SUPPLEMENTARY MATERIAL

Toxicity of the main electronic cigarette components, propylene glycol, glycerin, and nicotine, in Sprague-Dawley rats in a 90-day OECD inhalation study complemented by molecular endpoints

Blaine Phillips*[,a], Bjoern Titz*[,b], Ulrike Kogel*[,b], Danilal Sharma*[,a], Patrice Leroy[b], Yang Xiang[b], Grégory Vuillaume[b], Stefan Lebrun[b], Davide Sciuscio[b], Jenny Ho[a], Catherine Nury[b], Emmanuel Guedj[b], Ashraf Elamin[b], Marco Esposito[b], Subash Krishnan[b], Walter Schlage[c], Emilija Veljkovic[a], Nikolai V. Ivanov[b], Florian Martin[b], Manuel C. Peitsch[b], Julia Hoeng[b], Patrick Vanscheeuwijck[#,b]

[a] Philip Morris International Research Laboratories, 50 Science Park Road, Singapore, Singapore (part of Philip Morris International group of companies)

[b] Philip Morris International Research and Development, Philip Morris Products S.A., Quai Jeanrenaud 5, 2000 Neuchatel, Switzerland (part of Philip Morris International group of companies)

[c] Biology consultant, Max-Baermann-Str. 21, 51429 Bergisch Gladbach, Germany

* B.P., B.T., U.K., and D.S. contributed equally to this manuscript.

[#]**Corresponding author:**

Patrick Vanscheeuwijck, PhD

E-mail: Patrick.vanscheeuwijck@pmi.com

Tel: +41 (58) 242 2511

Fax: +41 (58) 242 2811

Supplementary Material and Methods

Determination of PG and VG concentrations

For the determination of PG and VG concentrations in the diluted test atmosphere, samples were collected using a Cambridge filter. Sample collection was carried out for 30 min at a sampling flow rate of 1.5 L/min for Low PG/VG chambers (45 L total volume sampled) and 0.7 L/min for Medium and High PG/VG chambers (21 L total volume sampled). After sample collection, the filters were transferred to glass tubes and processed by adding 25 mL of extraction solution (0.3 mg/mL of N-heptadecane in 2-propanol). The filters in the extraction solution were vortexed for 2 min and then shaken for 30 min at 400 rpm. The mixture was centrifuged for 5 min at 450 g, and 1 µL of the extract was analyzed using an Agilent Technologies gas chromatography (GC) system fitted with a flame ionization detector and DB-WAXetr GC column with 0.5-µm film thickness, 30 m × 0.25 mm I.D. (J&W Scientific Inc., Folsom, CA, USA). The GC temperature program was as follows: After equilibration at 110°C, the temperature was held for 1 min, increased by 10°C/min to 150°C, held for 3 min, increased by 80°C/min to 240°C, then held again for another 6 min. All quantification was based on a series of calibration standards containing N-heptadecane as the internal standard, along with USP testing specification-grade PG (≥99.5%, Sigma-Aldrich, St. Louis, MO, USA) at 0.16–1.6 mg/mL and VG (≥99%, Sigma-Aldrich) at 0.2–2 mg/mL. All quantitated extracts were within this range and diluted if necessary.

Transcriptomics analysis

RNA from the RNE, left lung and liver samples was isolated using a miRNeasy extraction kit (Qiagen, Hilden, Germany), with one modification: RLT buffer (Qiagen) with beta-mercaptoethanol was used to lyse the lung slices. RNA was processed for hybridization on GeneChip® Rat Genome 230 2.0 arrays (Affymetrix, Santa Clara, CA, USA) using a High-Throughput 3' *In Vitro* Transcription PLUS kit (Affymetrix). All samples were required to have a RIN value above 6.

Raw data files were processed in the custom Chip Description File environment Rat2302_Rn_ENTREZG v19.0.0 (rat2302rnentrezgcdf) (Dai et al., 2005), and normalized using frozen robust microarray analysis (fRMA) (McCall et al., 2010). The normalization vector needed for fRMA was created using a set of 1,023 microarray rat samples from 11 tissues and the R package frmaTools version 1.18.0. Quality controls, including log-intensities, normalized unscaled standard error (NUSE), relative log expression (RLE), median absolute value RLE (MARLE), and pseudo-images as well as raw image plots, were performed with the affyPLM package (Bioconductor suite) (Bolstad et al., 2013). The following quality control criteria had to be met by each CEL file before undergoing downstream analysis:

a) Pseudo-image does not display a spatial pattern covering approximately 10% of the image.

b) Median NUSE < 1.05

c) $|\text{Median RLE}| < 0.1$

d) $|(\text{MARLE}-\text{median}(\text{MARLE}))|/\ (1.4826 * \text{mad}(\text{MARLE})) < 1/\sqrt{0.01}$; (where mad is the median absolute deviation)

Following the quality control procedures, raw *p*-values were generated for contrasts between vehicle and exposed groups with the *limma* package (Smyth, 2004), and adjusted using the Benjamini-Hochberg false discovery rate (FDR) multiple test correction (Gentleman et al., 2004). In addition, the overall nicotine effect was statistically assessed as the contrast between the average Nicotine + (PG/VG) vs. vehicle and the average (PG/VG) vs. vehicle effect.

Proteomics analysis

Proteome alterations were assessed by isobaric tag-based quantification using the iTRAQ® approach (Titz et al., 2015a). For this, rat right RNE samples were homogenized and sonicated in tissue lysis buffer (BioRad, Hercules, CA, USA) in random order and the proteins were precipitated with acetone. Frozen rat

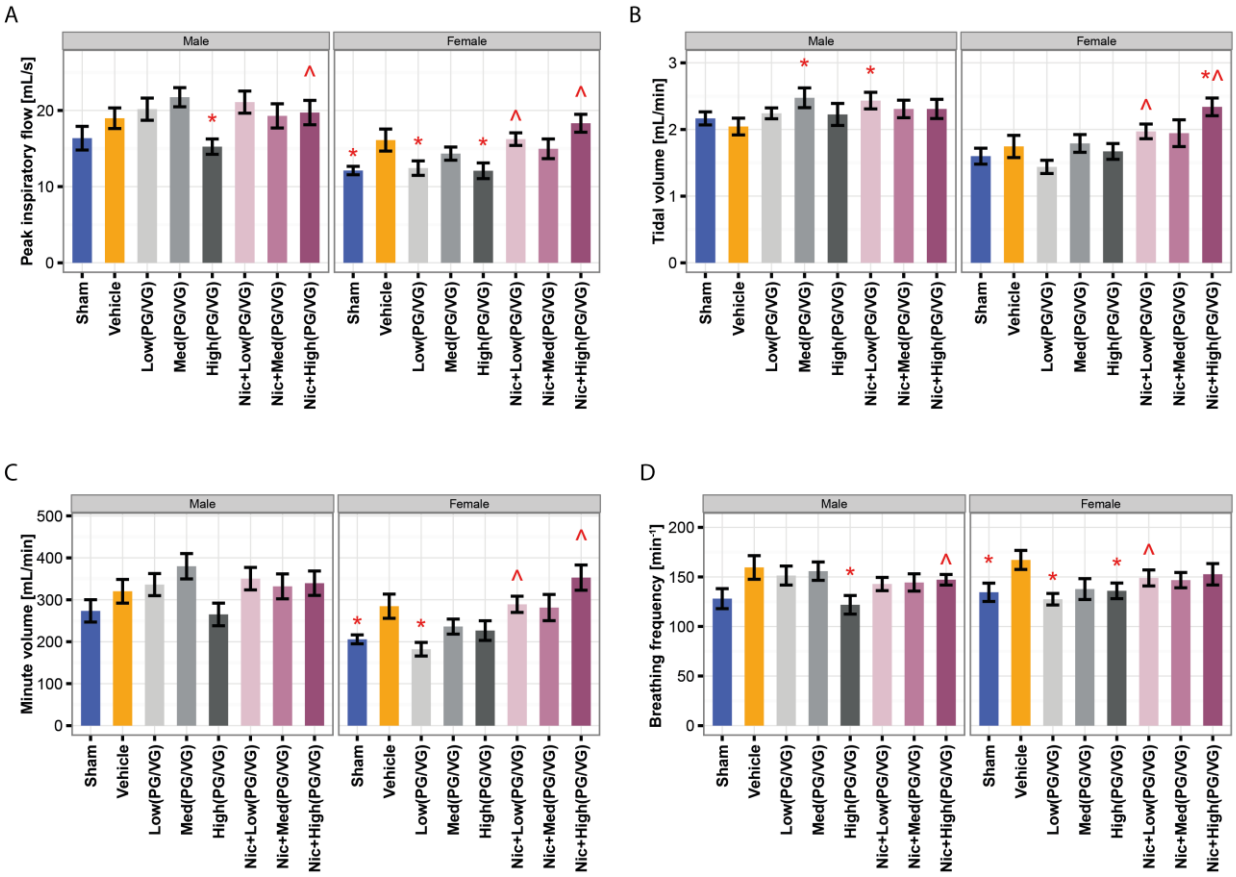
lung tissue slices were homogenized using a bead-assisted procedure in a Tissue Lyser II (Qiagen) in tissue lysis buffer before acetone precipitation. Liver tissue samples were homogenized similarly in tissue lysis buffer (BioRad Laboratories, Hercules, CA, USA) in random order. Protein precipitates were resuspended in 0.5 M triethylammonium bicarbonate (TEAB, Sigma-Aldrich), 1 M urea (Sigma-Aldrich), and 0.1% SDS (Sigma-Aldrich). Next, 50 µg of the suspension were processed using the iTRAQ 8-plex labeling procedure according to the manufacturer's instructions (AB Sciex, Framingham, MA, USA). Trypsin-Lys C mix (Promega, Madison, WI, USA) was added to the samples in a 1:10 trypsin to protein ratio (w/w), followed by overnight digestion at 37°C. The trypsin-digested samples were labeled with reporter-ion tags for the different exposure groups. For this, a single multiplexed iTRAQ set layout was defined that accommodated one sample of each exposure type: Each 8-plex labeling set included one sample replicate of each sample type (Sham, Vehicle, Low (PG/VG), Med (PG/VG), High (PG/VG), Nicotine + Low (PG/VG), Nicotine + Med (PG/VG), and Nicotine + High (PG/VG)). For each of the six replicate sets, one sample of each type was randomly selected. Within each replicate labeling set, the sample for channel mapping was randomized.

All labeled samples that belonged to one iTRAQ replicate set were pooled and dried in a SpeedVac concentrator (RVC 2-25 CD Plus, Martin Christ, Osterode am Harz, Germany). Samples were desalted using 0.5-mL bed detergent removal columns (Pierce, Rockford, IL, USA) and then using 1-cc C18 reversed-phase Sep Pak columns (Waters, Milford, MA, USA) according to the manufacturers' manuals. Samples were dried in a SpeedVac evaporator and resuspended in nanoLC buffer A (5% acetonitrile and 0.2% formic acid, Sigma-Aldrich). Samples were analyzed in random order using an Easy nanoLC 1000 instrument (Thermo Fisher Scientific, Waltham, MA, USA) connected online to a Q Exactive mass-analyzer (Thermo Fisher Scientific). Peptides were separated on a 50-cm Acclaim™ PepMap™ 100 C18 LC column (2-µm particle size; Thermo Fisher Scientific) at a flow rate of 200 nL/min, with a 200-min gradient from nanoLC buffer A to 40% acetonitrile with 0.2% formic acid. Each sample was injected twice with two different analytical methods, one fast and one sensitive method, as described previously (Kelstrup et al., 2012), on the same

column. The outputs of both mass spectrometry runs were combined as merged mass-lists and searched against the rat reference proteome set (UniProt, version October 2013, canonical isoforms only) using Proteome Discoverer version 1.4.0.288 software (Thermo Fisher Scientific). SequestHT implemented in Proteome Discoverer was used as the search tool, and iTRAQ reporter-ion intensities were determined from Proteome Discoverer. The Percolator node of Proteome Discoverer was used to estimate peptide-level adjusted *p*-values (q-values). iTRAQ peptide-level quantification data were exported and further processed in the R statistical environment (R Development Core Team, 2007). The quantification data were filtered for a q-value < 0.01 and for "unique" quantification results, as defined in Proteome Discoverer. A global variance-stabilizing normalization was performed with the corresponding Bioconductor package in R (Huber et al., 2002; Hultin-Rosenberg et al., 2013). Each iTRAQ reporter-ion set was normalized to its median, and protein expression values were calculated as the median of these normalized peptide-level quantification values (Herbrich et al., 2013). To detect differentially abundant proteins, a linear model was fitted for each exposure condition and the vehicle group, and *p*-values were calculated from moderated *t*-statistics with the empirical Bayes approach (Gentleman et al., 2004). In addition, the overall nicotine effect was assessed statistically as the contrast between the average Nicotine + (PG/VG) vs. vehicle and the average (PG/VG) vs. vehicle effect. The FDR method was then used to correct for multiple testing effects. Proteins with an adjusted *p*-value < 0.05 were considered differentially abundant.

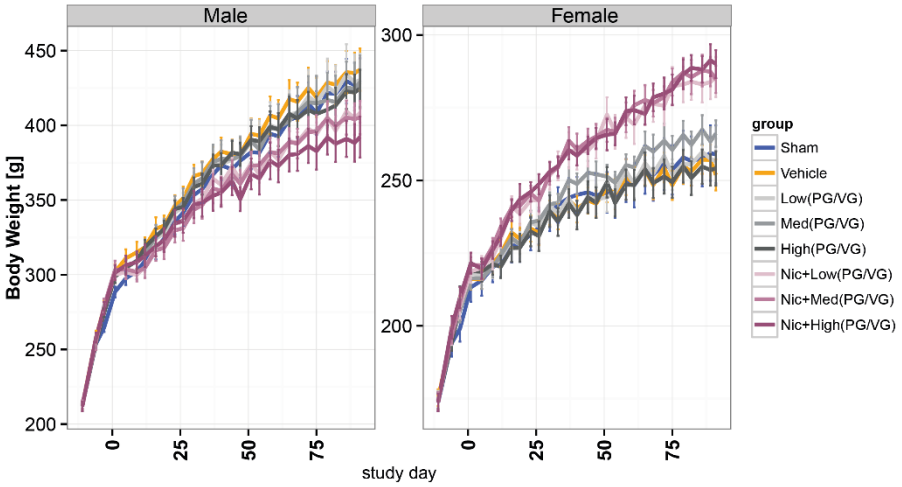
Supplementary Figures

Supplementary Figure 1. Respiratory physiology parameters. (A) Peak inspiratory flow. (B) Tidal volume. (C) Minute volume. (D) Breathing frequency. Statistically significant differences from the vehicle group are represented by asterisks (*); statistically significant differences between groups with and without nicotine (at the same PG/VG concentration) are represented by carets (^) ($p < 0.05$).

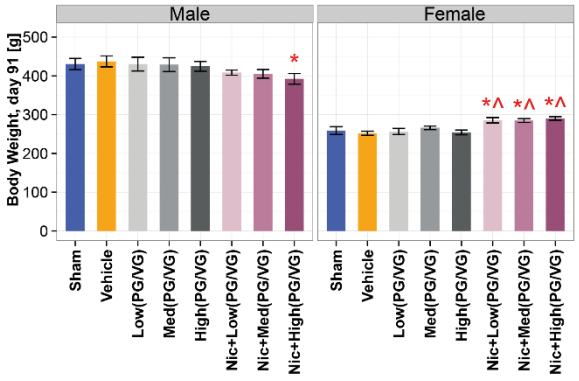


Supplementary Figure 2. Body weight and food consumption. (A) Body weight development over course of the study. (B) Body weights of male and female rats on day 91 of the study, at the end of the study period (mean ± SEM, N = 10). (C) Mean food consumption of male and female rats. Values were calculated as g/(day × 100 g body weight) (mean ± SEM) and were measured weekly; determinations were performed per group. PG, propylene glycol; VG, vegetable glycerin; Nic, nicotine. Statistically significant differences from the vehicle group are represented by asterisks (*); statistically significant differences between groups with and without nicotine (at the same PG/VG concentration) are represented by carets (^) ($p < 0.05$).

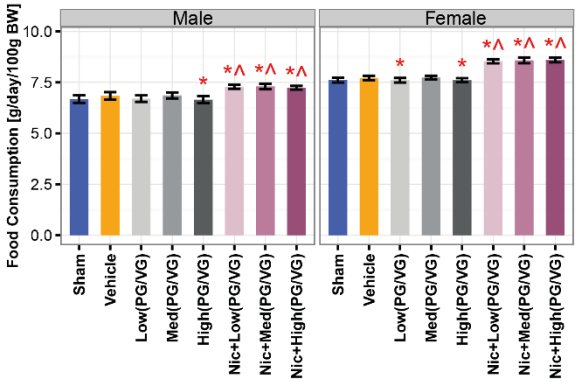
A



B

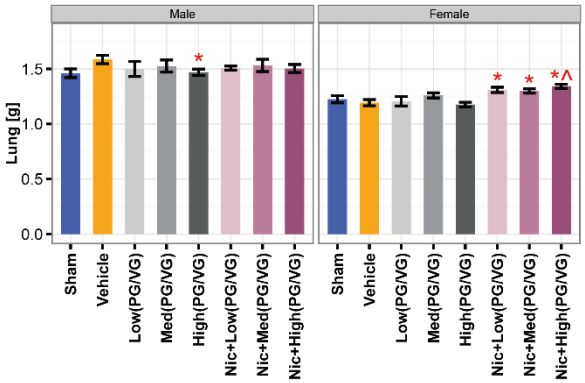


C

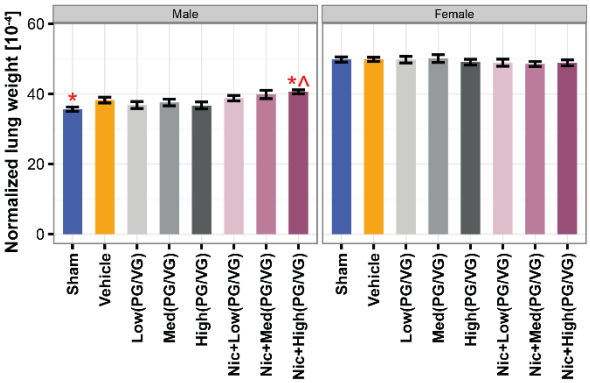


Supplementary Figure 3. Lung weights. (A) Absolute and (B) normalized lung weights to body weights for male and female rats (mean ± SEM, N=10). Statistically significant differences to the vehicle group are represented by asterisks (*); statistically significant differences between groups with and without nicotine (at the same PG/VG concentration) are represented by carets (^) (p-value < 0.05).

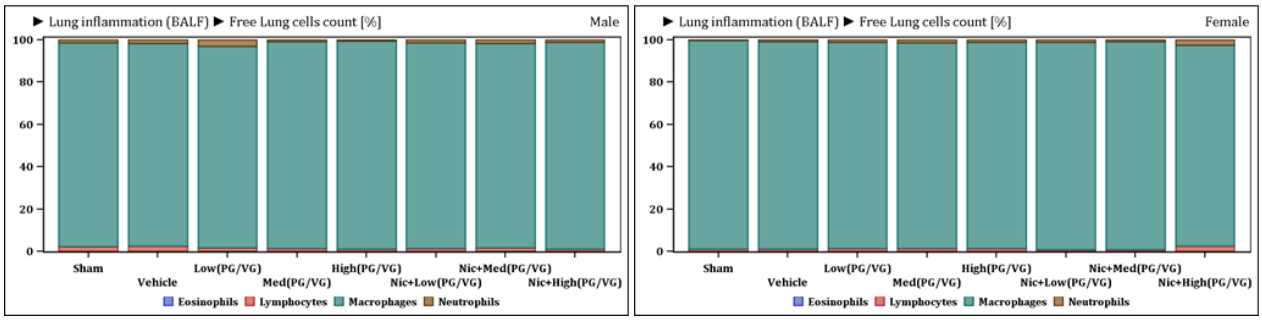
A



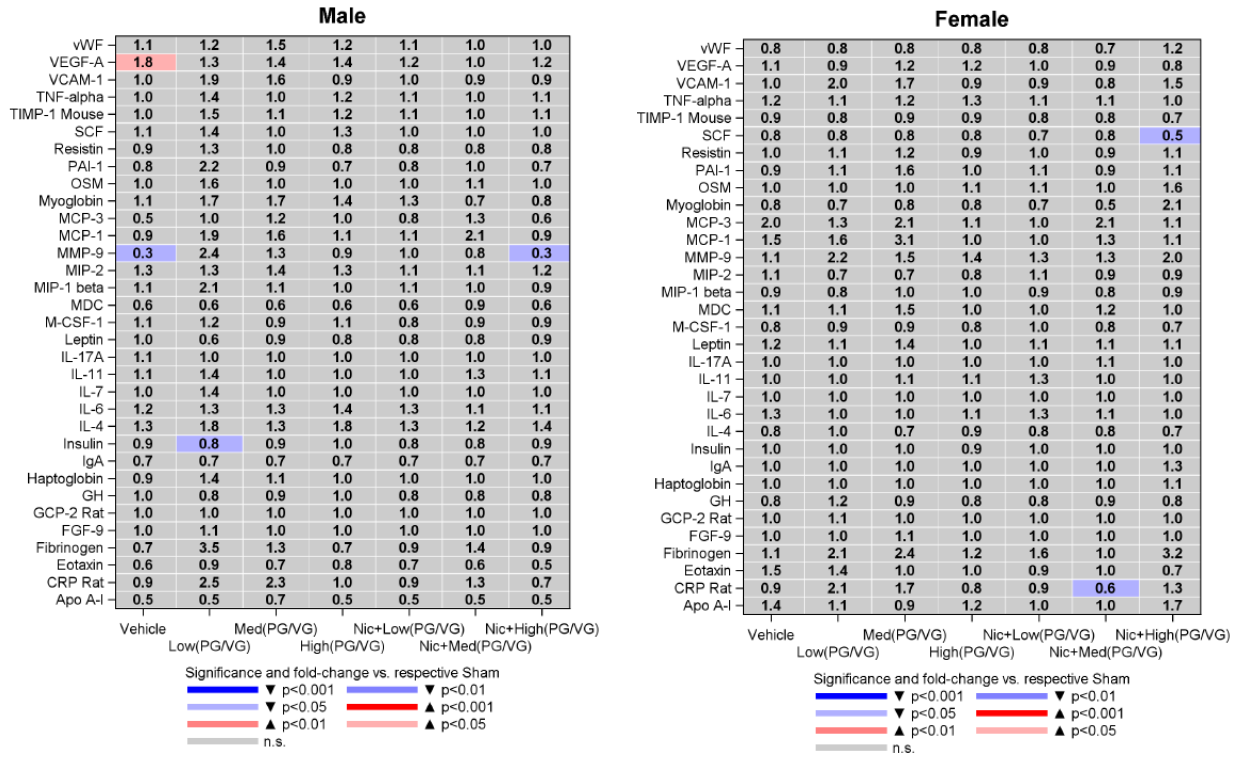
B



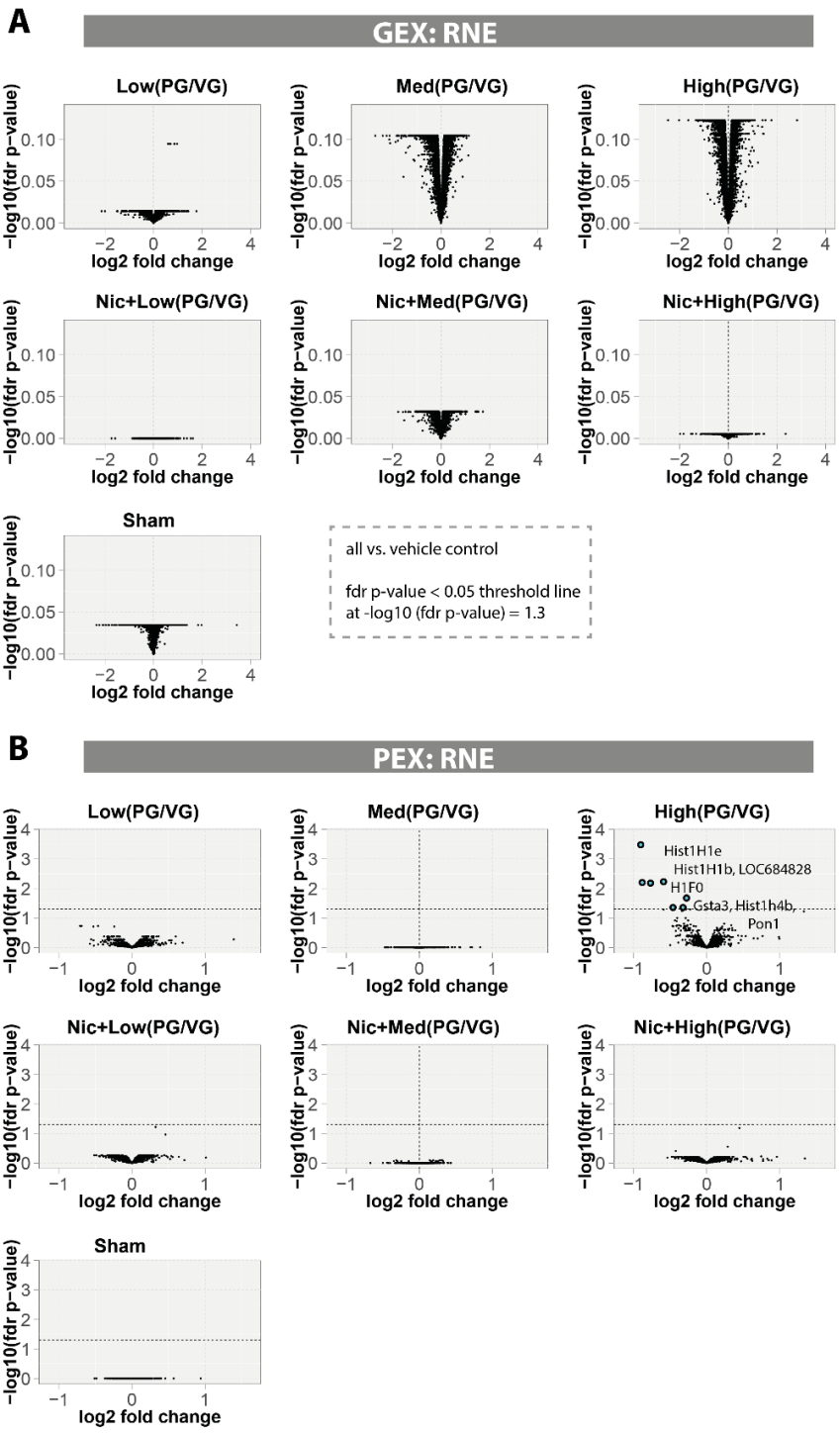
Supplementary Figure 4. Differential free lung cell counts in BALF (relative). Relative free lung cell counts in BALF from male (left panel) and female (right panel) rats exposed to different concentrations of the aerosols after 90-day exposure, N=10. PG, propylene glycol; VG, vegetable glycerin; Nic, nicotine.

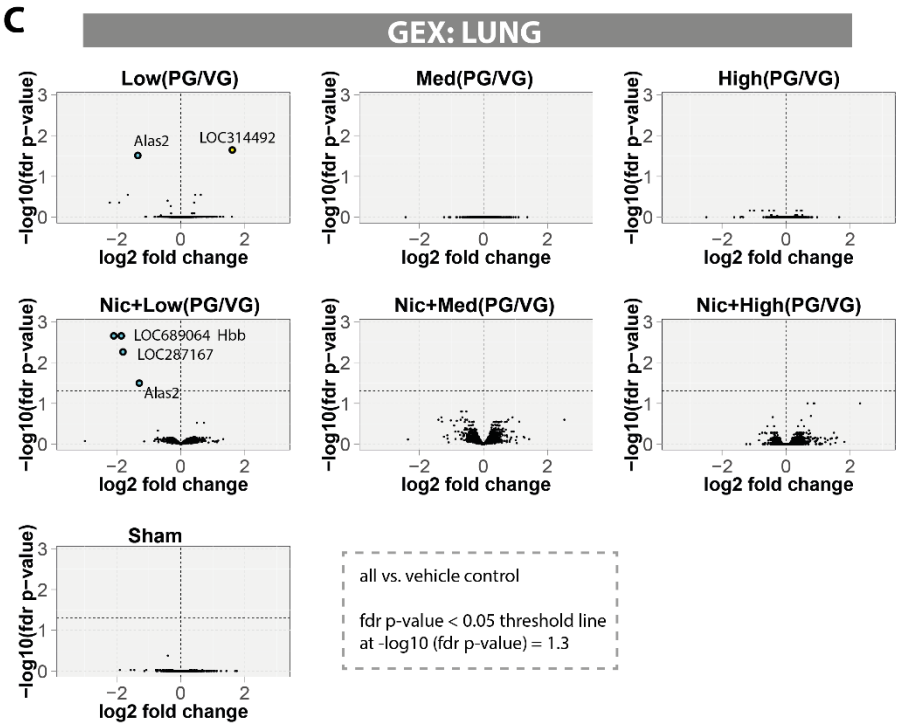


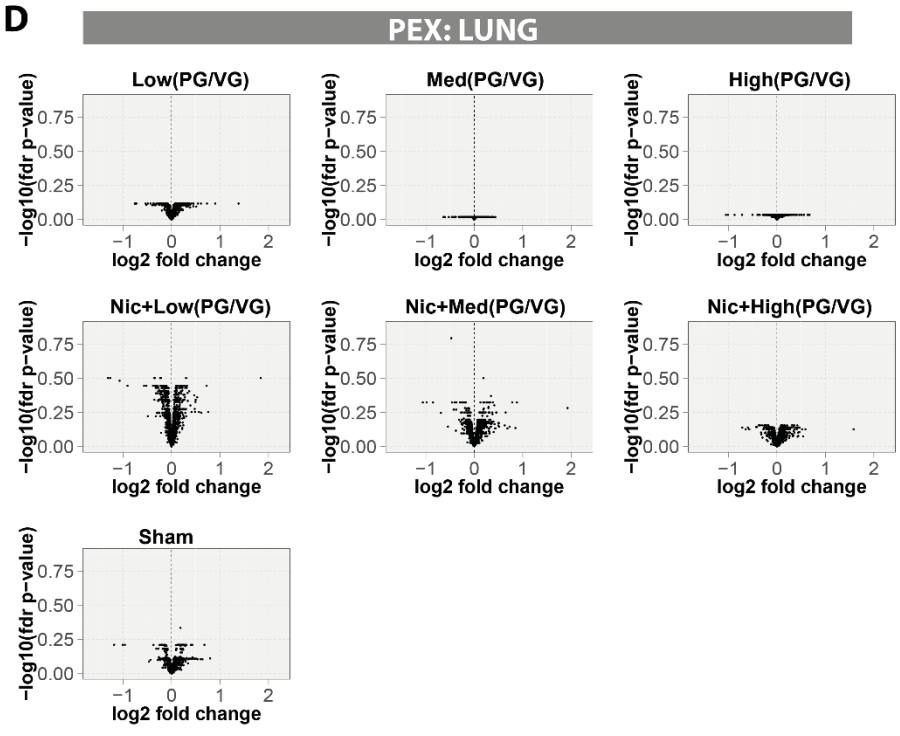
Supplementary Figure 5. Multi-analyte profiling of BALF. Heatmap of analytes present in BALF after 90-day exposure. The data indicate fold-change (numbers) and significance (colored boxes) in analyte concentrations in male (left panel) and female (right panel) rats in exposed groups compared with the sham group. Only analytes with values above the lower limit of quantification for male and female rats from all the groups are shown. PG, propylene glycol; VG, vegetable glycerin; Nic, nicotine.



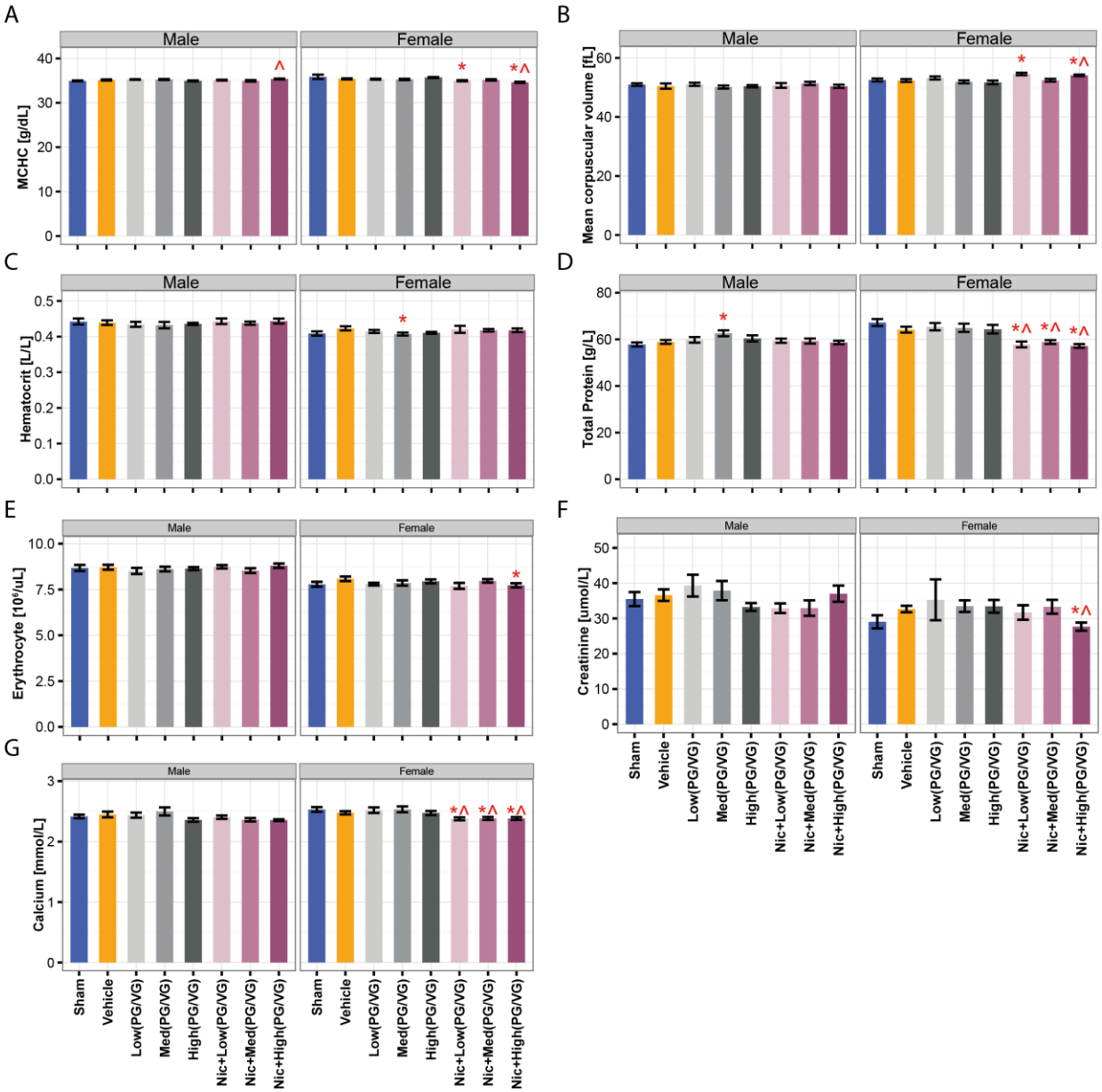
Supplementary Figure 6. Gene (GEX) and protein (PEX) expression volcano plots for respiratory tract organs. Volcano plots show the amplitude (log2 fold-change, x-axis) and significance (-log10 fdr-adjusted p-value, y-axis) for each quantified transcript/protein. Transcripts/proteins with an fdr-adjusted p-value < 0.05 are considered significant and are shown as yellow (up) and cyan (down) dots above the dotted significance threshold line. (A) GEX for respiratory nasal epithelium, (B) PEX for respiratory nasal epithelium, (C) GEX for lung tissue, and (D) PEX for lung tissue.


A
GEX: RNE
Low(PG/VG)
Med(PG/VG)
High(PG/VG)
Nic+Low(PG/VG)
Nic+Med(PG/VG)
Nic+High(PG/VG)
Sham
−log10(fdr p−value)
log2 fold change
all vs. vehicle control
fdr p-value < 0.05 threshold line
at -log10 (fdr p-value) = 1.3
B
PEX: RNE
Hist1H1e
Hist1H1b, LOC684828
H1F0
Gsta3, Hist1h4b,
Pon1

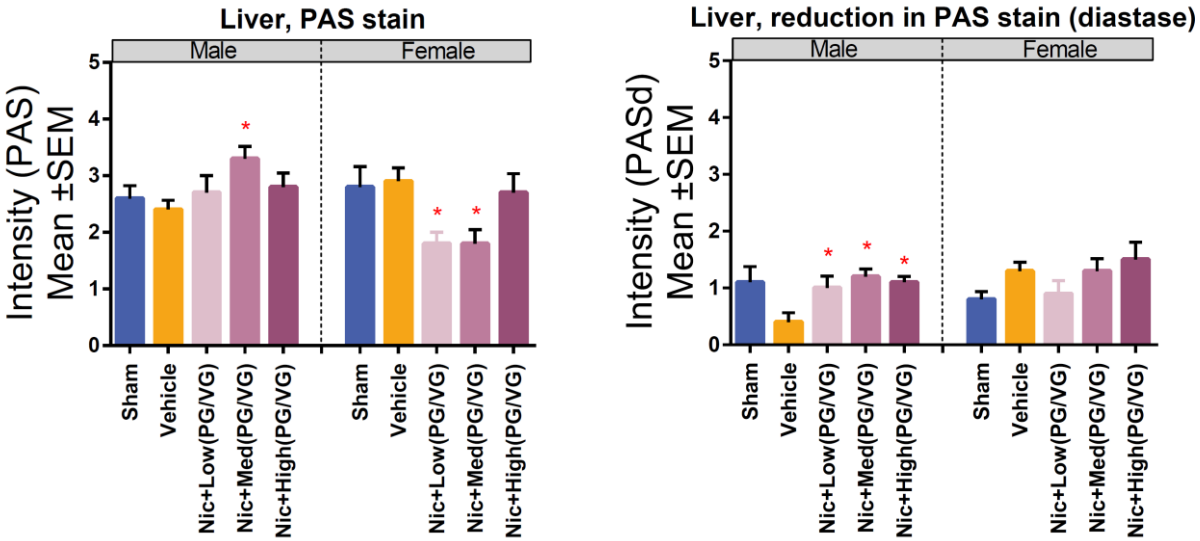

C
GEX: LUNG
Low(PG/VG)
Med(PG/VG)
High(PG/VG)
Nic+Low(PG/VG)
Nic+Med(PG/VG)
Nic+High(PG/VG)
Sham
Alas2
LOC314492
LOC689064 Hbb
LOC287167
−log10(fdr p−value)
log2 fold change
all vs. vehicle control
fdr p-value < 0.05 threshold line
at -log10 (fdr p-value) = 1.3


D
PEX: LUNG
Low(PG/VG)
Med(PG/VG)
High(PG/VG)
Nic+Low(PG/VG)
Nic+Med(PG/VG)
Nic+High(PG/VG)
Sham
−log10(fdr p−value)
log2 fold change

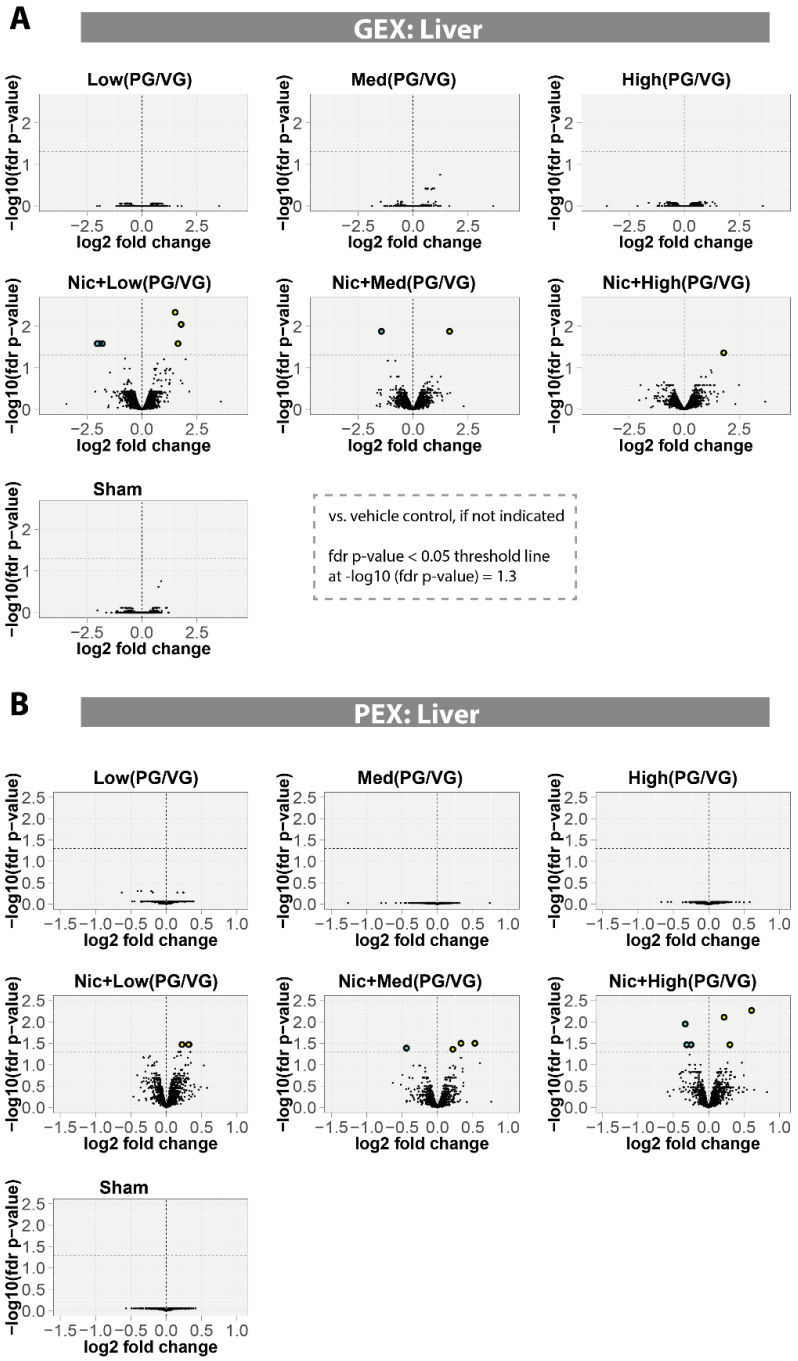
Supplementary Figure 7. Additional blood parameters. (A) Mean corpuscular hemoglobin concentration (MCHC), (B) mean corpuscular volume (MCV), (C) hematocrit, (D) total protein, (E) erythrocyte counts, (F) creatinine concentration, and (G) calcium concentration (mean ± SEM, N=10). Statistically significant differences to the vehicle group are represented by asterisks (*); statistically significant differences between groups with and without nicotine (at the same PG/VG concentration) are represented by carets (^) (p-value < 0.05).



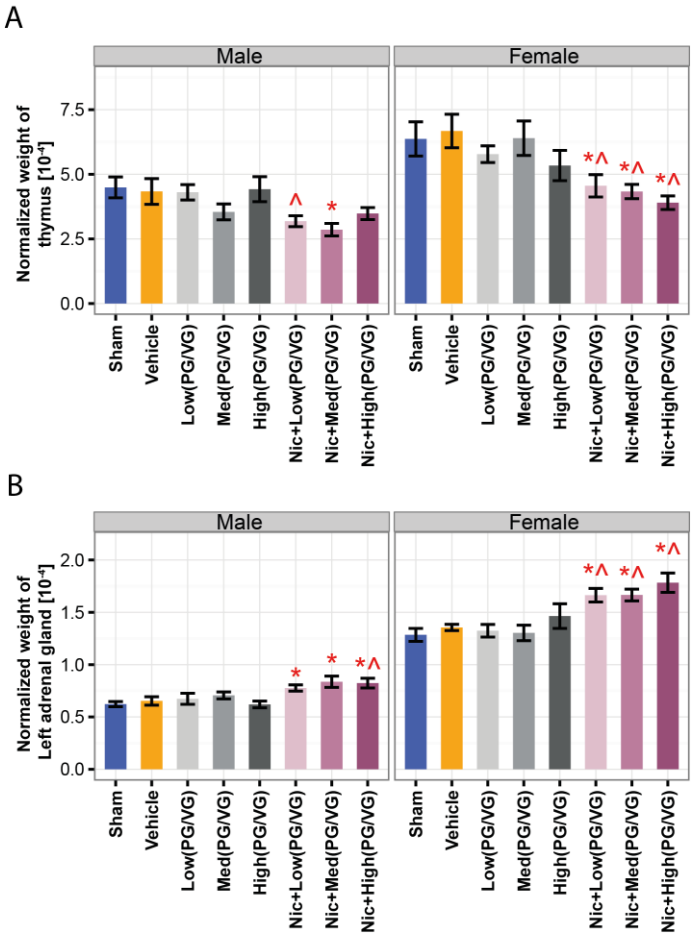
Supplementary Figure 8. PAS staining liver. Intensity scores for PAS stain (left) and reduction in the PAS stain (right) in liver. Statistically significant differences to the vehicle group are represented by asterisks (*) (p-value < 0.05).



Supplementary Figure 9. Gene (GEX) and protein (PEX) expression volcano plots for liver. Volcano plots show the amplitude (log2 fold-change, x-axis) and significance (-log10 fdr-adjusted p-value, y-axis) for each quantified transcript/protein. Transcripts/proteins with an fdr-adjusted p-value < 0.05 are considered significant and are shown as yellow (up) and cyan (down) dots above the dotted significance threshold line.



Supplementary Figure 10. Exposure-related systemic stress response. (A) Thymus and (B) left adrenal gland weight normalized to body weight for male and female rats (mean ± SEM, N = 10). Statistically significant differences from the vehicle group are represented by asterisks (*); statistically significant differences between groups with and without nicotine (at the same PG/VG concentration) are represented by carets (^) (p-value < 0.05).



Supplementary Tables

Supplementary Table 1. Analytical parameters for aerosol characterization.

Parameter	Assay principle	Determination schedule for exposure chambers 1 and 2 (Sham and vehicle)	Determination schedule for exposure chambers 3–5 (PG/VG) and 6–8 (Nic + PG/VG)
TPM	Gravimetric, after trapping on Cambridge filters	4 times per 6 h; 30 min per sampling [#]	4 times per 6 h; 30 min per sampling
Nicotine	Capillary gas chromatography after trapping on sulfuric acid-impregnated diatomaceous earth	1 time per week; 30 min per sampling	4 times per day, 30 min per sampling every day (chambers 6–8)
Propylene glycol	Capillary gas chromatography after trapping on and extraction from Cambridge filter pads	1 time per week; 30 min per sampling	4 times per day, 30 min per sampling
Vegetable glycerin	Capillary gas chromatography after trapping on and extraction from Cambridge filter pads	1 time per week; 30 min per sampling	4 times per day, 30 min per sampling
Temperature	Thermistor probe Pt100	Continuously	Continuously
Flow rate through exposure chamber	Pressure difference over the Venturi tube	Continuously	Continuously
Conductivity	Ion concentration of aerosol	1 time per week (chamber 2)	NA
Relative humidity*	Capacitive measurement	Continuously	Continuously
Particle size distribution	Spectrophotometric, using aerodynamic particle sizer	≥1 time per week (chamber 2)	≥1 time per week

*Relative humidity was measured only in the sham exposure chamber (chamber 1) because TPM may impede the performance of sensors in chambers with aerosols. PG, propylene glycol; VG, vegetable glycerin; Nic, nicotine; TPM, total particulate matter.

[#]TPM were measured 3 times per 6 h on every Wednesday because the sampling ports were used for PG and VG sampling.

Supplementary Table 2. Lipidomics data for lung, liver, and serum.
[separate Excel file]

Supplementary Table 3. Histopathological evaluation results.
[separate Excel file]

Supplementary Table 4. Data for other endpoints.
[separate Excel file]