

## Supplementary Information for

### Metabolic effects of air pollution exposure and reversibility

Sanjay Rajagopalan<sup>a,d†</sup>, Bongsoo Park<sup>b†</sup>, Rengasamy Palanivel<sup>a</sup>, Vinesh Vinayachandran<sup>a</sup>, Jeffrey A. DeIuliis<sup>a</sup>, Roopesh Singh Gangwar<sup>a</sup>, Lopa Das<sup>a</sup>, Jinhu Yin<sup>b</sup>, Youngshim Choi<sup>b</sup>, Sadeer Al-Kindi<sup>a</sup>, Mukesh K. Jain<sup>a,d</sup>, Kasper D. Hansen<sup>c</sup> and Shyam Biswal<sup>b\*</sup>

<sup>a</sup> Cardiovascular Research Institute, Case Western Reserve University, Cleveland, OH 44106; <sup>b</sup> Department of Environmental Health and Engineering; <sup>c</sup> Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205; <sup>d</sup> Harrington Heart and Vascular Institute, University Hospital Cleveland Medical Center, Cleveland, OH 44106.

†Contributed equally as first author.

#### Corresponding authors:

Sanjay Rajagopalan [sxr647@case.edu](mailto:sxr647@case.edu)

Shyam Biswal [sbiswal1@jhu.edu](mailto:sbiswal1@jhu.edu)

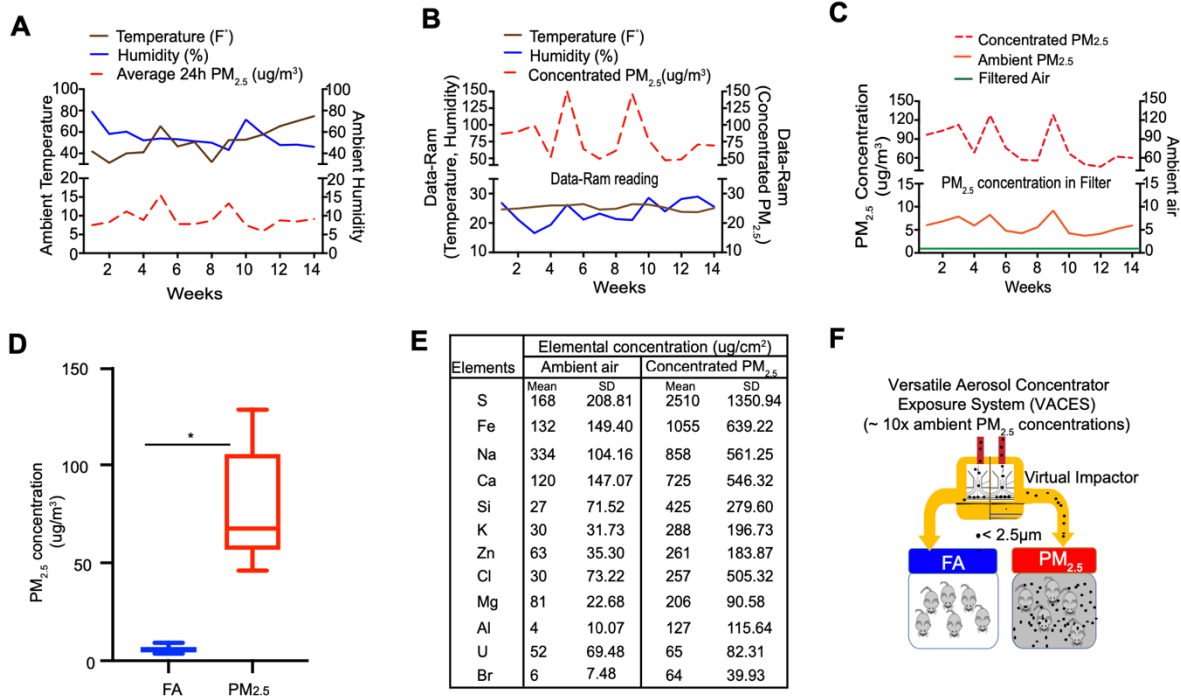
#### This PDF file includes:

Figures S1 to S15

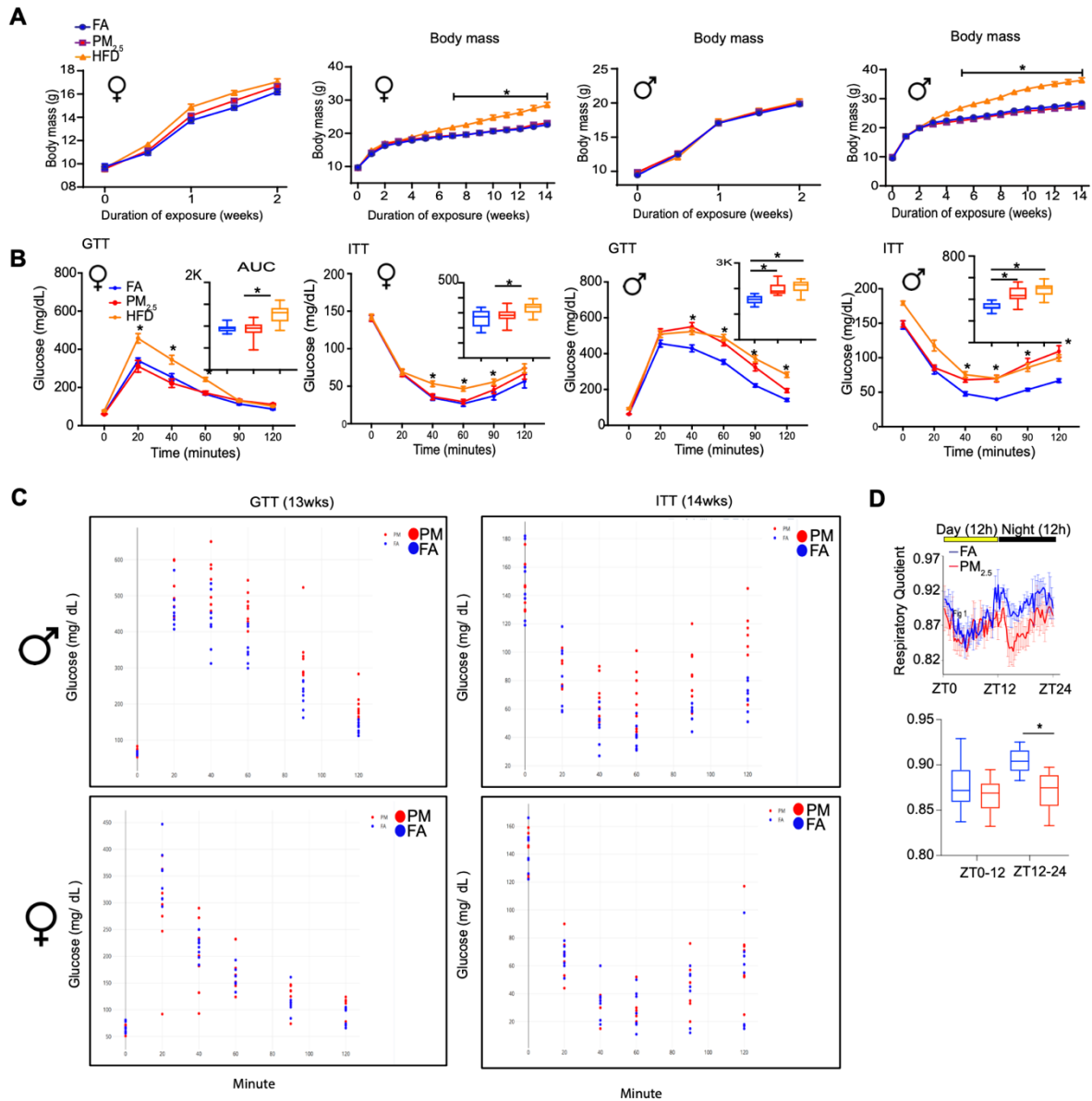
Legends for Datasets S1 to S15

Methodological Details

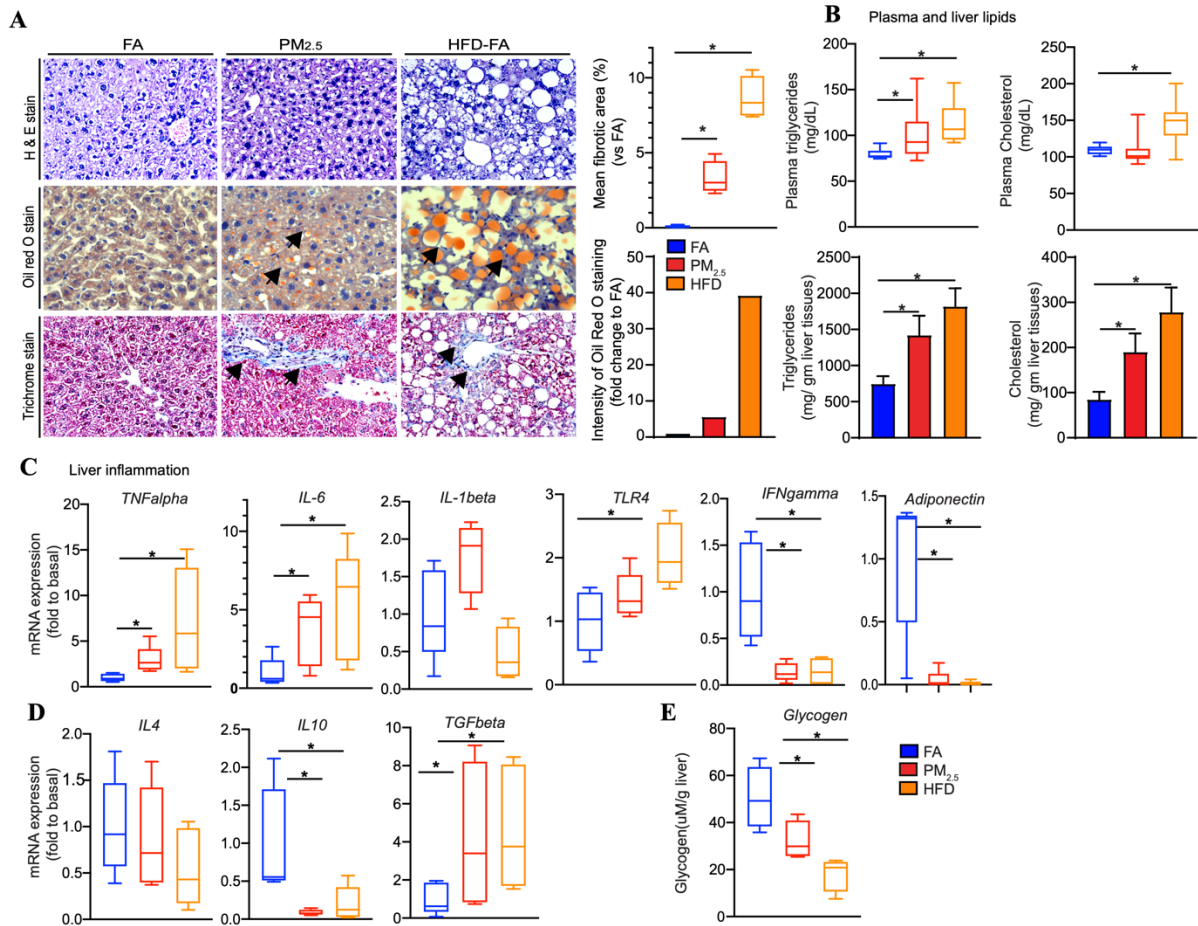
Tables T1 to T2



**Figure S1: Details of air pollution exposure** (A) Ambient temperature, humidity and average daily air quality during the time of exposure. The datasets were derived from the website [airnow.gov](http://airnow.gov) and [timeanddate.com](http://timeanddate.com) (show average weather conditions in Cleveland, Ohio). (B) Data-Ram readings using high sensitivity real-time telemetry, (DataRAM4, Thermo Scientific) indicates weekly averages of exposure chamber concentrations, temperature and humidity. (C) Average, exposed level of PM<sub>2.5</sub> (based on filter concentration) in the PM<sub>2.5</sub> and filtered air groups. (D) The average filter PM<sub>2.5</sub> concentration ( $\mu\text{g}/\text{m}^3$ ) during the entire duration of exposure. (E) Chemical composition of air pollution particles in ambient and concentrated air. All data are depicted as means SEM. (F) Schematic diagram illustrating PM<sub>2.5</sub> and filtered air (FA) exposure in mice.

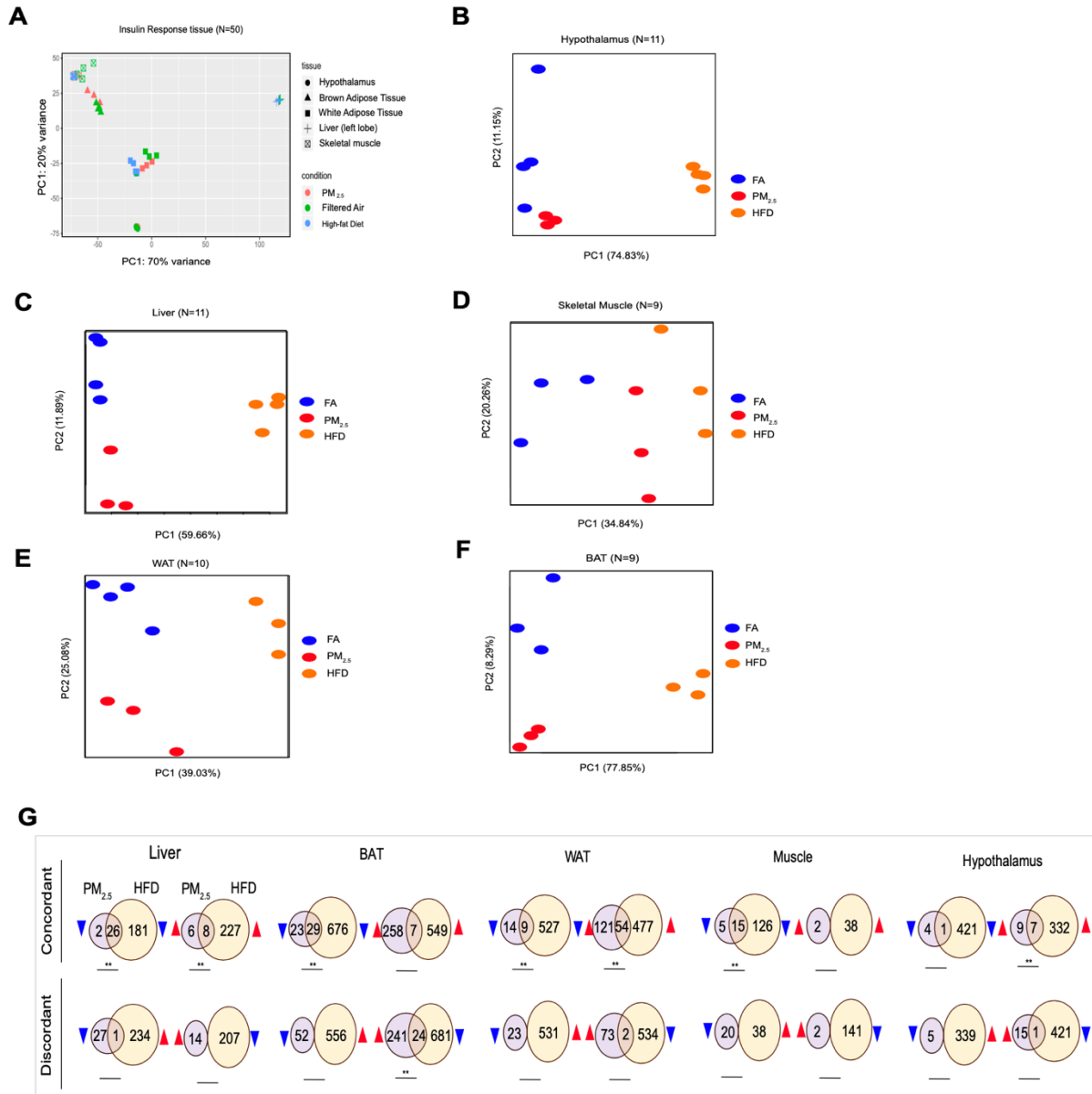


**Figure S2: The summary of glucose tolerance and insulin tolerance test** (A) Body mass of female and male mice exposed to either FA or PM<sub>2.5</sub> for 14 weeks (n=12/group). (B) Glucose tolerance test (GTT) and insulin tolerance test (ITT) in male (n=12) and female mice (n=12) of FA, PM<sub>2.5</sub> and HFD groups. All data are depicted as means  $\pm$  SEM. \* $<0.05$ , by unpaired t-test vs. mice fed ND and exposed with FA. (C) Dot plot depicts the individual variability in the GTT and ITT results in the FA and PM<sub>2.5</sub> mice. (D) Respiratory quotient. Line graphs indicate averages of day and night cycles over a 48h period, bar graphs indicate the total value at the time points indicated (n=6).

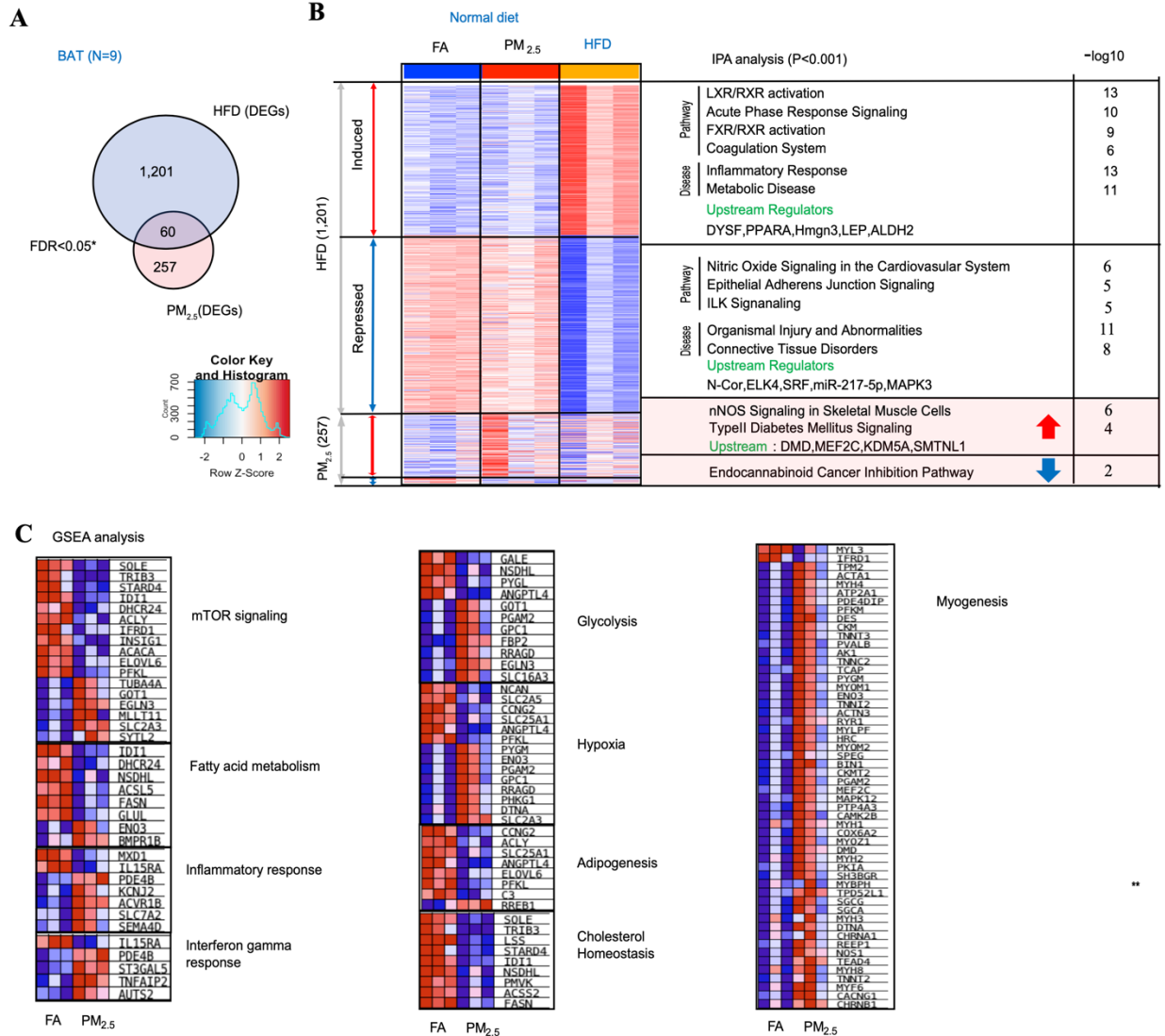


**Figure S3: Impact of PM<sub>2.5</sub> exposure and high-fat diet (HFD) on hepatic pathology and inflammation.**

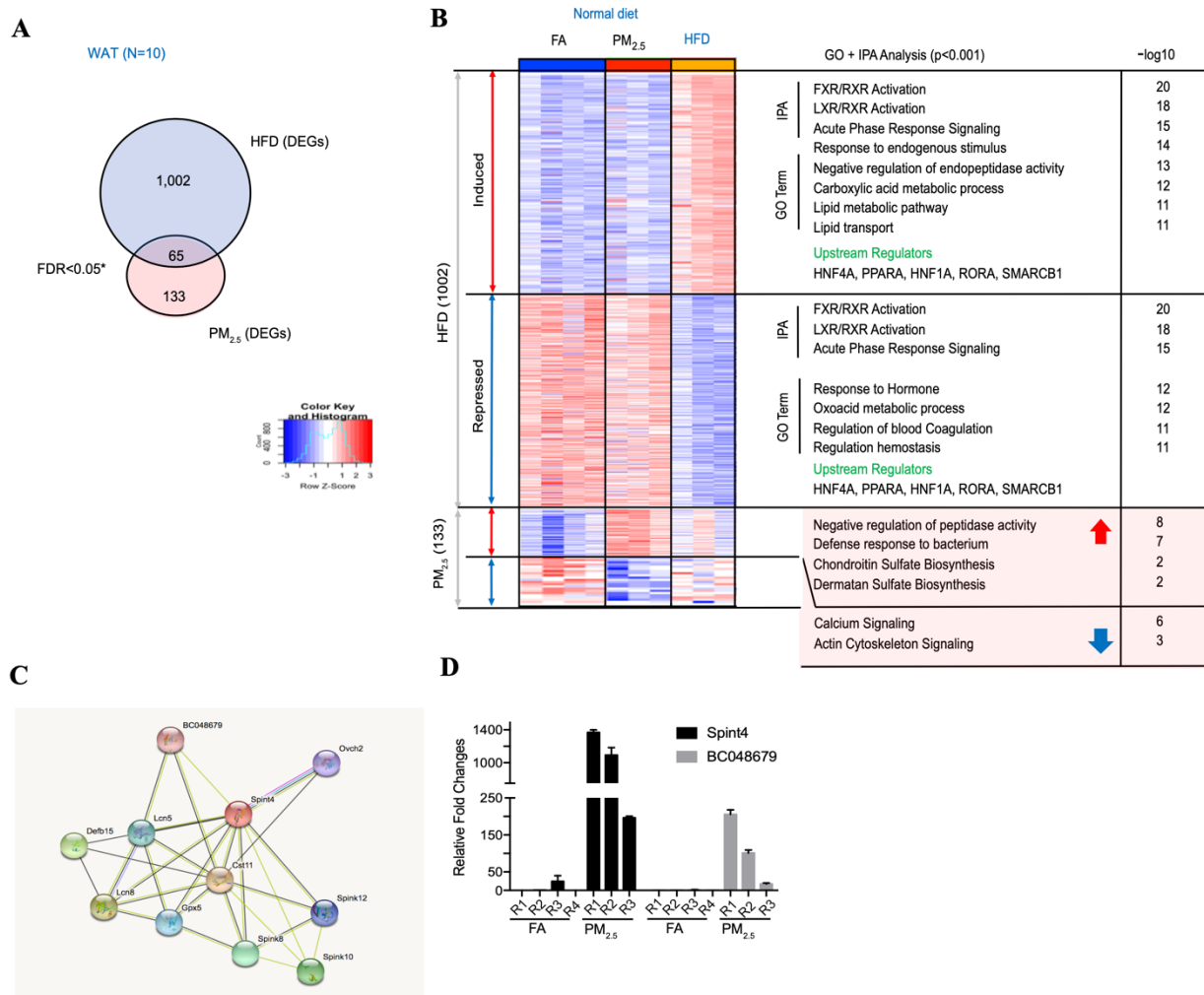
(A) Representative images of hematoxylin and eosin (arrow indicates lipid deposition hepatocyte enlargement), oil red O (arrow indicates lipid droplets in red), and Masson trichrome staining (arrow indicating fibrosis in blue), from liver sections of FA, PM<sub>2.5</sub> or HFD-FA mice exposed for 14 weeks (magnification= 20X, n=4). (B) Total triglycerides and cholesterol measured in plasma (n=10 per group) and liver tissue of FA, PM<sub>2.5</sub> and HFD mice (n=4), (C) Hepatic mRNA gene expression levels of M1 (TNF- $\alpha$ , IL-6, IL1 $\beta$ , TLR4, IFN- $\gamma$  and adiponectin) and (D) M2 genes (IL-4, IL-10, and TGF- $\beta$ ). (E) Hepatic glycogen levels in all 3 study groups. Values shown as mean  $\pm$  SEM, \*p<0.05, unpaired t-test.



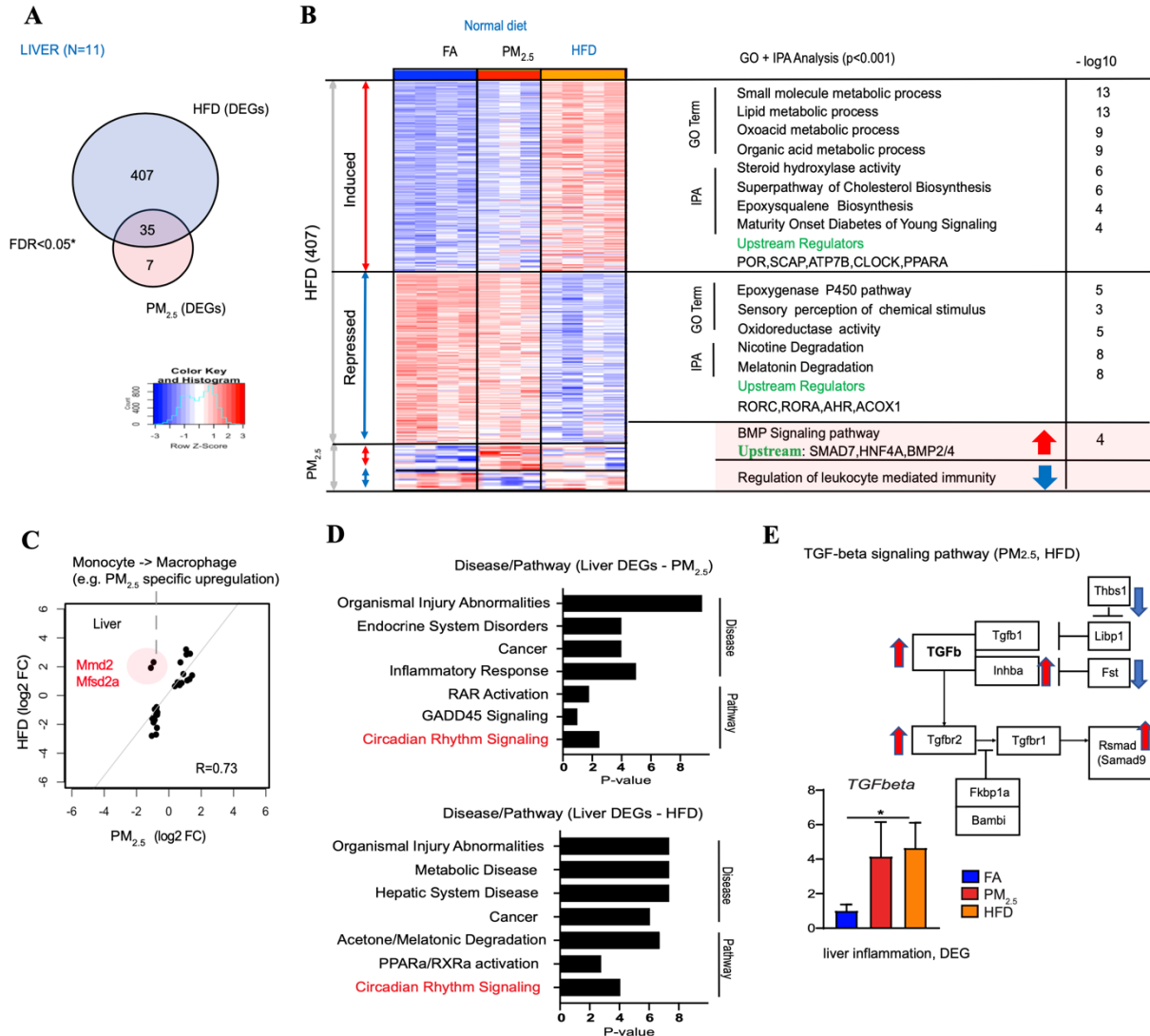
**Figure S4: Principal component analysis (PCA) plots of central and peripheral tissues samples (N=50).** (A) PCA plots of liver(n=11), muscle (n=9), BAT (n=9), WAT (n=10) and hypothalamus tissue (n=11) under all three experimental conditions Individual PCA plots depicting the variation in principle components across the different groups in (B) Hypothalamus, (C) Liver, (D) Skeletal muscle, (E) White adipose tissue, and (F) Brown adipose tissue. All PCA plots were plotted using FA, PM<sub>2.5</sub>, and HFD transcriptome datasets. (G) Venn diagrams depicting common and unique genes from DEGs of FA, PM<sub>2.5</sub> and HFD-fed mice. Hypergeometric test was performed to find significance of overlapped DEGs.



**Figure S5: BAT transcriptome summary for PM<sub>2.5</sub> and HFD specific Differentially Expressed Genes (DEGs)** (A) Venn diagram of FA vs PM<sub>2.5</sub> DEGs and Chow vs HFD DEGs. (B) Heatmap of statistically significant differentially expressed genes, and functional annotation for HFD and PM<sub>2.5</sub> specific DEGs. IPA analysis performed to find DEG associated pathways grouped by HFD induced or repressed expression, and PM induced or repressed expression. (C) Gene Set Enrichment Analysis (GSEA) using up-regulated and down-regulated DEGs (FDR < 0.1) in PM<sub>2.5</sub>.



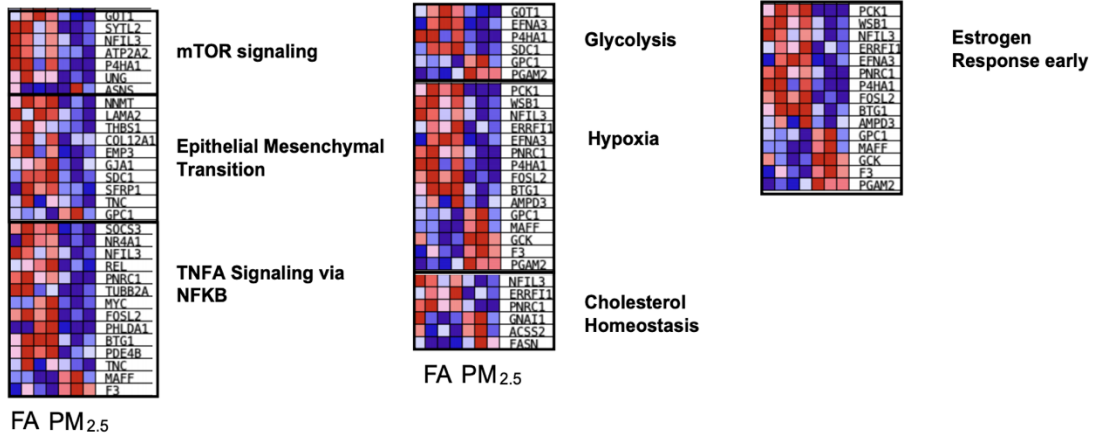
**Figure S6: White adipose tissue (WAT) transcriptome summary for PM<sub>2.5</sub> and HFD specific Differentially Expressed Genes (DEGs) using GO term and IPA.** (A) Venn diagram of PM<sub>2.5</sub> and HFD DEGs in WAT. 65 DEGs overlapped between the PM<sub>2.5</sub> and HFD conditions. (B) Heatmap of DEGs, and functional annotation for HFD and PM specific DEGs using IPA and GO term analysis. (C) Protein network visualization from the STRING database. (D) qPCR validation for high-fold change genes in WAT. The bar graph depicts the individual expression level quantified by qPCR.



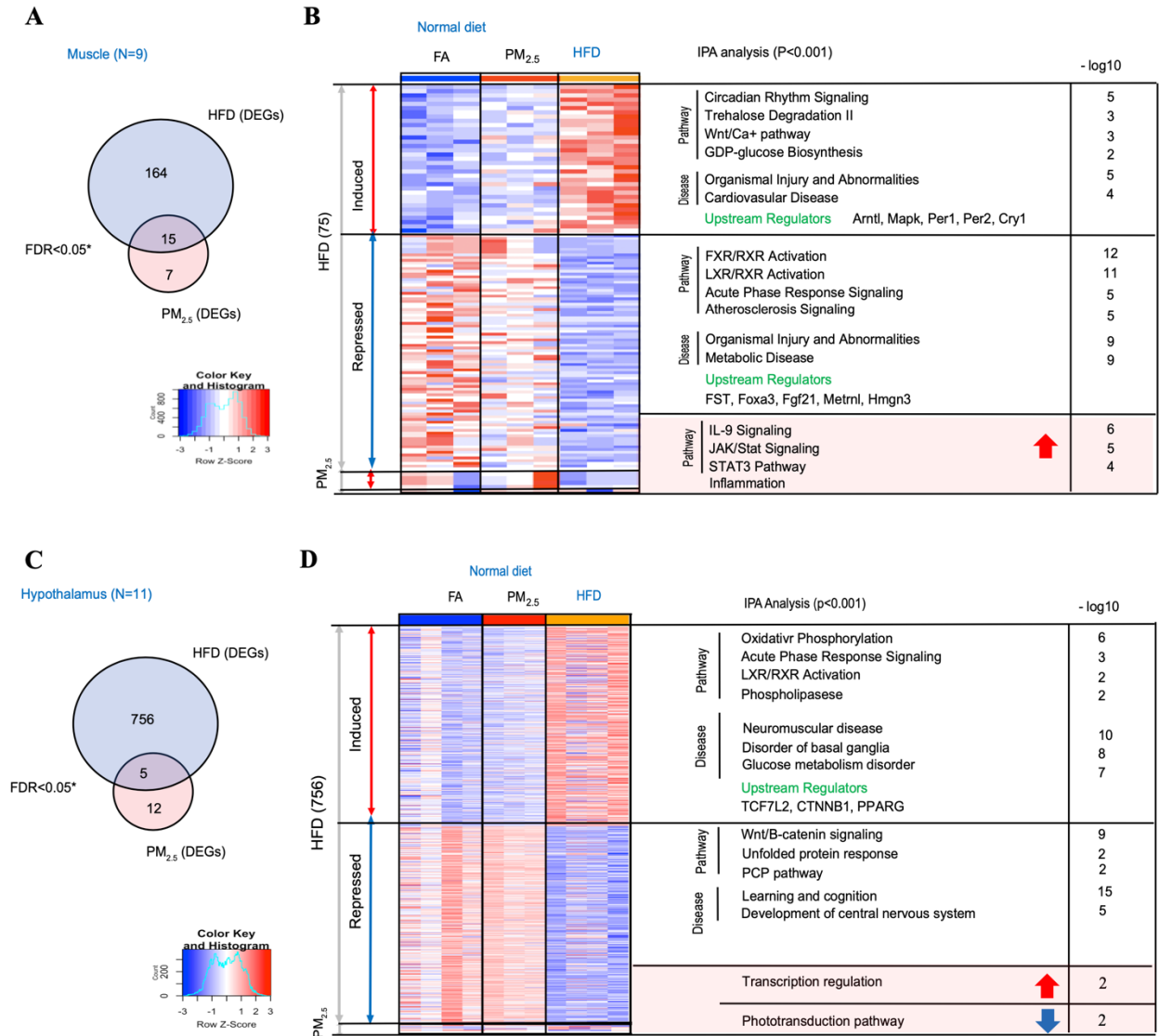
**Figure S7: Liver transcriptome summary for PM<sub>2.5</sub> and HFD specific Differentially Expressed Genes (DEGs)** (A)Venn Diagram of DEGs from the PM<sub>2.5</sub> and HFD liver samples with 35 DEGs common to both PM<sub>2.5</sub> and HFD. (B) Heatmap of statistically significant differentially expressed genes, and functional annotation for HFD and PM<sub>2.5</sub> specific DEGs using IPA and GO term analysis. IPA analysis performed to find DEG associated pathways using DEGs grouped by HFD induced or repressed expression, and PM<sub>2.5</sub> induced or repressed expression. (C) Scatterplot of correlations between overlapping PM<sub>2.5</sub> and HFD DEGs. (D) Disease and canonical pathways of liver DEGs. (E) TGF-beta signaling pathway was matched with DEG.



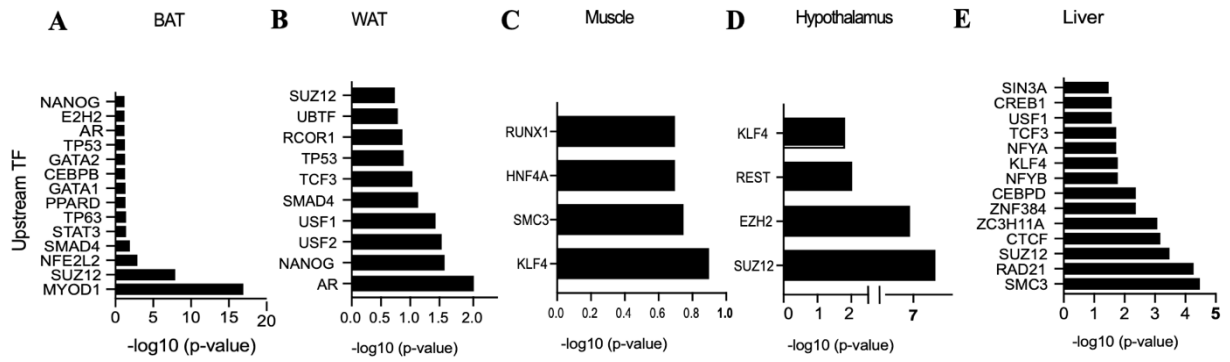
**A** GSEA analysis (FA vs PM<sub>2.5</sub>)



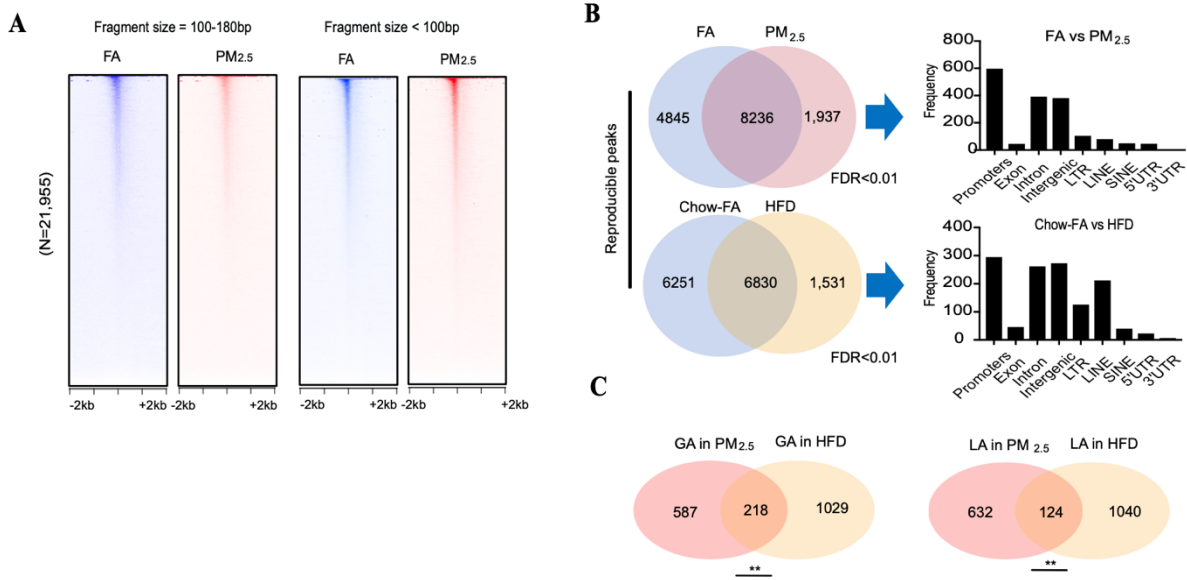
**Figure S8: GSEA analysis using differentially expressed genes (DEGs) in liver. (A)** Gene Set Enrichment Analysis (GSEA) using up-regulated and down-regulated DEGs (FDR<0.1) in PM<sub>2.5</sub>.



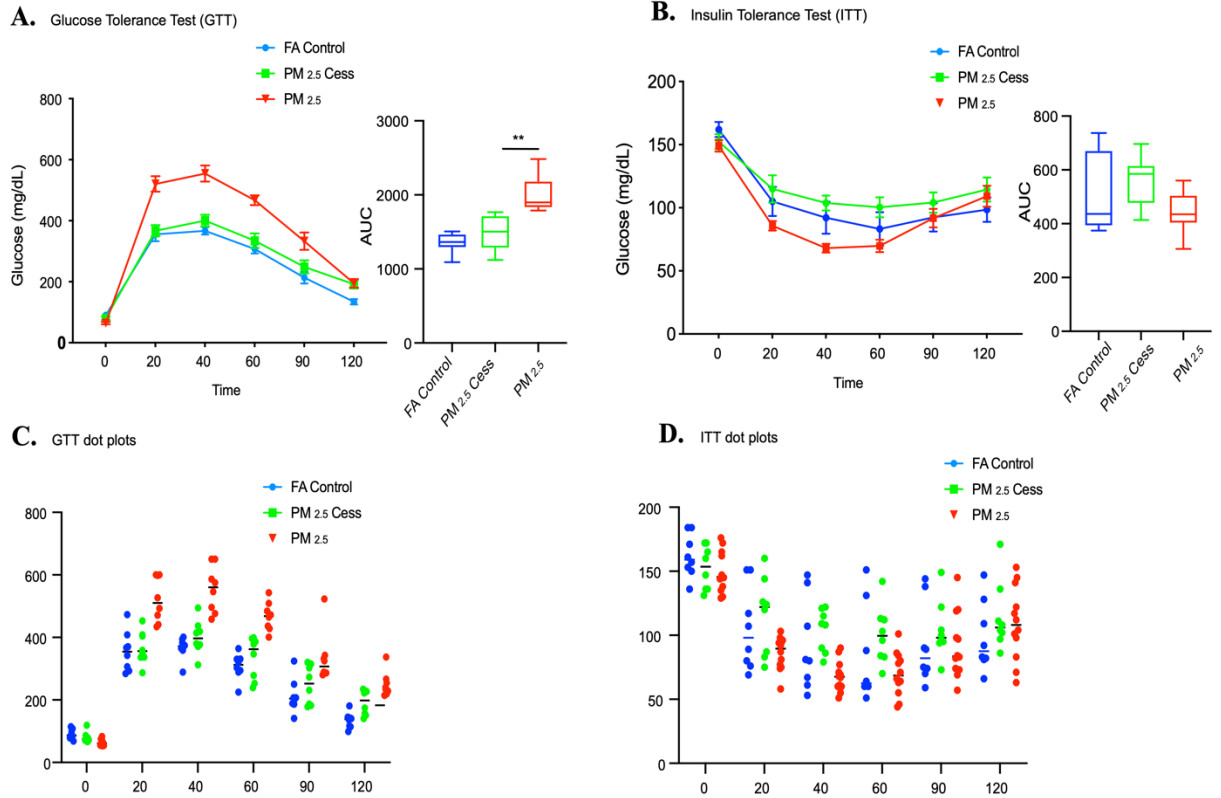
**Figure S9: Skeletal muscle and hypothalamus transcriptome summary for PM<sub>2.5</sub> and HFD specific Differentially Expressed Genes (DEGs).** (A) Venn diagram of DEGs from PM<sub>2.5</sub> and HFD muscle transcriptome datasets. 164 DEGs were HFD specific, 7 DEGs PM<sub>2.5</sub> specific while 15 DEGs overlapped. (B) Heatmap of statistically significant differentially expressed genes, and functional annotation for HFD and PM<sub>2.5</sub> specific DEGs using IPA and GO term analysis. (C) Venn diagram of DEGs from PM<sub>2.5</sub> and HFD muscle transcriptome datasets. 756 DEGs were HFD specific, 12 DEGs PM<sub>2.5</sub> specific while 5 DEGs overlapped. (D) Heatmap of statistically significant differentially expressed genes, and functional annotation for HFD and PM<sub>2.5</sub> specific DEGs using IPA and GO term analysis.



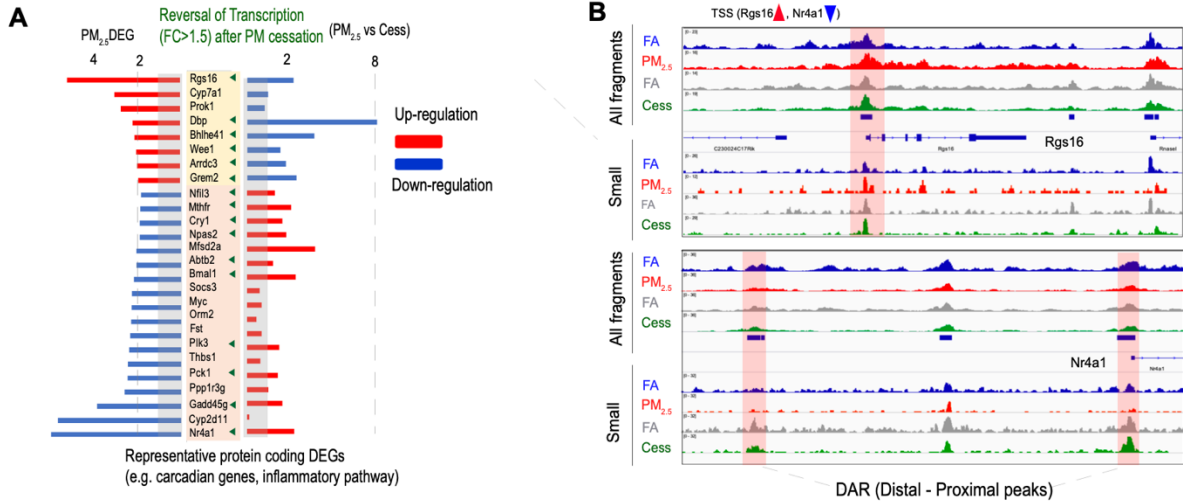
**Figure S10: Transcription Factor Enrichment Analysis (TFEA)** (A). Enriched up-stream TF in BAT sorted by p-value. (B) Enriched up-stream TF in WAT. (C) Enriched up-stream TF in skeletal muscle. (D) Enriched up-stream TF in hypothalamus. (E) Enriched up-stream TF in liver.



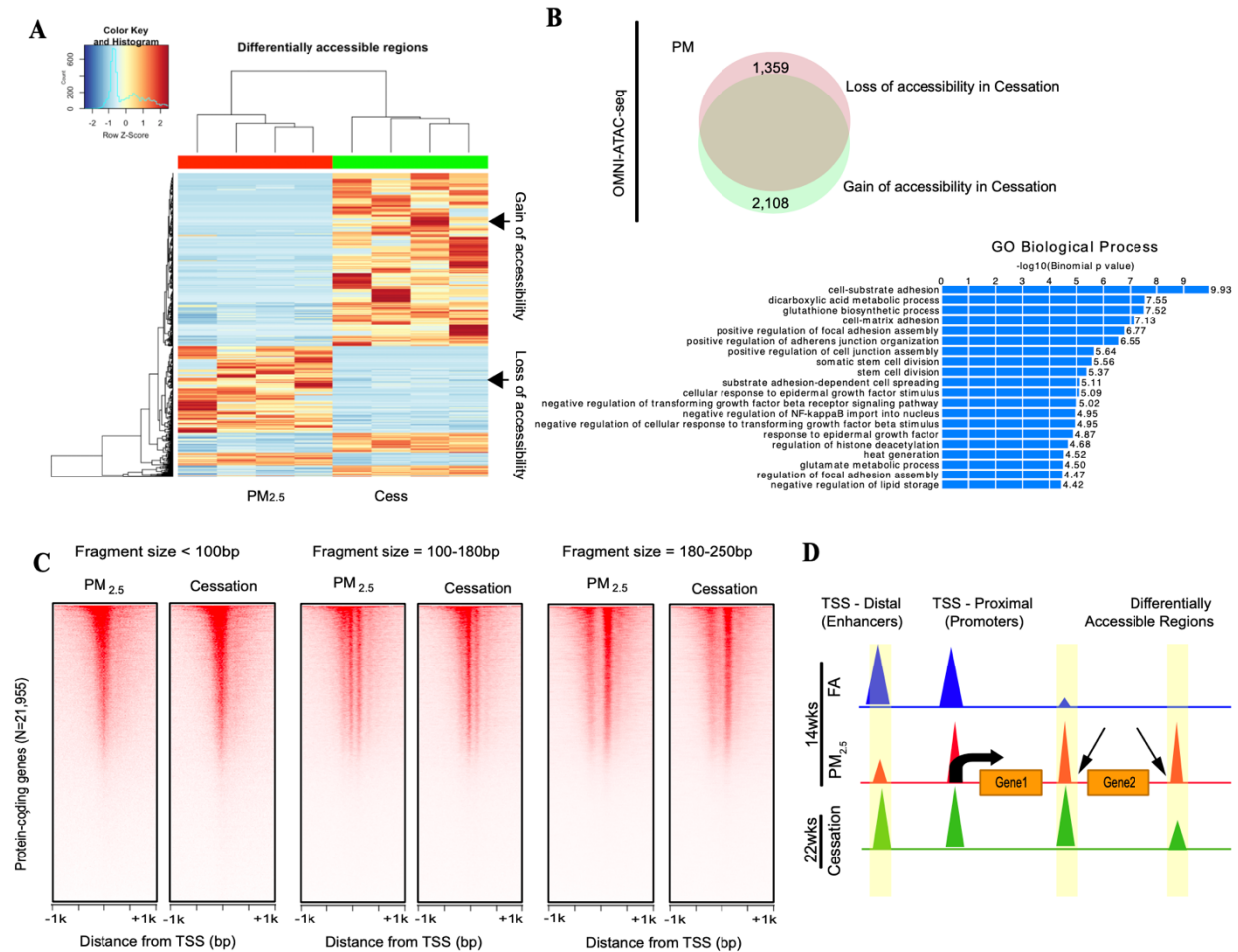
**Figure S11: Global chromatin accessibility changes according to fragment size distribution.** (A) Liver sample ATAC-seq differential fragment size analyses difference between FA (n=4, merged datasets) and PM<sub>2.5</sub> (n=3, merged datasets) liver ATAC datasets. Values < 100 bp fragment size represents nucleosome-free open chromatin, the 100-180bp as intermediate size, and the 180-250 bp fragments as mono-nucleosome size chromatin. (B) Venn diagram showing the PM<sub>2.5</sub> specific (1,937) and HFD specific (1,531) reproducible open chromatin from ATAC-seq liver tissue (n=12), and genomic distribution (C) The summary of the gain of accessibility (GA) and loss of accessibility (LA) in PM<sub>2.5</sub> and HFD. Hypergeometric test was performed to find significance of overlapped DARs.



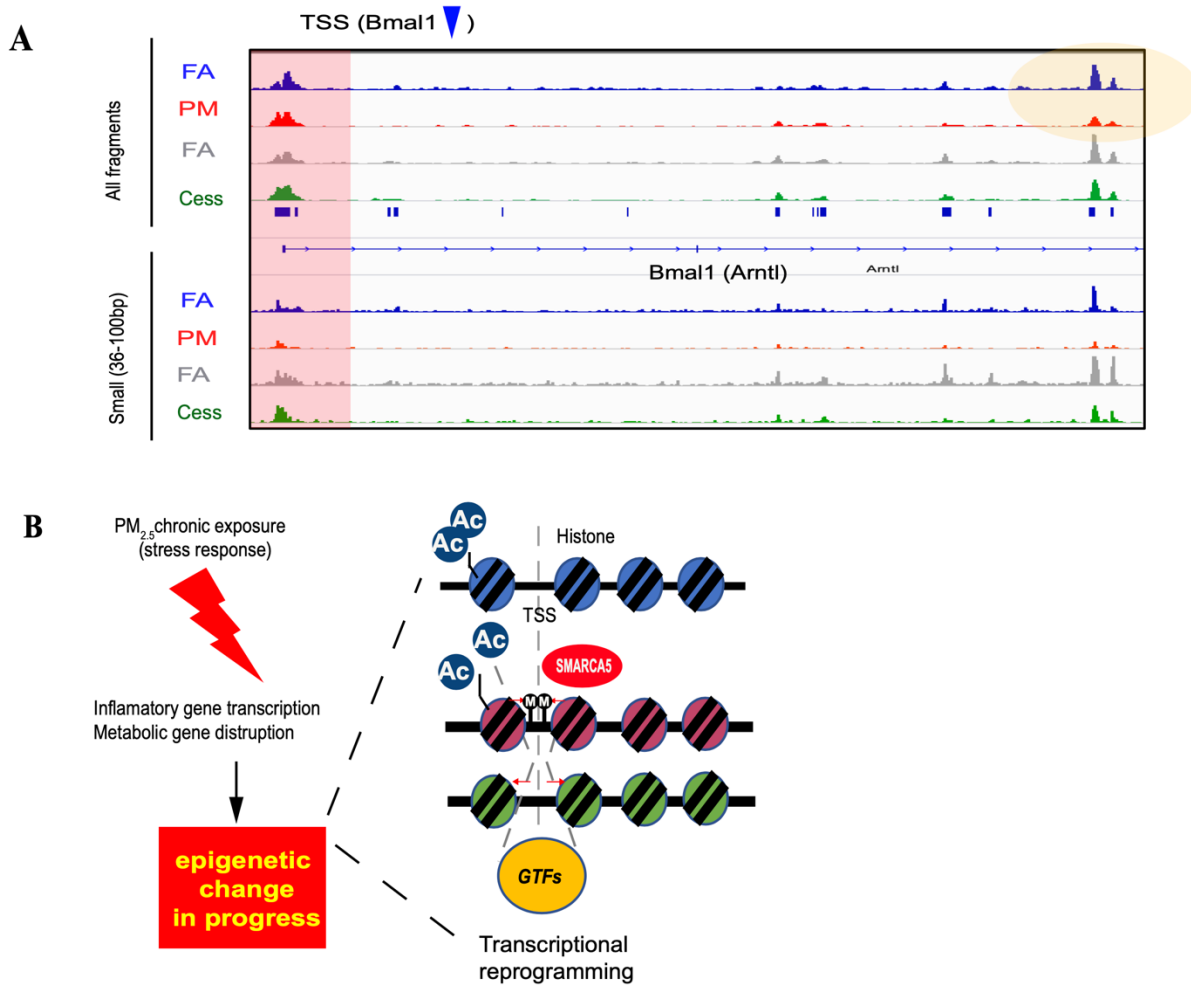
**Figure S12: Effect of PM<sub>2.5</sub> Cessation on glucose tolerance testing (GTT) and insulin tolerance testing (ITT).** (A) and (B) indicates GTT/ITT analysis from male mice exposed to FA control (n=8), PM<sub>2.5</sub> for 14wks (n=8) and 8 weeks of cessation (n=8). One-way ANOVA test performed using AUC values, and we observed significance phenotypic reversal in GTT. We tested ITT using additional four mice in PM<sub>2.5</sub> exposed mice (n=12). (C) and (D) depicts the dot plot of the individual variability in the GTT and ITT results.



**Figure S13. Impact of exposure cessation vs continued PM<sub>2.5</sub> exposure in liver.** (A) Reversal of liver DEGs upon 8-week PM<sub>2.5</sub> cessation (14-22wks). The length of the bars indicates the degree of gene expression change (B) Genome browser screenshots of Rgs16 and Nr4a1 from OMNI ATAC-seq datasets to visualize all fragments or small fragments under conditions of FA, PM<sub>2.5</sub> and PM<sub>2.5</sub> cessation.



**Figure S14: Chromatin accessibility and nucleosome positioning change in the cessation.** (A) Differentially accessible regions analysis and heatmap (B)  $PM_{2.5}$  specific (3,467) open chromatin from OMNI ATAC-seq in liver tissue ( $n=8$ ), and genomic distribution. GREAT cis-regulatory element analysis using differentially accessible regions (DARs). (C) OMNI-ATAC seq from liver of 14 wk.  $PM_{2.5}$  exposed ( $n=4$ ) and 8wk  $PM_{2.5}$  cessation mice ( $n=4$ ). (D) Schematic diagram representing mechanisms of epigenome changes in male mice subject to FA/ $PM_{2.5}$  or  $PM_{2.5}$  cessation.



**Figure S15: Genome browser shot with the circadian rhythm gene (*Arntl1/Bmal1*), and model of epigenetic mechanism (A) Potential enhancer site (gene body) that is >1kb distance from TSS was differentially accessible regions. (B) Schematic diagram that illustrates the potential epigenetic pathways (DNA methylation, Nucleosome positioning, and DNA accessibility)**



**Table S1: Summary of sequenced transcriptome samples for the study**

| Tissue types          | FA (M*) | PM <sub>2.5</sub> (M) | HFD (M) | Total (N) |
|-----------------------|---------|-----------------------|---------|-----------|
| Liver (14ws)          | 4       | 3                     | 4       | 11        |
| Liver (reversal,5mth) | 3       | 4                     | -       | 7         |
| White adipose         | 4       | 3                     | 3       | 10        |
| Brown adipose         | 3       | 3                     | 3       | 9         |
| Quad (muscle)         | 3       | 3                     | 3       | 9         |
| Hypothalamus          | 4       | 3                     | 4       | 11        |
| Total samples         | 21      | 19                    | 17      | 57        |

\*M: Male

**Table S2: hypergeometric analysis for Figure 2D**

| Tissue types        | Odds ratios | Expressed<br>genes in PM <sub>2.5</sub> | P-value   |
|---------------------|-------------|---|-----------|
| Liver (down)        | 55.83       | 12,491                                  | 2.2e-16   |
| Liver (up)          | 31.39       | 12,491                                  | 3.045e-09 |
| BAT (down)          | 10.79       | 13,656                                  | 2.2e-16   |
| BAT (up)            | 0.74        | 13,656                                  | 0.5262    |
| WAT (down)          | 12          | 16,674                                  | 4.32e-07  |
| WAT (up)            | 9.68        | 16,674                                  | 2.2e-16   |
| Muscle (down)       | 74.31       | 13,996                                  | 2.2e-16   |
| Muscle (up)         | 0           | 13,996                                  | 1         |
| Hypothalamus (down) | 6.42        | 13,566                                  | 0.1682    |
| Hypothalamus (up)   | 17.49       | 13,566                                  | 9.562e-07 |

## **Methodological Details**

### ***Air pollution exposure and phenotyping***

Male and female C57BL/6J mice (3 weeks) were purchased from Jackson Laboratory (n=64/group). At 3 weeks of age, mice were initially housed in groups and maintained at 21°C on a 12-hr light/12-hr dark cycle, to help acclimatize them to the new environment; they had free access to water and were fed either a regular chow or a high-fat diet with 60% of its calories derived from lipid (Research Diet, NJ, USA; D12492). Mice fed normal chow were exposed through inhalation to either filtered air (FA) or concentrated PM<sub>2.5</sub> (~10x ambient level/ ~60-120ug/m<sup>3</sup>) for 6 hours/day, 5 days/week for 14 weeks. A group of high fat fed animals were exposed to filtered air alone to serve as high-caloric diet controls. Weights were documented weekly, along with assessment of measures of glucose homeostasis and insulin responses. Inhalation exposure was carried out in a Versatile Aerosol Concentrator and Exposure System (VACES) air pollution exposure facility at Case Western Reserve University Animal Facility per IACUC protocol (2016-0319). The design of VACES has been described previously and provides stable concentrations of PM<sub>2.5</sub> which is roughly 10x of the ambient exposure (1, 2). All animal procedures and experiments were approved by the IACUC committee before they were undertaken.

***Exposure Cessation Experiments:*** For these experiments we utilized a set of mice that had undergone exposures to FA or PM<sub>2.5</sub> and fed normal chow diet for 14 weeks. Mice fed high fat diet exposed to FA, served as controls for a contrasting environmental exposure over the same 14-week duration. At the end of 14 weeks of exposure, 8-12 mice in the PM<sub>2.5</sub> group were switched to FA. The mice in the 3 groups were sacrificed at the end of 8 weeks and various tissues (liver, BAT, WAT, skeletal muscle) were isolated.

### ***Evaluation of insulin resistance and Glucose Homeostasis***

Body weight and food intake was monitored on a weekly basis. Glucose tolerance test (GTT): GTT was performed at weeks 2, 9, 13, and 22, during the course of the experimental period. After 13 weeks of PM<sub>2.5</sub> exposure, fasting glucose and post bolus blood glucose were measured. Briefly, mice were fasted overnight with access to water following which their blood glucose levels were evaluated (time=0). A sterile glucose solution was injected intraperitoneally at 2 g/kg of body mass. Glucose was measured in tail blood at times: 15, 30, 60, 90, and 120 minutes post glucose injection. Blood glucose levels were determined using the handheld Bayer Contour glucometer, and both kinetic curve and area under the curve (AUC) were generated to evaluate the glucose levels in different treatment groups. Insulin tolerance tests (ITT) were performed following a 6 hour fast. Insulin was injected at 1 U/kg, followed by an assessment of blood glucose on the same timescale as GTT.

### ***Whole Body Metabolism and FDG-PET for Brown Adipose Tissue Glucose Uptake***

Energy expenditure and RER levels were measured using the Columbus Instruments Laboratory Animal Monitoring System (CLAMS) (Columbus, OH, USA) over 48 hours. 18F-Fluorodeoxyglucose (FDG) was synthesized by nucleophilic substitution method using an FDG synthesizing instrument. Positron emission tomography (PET) was performed using an advance scanner (Philip Mosaic PET Scanner). Eight hours fasted mice were restrained and injected with 0.9% saline (control) or insulin (0.75 units/kg) diluted in 0.9% saline for 5 min, and then received intravenous administration of FDG (200-300 µCi/mice). Mice underwent small-animal PET and microcomputed tomography, and whole-body PET images were acquired 30 min later using an acquisition time of 15-30 min. FDG uptake quantification was performed by region-of-interest analysis using Carimas II Research Workplace software.(3, 4) The PET analysis was performed in a blinded manner by an independent observer in a core laboratory without knowledge of the group assignments.

### ***Tissue collection and DNA/RNA Extraction***

Mice were euthanized after 14 weeks of PM<sub>2.5</sub> exposure with CO<sub>2</sub> gas at room temperature. Tissues were snap-frozen in liquid nitrogen and moved to -80°C until further processing. The left liver lobe was collected to assess liver transcriptome, while remaining liver sections were utilized for other assays including multi-omics analysis. DNA/RNA was extracted from tissues using a QIAGEN All Prep kit after pulverizing the tissue using a Freezer Mill (6775 Freezer/Mill® Cryogenic Grinder). RNA quantity and quality were checked using NanoDrop (Thermo Fisher) and BioAnalyzer (Agilent Santa Clara, CA USA), respectively. Samples that had a RIN (RNA Integrity Number) value of 6.8 or higher were screened for this study. Additionally, quadriceps muscle, subscapular brown adipose tissue (BAT) and epididymal adipose (white adipose tissue, WAT) were collected to evaluate relevant tissue-specific genes expression.

## **RNA-Sequencing**

The extracted RNA was quantified by NanoDrop (Thermo Fisher Scientific, MA) and the quality was assessed using a 2100 Bioanalyzer (Agilent Technologies, CA). The Agilent Bioanalyzer is a microfluidics platform used for sizing, quantification, and quality control for RNA (and DNA/ proteins) and provides an “RNA Integrity Number” (RIN), which quantifies the fragmentation of the RNA sample. The RNA samples were selected for sequencing if RIN value was more than 6.5. On average, 500-1,000mg of RNA samples were used for library preparation and double-stranded cDNA generation. We utilized the NEBNext® Ultra RNA Library Prep Kit (New England BioLabs, Inc, Ipswich, MA) for liver tissues and a TrueSeq RNA Library Prep Kit (Illumina, San Diego, CA) to generate strand-specific libraries for all other mentioned tissues. The library was amplified by 15-cycle according to the manufacturer protocol. After PCR primers removal with Agencourt AMPure PCR purification kit (Beckman Coulter, CA), the sequencing library was quantified on TapeStation (TapeStation Instrument). The prepared library was sequenced by HiSeq series sequencer including HiSeq4000, HiSeqX (Illumina, San Diego, CA). The raw BCL files were converted into FastQ files using CASAVA 1.8.2 (CASAVA).

Transcriptome datasets were derived from tissues and treatments as summarized in Supplementary Table 1. We also included the transcriptome profiles from the liver after 8 weeks of exposure cessation, A total of 5.8 billion reads were used, on an average of 50 million reads per sample. For transcriptome analysis we used STAR Gencode M13 reference features. Prior to sequence alignment, we used trim galore (version 0.4.3) with cutadapt package (version 1.12) (5) for sequence trimming and to improve data quality. We then mapped sequencing reads to the mouse reference genome (mm10) using STAR aligner (6), and calculated the raw count using FeatureCounts package (gene-level) (4). We performed Principle Component Analysis (PCA) for testing biological reproducibility within replicates and to identify the variability between treatments. All datasets are deposited under GEO accession number are accessible to the scientific community. All protocols and analyses were performed by the Johns Hopkins University and University of Michigan sequencing core according to the standard operation procedures.

### ***Differentially Expressed Genes (DEGs) and functional annotations***

For transcriptome analysis, we assayed for protein-coding gene expression (Gencode M13- Freeze date Oct 2016). We used Rsubread and feature Counts to generate a gene-by-sample matrix of reading counts that was analyzed using edgeR after lane normalization and removing unwanted variation (RUVg) (7). The output of this analysis is a set of differentially expressed genes that are distinguished between experimental groups. We applied the limma based edgeR method to determine differentially expressed transcripts; cutoff:  $\log_2FC > 0.8$ , CPM $>2$ , FDR  $< 0.05$ . A low threshold in fold change (*low-fold change DEGs*:  $\log_2FC$  between 0.8 and 1.0) was prioritized due to the potentially small effect size of environmental exposure on epigenome or transcription. We also applied a higher fold change cutoff for stringent analysis ( $\log_2FC > 1.0$ ) was also applied for GO and Pathway analyses using DEGs and control genes. For gene set enrichment analysis, we also utilized the relaxed cutoff (FDR $<0.1$ ) to include more candidate DEGs with a minimum of five genes in the enriched gene sets. R package TopGO (8) was utilized to process GO term enrichment analysis. In addition, IPA (Ingenuity Systems Inc., Redwood City, CA) was conducted to identify enriched pathways from DEGs and upstream regulators (TF, enzyme, receptors). To validate our findings Transfac, and HOMER analysis (9) were performed to confirm upstream regulators from the DEGs list.

### ***Bioinformatics analysis for RNA sequencing data***

Raw sequencing reads were processed using TaRGETII RNA-seq pipeline which was prepared large size RNA-seq sample processing ([https://github.com/Zhang-lab/RNA-seq\\_QC\\_analysis](https://github.com/Zhang-lab/RNA-seq_QC_analysis)). Prior to sequence alignment, we used trim galore (version 0.4.3) with cutadapt package (version 1.12) (5) for sequence trimming and to improve data quality. We then mapped sequencing reads to the mouse reference genome (mm10) using STAR aligner (6), and calculated the raw count using FeatureCounts package (gene-level). The RNA-seq read counts for genes were normalized and log-transformed into using limma, edgeR packages. We performed Principle Component Analysis (PCA) for testing biological reproducibility within replicates and to identify the variability between treatments. Heatmap and volcano plots were generated by R packages (gplots, RColorBrewer). The differential gene expression analysis between FA and PM (or

HFD) was performed by limma package after correction of potential batch effect using RUVg (7). We applied the limma based edgeR method to determine differentially expressed transcripts; cutoff:  $\log_2FC > 0.8$ ,  $CPM > 2$ , and  $FDR < 0.05$ . For gene set enrichment analysis (GSEA), we also utilized the relaxed cutoff ( $FDR < 0.1$ ) to include more candidate DEGs with a minimum of five genes in the enriched gene sets.

### ***Open chromatin signatures from liver samples***

10mg of frozen liver powder was used to prepare ATAC-seq library. ATAC-seq protocol has been applied (10), preparing nuclei prep and applying transposase reaction. The library was purified using 1X AMPure beads (100-500bp fragment size). We utilized Hi-Seq system to sequence the libraries. In addition, we also performed OMNI ATAC-seq using frozen liver tissue from all groups including PM<sub>2.5</sub> cessation group (11), in order to increase the signal-to-background ratio and the number of enriched peaks. Briefly, all reads were trimmed using cutadapt package (5), and trimmed read (>36bp minimum alignment length) were mapped against mm10 genome using BWA aligner (12). We used de-duplicated and uniquely mapped reads for peak calling analysis after excluding black-list regions defined by ENCODE. The candidate peaks were predicted by MACS peak calling software (13). We used predicted open chromatin peaks where at least two biological replicates were reproducible for the downstream motif binding analysis using IDR cutoff 0.05 (14). In addition, we also applied the limma based edgeR method to determine differentially accessible regions (DARs); cutoff:  $\log_2FC > 1.0$ ,  $CPM > 2$ ,  $p\text{-value} < 0.01$ . TaRGETII RNAseq, and ATAC-seq pipelines are freely available via git hub and docker image developed by Bo Zhang lab at WASHU (<https://github.com/Zhang-lab/TaRGET-II-ATACseq-pipeline>) as a part of TaRGETII consortium Bioinformatics Working Group. Finally, the candidate open chromatin regions were submitted to search potential transcription factor binding sites using HOMER (9) software (mm10 -size 150 -len 11). We performed differential motif binding analysis (15) with available ChIP-seq datasets (16) and GREAT analysis.

### ***Statistical analysis***

For all phenotypic analyses, a  $p\text{-value} < 0.05$  was considered significant (\*). For statistics for glucose tolerance and insulin tolerance test, we performed a two-way ANOVA test using Prism software version 8 (GraphPad Software, San Diego CA). False discovery rate (FDR) correction was used to account for Type I error using adjusted BH method. For differential gene analysis, we used the negative binomial test with a cutoff of the false discovery rate ( $FDR < 0.05$ ). For batch effect correction, we used 5,000 empirically selected not significantly changed genes ( $FDR < 0.05$ ) as a negative control gene set. We used the same method for differential binding analysis from ATAC-seq data, however, we additionally performed IDR ( $FDR < 0.05$ ) to check the reproducibility of enriched ATAC peaks. We used predicted open chromatin peaks where at least two biological replicates were reproducible for the downstream motif binding analysis. To determine differentially accessible regions, we used the negative binomial test with edgeR (cutoff:  $\log_2FC > 1.0$ ,  $CPM > 2$ ,  $p\text{-value} < 0.01$ ). The hypergeometric analysis was done by using fisher.test using R using 2x2 matrix of overlapped genes, unique to group HFD-FA, unique to group PM<sub>2.5</sub>, and total expressed genes in PM<sub>2.5</sub>.

## REFERENCES

1. Maciejczyk P, and Chen LC. Effects of subchronic exposures to concentrated ambient particles (CAPs) in mice. VIII. Source-related daily variations in in vitro responses to CAPs. *Inhal Toxicol*. 2005;17:243-53.
2. Sun Q, Yue P, Deiluiis JA, Lumeng CN, Kampfrath T, Mikolaj MB, et al. Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation*. 2009;119(4):538-46.
3. Hulsmans M, Sager HB, Roh JD, Valero-Munoz M, Houstis NE, Iwamoto Y, et al. Cardiac macrophages promote diastolic dysfunction. *The Journal of experimental medicine*. 2018;215(2):423-40.
4. Olsen JM, Csikasz RI, Dehvari N, Lu L, Sandstrom A, Oberg AI, et al. beta3-Adrenergically induced glucose uptake in brown adipose tissue is independent of UCP1 presence or activity: Mediation through the mTOR pathway. *Mol Metab*. 2017;6(6):611-9.
5. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnetjournal*. 2011;17(1).
6. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
7. Risso D, Ngai J, Speed TP, and Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol*. 2014;32(9):896-902.
8. J AAaR. topGO: Enrichment Analysis for Gene Ontology. <http://bioconductor.org/packages/release/bioc/html/topGO.html>.
9. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell*. 2010;38(4):576-89.
10. Buenrostro JD, Wu B, Chang HY, and Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*. 2015;109:21 9 1-9.
11. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*. 2017;14(10):959-62.
12. Li H, and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-60.
13. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
14. Q Li JB, H Huang, PJ Bickel. Measuring reproducibility of high-throughput experiments. *Ann Appl Stat*. 2011;5:1752-79.
15. Tripodi IJ, Allen MA, and Dowell RD. Detecting Differential Transcription Factor Activity from ATAC-Seq Data. *Molecules*. 2018;23(5).
16. Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res*. 2018;46(D1):D252-D9.