates progressive changes in pathways associated with inflammation and diabetes., Toxicology in Vitro, Volume 83, 2022, 105396, https://doi.org/10.1016/j.tiv.2022.105396, [1] |
Value of the Data

- The data provided here represents the first RNAseq data reported for exposure of human cells to PCB126, an important dioxin-like persistent organic pollutant known to activate the aryl hydrocarbon receptor (AhR).
- The data provides a full transcriptional profile of changes that occur over time in preadipocytes that have been exposed to PCB126. Changes that occur over time could reveal novel pathways that are activated directly or secondarily upon exposure to PCB126.
- Toxicologists and molecular biologists will be able to mine this data to unravel the cascade of gene pathways that are activated and altered upon acute exposure to PCB126 to help identify how and why PCB126 contributes to metabolic disease progression.
- Datasets could be used to generate hypothesis for how other dioxin-like compounds such as PCB118 and PCB77 alter preadipocyte gene expression.
- Gene expression changes found here in human preadipocytes could be compared to other experiments that used primary or immortalized rodent cell lines to study PCB126. These studies could provide important insight into how different species respond similarly or differently to PCB126 exposure.
- The gene expression changes found to be elicited by PCB126 exposure of human preadipocytes provide clues as to how exposure to PCB126 or other dioxin-like compounds disrupts adipose function to cause metabolic syndrome.

1. Data Description

Normal human preadipocytes (NPAD) derived from subcutaneous adipose tissue were exposed to 10 μM PCB126 or DMSO over a time course and then subjected to RNA-Sequencing. The 10 μM concentration was chosen based on our previous studies demonstrating that this concentration was non-cytotoxic but caused significant inhibition of adipogenesis [2]. The specific time points of 9 h, 24 h, and 72 h were chosen to define the temporal changes in gene expression that occur after PCB126 exposure. This time frame is based on our previous findings demonstrating induction of the cytochrome p450 genes at the early time point of 9 h with a delayed inflammatory response that started at 24 h that was near maximal by 72 h [3]. Raw data obtained from an Illumina NovaSeq 6000 sequencer was converted to fastq format before being deposited in the Gene Expression Omnibus (GEO) database (Accession number: GSE193578). Table 1 outlines each raw data file and describes the treatment condition, either DMSO or 10 μM PCB126, as well as the duration of exposure to the treatment condition, 9-72 h. In addition to the raw counts in Table 1, the data was aligned to hg38 reference genome and the number of aligned reads and number of genes identified for each sample is reported in Table 2. After alignment, differentially expressed genes (DEG) were identified by comparing the 10 μM PCB126 treated NPAD data to the DMSO treated NPAD data for each of the 3 exposure durations. The list of DEG was filtered to only include genes that showed a log fold change ≥ 0.3 & adjusted p-value ≤ 0.05. Lists of raw counts for every gene and filtered DEG for each exposure duration and their corresponding log fold change and p-value are available in the files listed in Table 3 and can be found on GEO Accession number: GSE193578.

DEGs that had a p<0.05 and ≥0.3 log2 fold change comparing PCB126 treated to DMSO treated at all time points were selected, and a heatmap of their log fold change was generated showing those same genes for all 3 exposure durations (Fig. 1). Seventy-three genes were significantly upregulated and eighteen significantly downregulated at all three time points. The majority of these genes exhibited progressive upregulation or downregulation, respectively, over time. Many of the genes that were progressively upregulated are AhR-responsive genes (e.g., CYP1B1, TIPARP, CYP1A1) or proinflammatory genes (e.g., IL1B, IL1A, LIF).
### Table 1
List of accession number for each transcriptome in GEO database.

| Sample    | Treatment Condition | Exposure Duration | GEO Accession Number |
|-----------|---------------------|-------------------|---------------------|
| Veh_1_9H | DMSO                | 9 h               | GSM5814526          |
| Veh_2_9H | DMSO                | 9 h               | GSM5814527          |
| Veh_3_9H | DMSO                | 9 h               | GSM5814528          |
| Veh_4_9H | DMSO                | 9 h               | GSM5814529          |
| X126_1_9H| 10 μM PCB126        | 9 h               | GSM5814530          |
| X126_2_9H| 10 μM PCB126        | 9 h               | GSM5814531          |
| X126_3_9H| 10 μM PCB126        | 9 h               | GSM5814532          |
| X126_4_9H| 10 μM PCB126        | 9 h               | GSM5814533          |
| Veh_1_Day1| DMSO               | 24 h              | GSM5814534          |
| Veh_2_Day1| DMSO               | 24 h              | GSM5814535          |
| Veh_3_Day1| DMSO               | 24 h              | GSM5814536          |
| Veh_4_Day1| DMSO               | 24 h              | GSM5814537          |
| X126_1_Day1| 10 μM PCB126      | 24 h              | GSM5814538          |
| X126_2_Day1| 10 μM PCB126      | 24 h              | GSM5814539          |
| X126_3_Day1| 10 μM PCB126      | 24 h              | GSM5814540          |
| X126_4_Day1| 10 μM PCB126      | 24 h              | GSM5814541          |
| Veh_1_Day3| DMSO               | 72 h              | GSM5814542          |
| Veh_2_Day3| DMSO               | 72 h              | GSM5814543          |
| Veh_3_Day3| DMSO               | 72 h              | GSM5814544          |
| Veh_4_Day3| DMSO               | 72 h              | GSM5814545          |
| X126_1_Day3| 10 μM PCB126      | 72 h              | GSM5814546          |
| X126_2_Day3| 10 μM PCB126      | 72 h              | GSM5814547          |
| X126_3_Day3| 10 μM PCB126      | 72 h              | GSM5814548          |
| X126_4_Day3| 10 μM PCB126      | 72 h              | GSM5814549          |

### Table 2
Summary statistics of reads mapping for each sample after alignment.

| Sample    | # of Mapped Reads | # of Genes Mapped |
|-----------|-------------------|-------------------|
| Veh_1_9h  | 29,352,009        | 20,093            |
| Veh_2_9h  | 26,035,064        | 19,695            |
| Veh_3_9h  | 40,601,343        | 20,699            |
| Veh_4_9h  | 29,735,917        | 19,934            |
| X126_1_9h | 33,029,096        | 20,172            |
| X126_2_9h | 31,377,156        | 20,089            |
| X126_3_9h | 33,340,380        | 20,066            |
| X126_4_9h | 32,423,386        | 20,179            |
| Veh_1_Day1| 29,733,576        | 20,070            |
| Veh_2_Day1| 36,013,685        | 20,586            |
| Veh_3_Day1| 33,340,688        | 20,388            |
| Veh_4_Day1| 35,876,505        | 20,442            |
| X126_1_Day1| 35,647,938      | 20,429            |
| X126_2_Day1| 38,586,351       | 20,386            |
| X126_3_Day1| 33,551,975       | 20,100            |
| X126_4_Day1| 34,912,103       | 20,194            |
| Veh_1_Day3| 29,804,511        | 20,198            |
| Veh_2_Day3| 31,061,250        | 20,422            |
| Veh_3_Day3| 34,618,101        | 20,645            |
| Veh_4_Day3| 25,673,758        | 19,992            |
| X126_1_Day3| 33,644,941       | 20,143            |
| X126_2_Day3| 32,968,110        | 20,236            |
| X126_3_Day3| 32,465,246        | 19,879            |
| X126_4_Day3| 29,879,815        | 19,840            |
Fig. 1. Heatmap of DEG changes over time in PCB126 treated preadipocytes compared to DMSO treated preadipocytes. DEG with a $p < 0.05$ at all exposure durations were selected, and their log fold change was plotted in a heat map to show the progression of gene changes from 9-72 h. Color map represents log fold change, with a negative fold change representing a decrease in gene expression and a positive fold change representing an increase in gene expression in PCB126 exposed cells compared to DMSO exposed cells.
Table 3
Processed data files after alignment and differentially expressed gene analysis.

| File Name | Description of Analysis | Exposure Duration |
|-----------|-------------------------|-------------------|
| GSE193578_processed_raw_counts_all_genes_vehicle_PCB126.csv.gz | Raw Counts after alignment | All |
| GSE193578_IPG-DEG_pcb126_9hour_vs_veh_9hour.xlsx | Differential Gene Expression between DMSO and PCB126 treated cells. Filtered to include genes with log fold change ≥ | 9 h |
| GSE193578_IPG-DEG_pcb126_dayone_vs_veh_dayone.xlsx | | 24 h |
| GSE193578_IPG-DEG_pcb126_daythree_vs_veh_daythree.xlsx | | 72 h |

2. Experimental Design, Materials and Methods

2.1. Human Subcutaneous Preadipocyte Exposure and RNA-Extraction

For this dataset, an immortalized human preadipocyte line (NPAD) was used for all conditions. The cells were originally isolated from subcutaneous fat harvested from a non-diabetic female donor. The isolated cells were then immortalized and characterized as previously described [4]. NPADs were cultured on tissue culture plastic in Preadipocyte Growth Media (PGM2, Lonza) throughout the duration of the treatment. For treatment, 80,000 cells were plated in a 6-well plate and cultured until 90% confluent. Once near confluent, the media was changed to PGM2 supplemented either with dimethyl sulfoxide (DMSO) or 10 μM PCB126 (C12H5Cl5; InChI: 1S/C12H5Cl5/c13-8-2-1-6(3-9(8)14)7-4-10(15)12(17)11(16)5-7/h1-5H; InChI Key: REHONNLRQRTTF-UHFFFAOYSA-N; Canonical SMILES: 1 = CC(=C(=C(C=C)C)C)C)C)C)C)C)C)) dissolved in DMSO. The level of DMSO was held constant in all conditions at a level of 0.1% (v/v). The media remained on the cells until RNA harvesting. At each harvesting time point, the media was removed and cells were harvested using 1 mL of TRIzol Reagent (Invitrogen). Cells and TRIzol were pipetted several times to homogenize the sample and then transferred to Eppendorf tubes for further processing using a Qiagen RNEasy kit to isolate RNA for each sample. Each treatment condition and treatment duration was repeated 4 times to provide biological replicates.

2.2. RNA Library Preparation and Sequencing

To prepare RNA libraries for sequencing, 500 ng of RNA from each sample were incubated with oligo(dT) primer coated beads to enrich for polyA-containing transcripts. The enriched RNA pool was then fragmented, converted to cDNA, and ligated to indexes using an Illumina TruSeq stranded mRNA preparation kit (Cat. #RS-122-2101, Illumina). An Agilent Bioanalyzer 2100 was used to measure the molar concentrations of the indexed libraries and samples were pooled for sequencing. The concentration of each pool was measured using an Illumina Library Quantification kit (KAPA Biosystems) and sequenced on an Illumina NovaSeq 6000 genome sequencer at the Iowa Institute of Human Genetics. A minimum of 25 million paired-end 50 bp reads per sample were obtained.

2.3. Data Processing and Differential Expression Analysis

Reads were then converted from the native Illumina BCL format to fastq and processed using the `bcbio-nextgen` pipeline ([https://github.com/chapmanb/bcbio-nextgen](https://github.com/chapmanb/bcbio-nextgen)). For
alignment, the hg38 human genome was used as a reference and reads were aligned using hisat2 aligner [5,6]. For all samples ~95% of RNA-seq reads uniquely mapped to the reference genome with ~90% of mapped reads residing in an exonic region. For quality control, the ‘bcbio-nextgen’ pipeline runs MultiQC and returned no significant quality control problems for any of the samples reported here. In addition, the salmon aligner was used to quantify reads to the human transcriptome (GENCODE 39). Salmon transcript expression values were mapped to the gene level count estimates using tximport as described in (http://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html#transcript-abundance-files-and-tximport-tximeta) [7]. Non-normalized gene-level counts were then used as input to the DESeq2 model to identify differentially expressed genes (DEG) between DMSO and PCB126 treated NPADS at each exposure duration (DESeq2 performs within- and between-sample normalization internally) [8]. The code to reproduce the analysis can be found here: https://github.com/mchimenti/klingelhutz_rnaseq_july2020_pcb126. Excel tables displaying differentially expressed genes (log fold change ≥ |0.3| & adjusted p-value ≤ 0.05) comparing PCB126-treated versus DMSO-treated samples (averaged across 4 replicates of each) at the same time point were generated. The adjusted p-value was obtained using the Benjamin-Hochberg Method as implemented in DESeq2. A heatmap showing common differentially expressed genes across all time points was generated using GraphPrism.

**Ethics Statement**

The work described in this manuscript adheres to ethical publishing standards. The work did not involve human subjects. The immortal NPAD cell line used in this study has been published [4] and was derived from de-identified primary preadipocytes purchased from Lonza and obtained by consent.

**Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Data Availability**

Human Preadipocyte Response to PCB126 (Original data) (Gene Expression Omnibus).

**CRediT Author Statement**

Francoise A. Gourronc: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft; Brynn K. Helm: Investigation; Larry W. Robertson: Resources, Funding acquisition; Michael S. Chimenti: Data curation, Writing – original draft; Hans-Joachim Lehmler: Resources, Funding acquisition; James A. Ankrum: Conceptualization, Writing – original draft, Funding acquisition; Aloysius J. Klingelhutz: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft, Supervision, Funding acquisition.

**Acknowledgments**

The RNaseq sample data collection and data analyses were performed in the Genomics and Bioinformatics Divisions of the Iowa Institute of Human Genetics, respectively, which is supported, in part, by the University of Iowa Carver College of Medicine and the Holden Comprehensive Cancer Center (National Cancer Institute of the National Institutes of Health under
References

[1] F.A. Gourronc, B.K. Helm, L.W. Robertson, M.S. Chimenti, H. Joachim-Lehmler, J.A. Ankrum, A.J. Klingelhutz, Transcriptome sequencing of 3,3′,4,4′,5-Pentachlorobiphenyl (PCB126)-treated human preadipocytes demonstrates progressive changes in pathways associated with inflammation and diabetes, Toxicol. In Vitro: Int. J. Publ. Assoc. BIBRA 83 (2022) 105396, doi:10.1016/j.tiv.2022.105396.

[2] G. Gadupudi, F.A. Gourronc, G. Ludewig, L.W. Robertson, A.J. Klingelhutz, PCB126 inhibits adipogenesis of human preadipocytes, Toxicol. In Vitro: Int. J. Publ. Assoc. BIBRA 29 (1) (2015) 132–141, doi:10.1016/j.tiv.2014.09.015.

[3] F.A. Gourronc, L.W. Robertson, A.J. Klingelhutz, A delayed proinflammatory response of human preadipocytes to PCB126 is dependent on the aryl hydrocarbon receptor, Environ. Sci. Pollut. Res. Int. 25 (17) (2018) 16481–16492, doi:10.1007/s11356-017-9676-z.

[4] B.G. Vu, F.A. Gourronc, D.A. Bernlohr, P.M. Schlievert, A.J. Klingelhutz, Staphylococcal superantigens stimulate immortalized human adipocytes to produce chemokines, PloS One 8 (10) (2013) e77988, doi:10.1371/journal.pone.0077988.

[5] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, Nat. Methods 12 (4) (2015) 357–360, doi:10.1038/nmeth.3317.

[6] D. Kim, J.M. Paggi, C. Park, C. Bennett, S.L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, Nat. Biotechnol. 37 (8) (2019) 907–915, doi:10.1038/s41587-019-0201-4.

[7] C. Soneson, M.I. Love, M.D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences, F1000Research 4 (2015) 1521, doi:10.12688/f1000research.7563.2.

[8] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (12) (2014) 550, doi:10.1186/s13059-014-0550-8.