STUDY REPORT SYSTEMS TOXICOLOGICAL EVALUATION OF THS2.2 90 DAY INHALATION STUDY OECD 413 – OECD PLUS PHASE

(STUDY NUMBER **150082/150083**)

Table of Content

Table of Content 2
Table of Figures 4
Table of Tables 5
1 Approval 6
1.1 Study Director's Statement of Compliance 6
1.2 Quality Assurance Statement 7
2 General Information 7
2.1 Names and Addresses 7
2.2 Study Schedule 8
2.3 Test Guidelines 8
3 Study Plan Deviations/Amendments 9
4 Abstract/Executive Summary 11
4.1 Introduction 11
4.2 Study Objectives 12
4.3 Experimental Design 13
4.4 Results and Main Conclusions 13
5 Introduction 15
6 Objectives 17
7 Experimental Design 18
7.1 Test and Reference Items 18
7.2 Test System 18
7.3 Experimental Groups and Dosing 19
7.4 Sample Generation 20
7.5 Scope of Reporting 20
7.6 Lung Sample Cryoslicing 21
7.7 Respiratory Nasal Epithelium Processing for RNA Extraction 21
7.8 Quantity Determination of Isolated RNA and Protein 21
7.9 Quality Control of the Isolated RNA and Protein 22
7.9.1 Affymetrix GeneChip processing 22
7.9.2 Data processing for gene expression analysis 22
7.10 iTRAQ LC MS/MS Procedure 24

7.10.1 Sample lysis and labeling for quantitative proteomics 24
7.10.2 Definition of iTRAQ analysis sets 25
7.10.3 Mass spectrometry for identification and quantification 25
7.10.4 Number of replicates per group and protein identification 26
7.11 miRNA Microarray 28
7.11.1 miRNA processing 28
7.11.2 miRNA array data processing and quality control 28
8 Statistical Methods and Computational Methods 30
8.1 Sample Randomization Protocol 30
8.2 Computational Methods 30
8.2.1 Gene/protein-set analysis 30
8.2.2 Network Perturbation Amplitude 30
8.2.3 Biological Impact Factor 31
9 Results 34
9.1 Body Weight 34
9.2 RNE Transcriptome at the 90-day Exposure Time Point and after the 42-day Recovery Period in male and female rats 36
9.2.1 Differentially expressed genes in RNE after 90 days of exposure to 3R4F and THS2.2 or after the 42-day recovery period 36
9.2.2 Biological network perturbation amplitude analysis in RNE after 90 days of exposure to 3R4F and THS2.2, and after a 42-day recovery period 39
9.3 Lung Transcriptome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats 42
9.3.1 Differentially expressed genes in the lung transcriptome after 90 days of exposure to 3R4F or THS2.2, or after a 42-day recovery period 42
9.3.2 Biological network perturbation amplitude analysis in RNE after 90-days exposure to 3R4F and THS2.2 or after a 42-day recovery period 47
9.4 Lung Proteome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats 54
9.4.1 Differentially expressed proteins in lung tissue 54
9.4.2 Set enrichment analyses to determine the effects of the exposure conditions on groups of proteins 63
9.5 Lung miRNAome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats 67

9.5.1 Differentially expressed miRNAs in lung tissue 67
9.6 Liver Transcriptome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats 72
9.6.1 Differentially expressed genes in liver tissue after 90 days of exposure to 3R4F or THS2.2, or after a 42-day recovery period 72
10 Discussion and Conclusion 79
11 Archiving 82
12 Abbreviations 83
13 References and Sources 84
13.1 Literature 84
13.2 Philip Morris International's List of Standard Operating Procedures (SOP) and Work Instructions (WKI). 90
13.3 Philip Morris International's Internal Documents 90
14 Supplementary Tables and Supplementary Figures 92
14.1 Supplementary Tables 92
14.2 Supplementary Figures 99
15 Appendix 100

Table of Figures

Figure 1: Outline of the OECD and OECD Plus study designs 16
Figure 2: Median parent ion intensity of the reference mix sample in iTRAQ male samples 27
Figure 3: Volcano plots representing the expression profiles of differentially expressed genes in RNE samples 36
Figure 4: Significantly differentially expressed genes in the RNE of male rats 37
Figure 5: Heatmap of changes in gene expression in RNE 38
Figure 6: Biological impact of the different exposures in the RNE of male rats displayed as starplots 40
Figure 7: Heatmap of subnetwork perturbations in the RNE of male rats 41
Figure 8: Volcano plots showing the systems response profiles in lung transcriptome 42
Figure 9: Significantly differentially expressed genes in lung tissue of male and female rats. 44
Figure 10: Heatmap of changes in gene expression in lung tissue 46
Figure 11: Biological impact of the different exposures in lung tissue of male and female rats displayed as starplots 49

Figure 12: Heatmap of subnetwork perturbations in lung tissue of male and female rats 52
Figure 13: Volcano plots representing proteome response profiles in lung 56
Figure 14: Significantly differentially expressed proteins in lung of male and female rats 58
Figure 15: Fold-change and significance of the differentially expressed proteins across the exposure conditions 60
Figure 16: Enriched pathways for the differentially expressed proteins under the different exposure conditions 65
Figure 17: Volcano plots representing the expression profiles of differentially expressed miRNA in lung samples 68
Figure 18: Heatmap of changes in miRNA expression in lung tissue 69
Figure 19: Quantile-quantile plots (QQ plot) of miRNA differential expression 71
Figure 20: Significantly differentially expressed genes in liver tissue of male and female rats 74
Figure 21: Heatmap of changes in gene expression in liver tissue 76

Table of Tables

Table 1: Study Plan Deviations 9
Table 2: Exposure groups selected for the exploratoryomics OECD Plus study 19
Table 3: Collected fluids and tissues and their intended use 20
Table 4: Sample sizes used for the transcriptomics analysis 23
Table 5: Number of valid replicates of lung proteomics profiles 26
Table 6: Number of proteins in the lung samples considered for differential expression analysis. 27
Table 7: Number of valid replicates of lung miRNA profiles 29
Table 8: “Unique” genes that were differentially expressed only in lung of female rats after THS2.2 exposure. 47
Table 9: “Unique” proteins that were differentially expressed only after THS2.2 treatment 61
Table 10: “Unique” genes that were differentially expressed genes in liver only after THS2.2 exposure. 77
Table 11: Overarching summary of results 81

1 Approval

1.1 Study Director's Statement of Compliance

The accountability of this study was with the study director. The study director takes accountability for the study having been executed as defined in the study plan.

The laboratory heads take accountability for the data generated in their laboratories.

This document has been approved by:

Name	Justification	Date / Signature
Hoeng, Julia	Approved by Program Leader, Systems Toxicology	19 April 2016 Julia Hoeng
Ivanov, Nikolai	Approved by Manager, Research Technology / Test Facility Manager/	19 April 2016
Emmanuel Guedj	Approved by Supervisor, Gene Expression	19 APR 2016
Stephanie Boue	Approved by author study plan transcriptomics/SME/ study director	19 April 2016
Elamin, Ashraf	Approved by Supervisor MS-based Proteomics / study director proteomics	19 Apr 2016
Ulrike Kogel	Approved by workpackage leader OECD plus part	19. April 2016 U. Kogel

1.2 Quality Assurance Statement

Not Applicable

2 GENERAL INFORMATION

2.1 Names and Addresses

Sponsor	Manuel Peitsch Philip Morris Products S. A. Quai Jeanrenaud 5 2000 Neuchâtel Switzerland
Program Leader Systems toxicology	Julia Hoeng Quai Jeanrenaud 5 2000 Neuchâtel Switzerland
Test Facility	Philip Morris Products S. A. Research & Development PMI Product Testing Quai Jeanrenaud 5 2000 Neuchâtel Switzerland
Test Facility Management	Nikolai Ivanov Genomics Research and Technology Laboratories Philip Morris Products S. A. Research & Development Quai Jeanrenaud 5 2000 Neuchâtel Switzerland

2.2 Study Schedule

Experimental starting date:	21 May 2013 (Lung slicing)
Experimental completion date:	05 September 2014 (Proteomics on left lung)

2.3 Test Guidelines

NA

3 STUDY PLAN DEVIATIONS/AMENDMENTS

During the execution of study 150082/150083, study plan amendment and deviations were raised. These are summarized in the table below (Table 1). For further details, the signed amendment can be found in the Appendix.

Table 1: Study Plan Deviations

Number	Description	Justification of impact on study data and integrity
1	Deviation described in the chapter: AMENDMENT TO THE OECD (+) ZURICH PROTEOMICS SP (150083), which is an amendment of the study plan 15006_150083 THS SP	To extend the analyses, we amended the original plan to use only samples generated from male rats to also use samples derived from female rats for proteomics analysis. More comprehensive analysis.
2	In the study plan the tested product was named ZURICH or ZRH. In the report we used exclusively the name "tobacco heated product" or THS (2.2).	Change in naming convention. No impact on data quality.
3	In the study plan (15006_150082 THS SP Chapter 3.2) secondary objectives were stated, which were not addressed in the reporting phase: • Correlation with body weight / food consumption (especially liver gene expression). • Correlation with mediators concentrations measured in BALF. • Correlation with immunohistochemical staining to analyze proliferation and apoptosis in lung. There may also be immunohistochemical staining of nasal epithelium (to be confirmed). • Correlation with proteomics profiles in blood, nasal epithelium, and lung. • To estimate whether there is a difference in gene expression between the different right lung lobes in male rats after 28 days of exposure. To compare right and left lung lobes at 90 days. • To compare THS2.2 aerosol and conventional cigarette smoke effects between male and female rats.	So, far the correlations have not been performed and are not in scope of the report. No impact on data quality.

(Table continues)

Number	Description	Justification of impact on study data and integrity
	• To establish a reference with 3R4F at 28 day to be used to confirm the validity of the tests for the Nike study (PMI internal report 150122, *2012*) Correlations have not been conducted.	
4	The study plan states that "Blood samples will be collected and analysed for their mRNA profile by microarray, their miRNA profile by Affymetrix array, and their DNA methylation profile", and also "Possible miRNA profiling and DNA methylation analysis will depend on method set-up and results in other experiments and may be the subject of an amendment of this study plan".	This is to clarify that miRNA isolation and analysis was only done from lung tissue but not from blood samples, and will therefore not be scope of the report. Also, DNA methylation will not be scope of the report. No impact on data quality.
5	The study plan (15006_150082 THS SP) stated "the concentration needs to amount to at least 50 ng/µl." This amount is not correct, the sentence should be "The concentration needs to amount to at least 33.3 ng/µl".	No impact on study data.
6	The proteomics study plan (15006_150083 THS SP) included this secondary objective: "To verify and quantitate selected protein targets in the left lung tissues (and serum if applicable)". Verification experiments using selected ion monitoring were not conducted for this study.	In the systems toxicology context, it was decided to focus on the global response rather than on specific biomarkers. No impact on primary objective.
7	The proteomics study plan (15006_150083 THS SP) contains the wrong number in the header. It should be 150083	No impact on data.

These amendment and deviations have been integrated in the description of methods and sample sizes.

4 ABSTRACT/EXECUTIVE SUMMARY

4.1 Introduction

It has been hypothesized that the chemical composition of aerosols generated by burning or heating tobacco and their toxicological effects depend to a great extent on temperature (Patskan 2003). The current version of the heat-not-burn product (THS2.2), which is operated at a temperature below 350°C, generates largely reduced amounts of harmful and potentially harmful constituents relative to the Kentucky Reference cigarette, 3R4F, under the smoking machine conditions.

To characterize toxicity and assess the hazard of THS2.2, a battery of chemical and toxicological assays was conducted. Among the studies performed, the 90-day inhalation study, as described in the Organisation for Economic Co-operation and Development (OECD) Test Guideline 413 (TG 413) (OECD 2009b), aimed to characterize the adverse effects caused by repeated daily inhalation exposure to THS2.2 for 90 days. Ninety-day inhalation studies have been used widely to assess the impact of cigarette modifications (Gaworski 2011a, Gaworski 2009, Gaworski 2011b, Renne 2006, Stavanja 2008, Theophilus 2007, Vanscheeuwijck 2002) or new cigarette designs, including designs in which tobacco is heated rather than burned (Moennikes 2008, Terpstra 2003, Werley 2008a). The standard study design has proven to be suitable for providing information on the hazards likely to arise from repeated exposure by the inhalation route, and has been shown to be capable of distinguishing between categories of products for systemic toxicological parameters as well as histopathological changes in the respiratory tract (e.g., (Werley 2008a)).

To obtain additional insights into the molecular mechanisms underlying the toxicity, we have added an additional study arm to generate transcriptomic, miRNAomic, and proteomic data. Experience has demonstrated that toxicological markers are detectable only after substantial damage has already occurred and therefore are considered to lack sensitivity (Ellinger-Ziegelbauer 2011). The combined approach of omics and conventional toxicology has been shown to be a useful tool for mechanistic investigation and for the identification of putative biomarkers (Cutler 1999, Ellinger-Ziegelbauer 2011).

4.2 Study Objectives

In the framework of a pre-clinical assessment program on the second generation of tobacco heating systems, THS2.2, a 90-day inhalation study was performed according to OECD TG 413 (OECD 2009b) to determine whether adult smokers who participate in clinical studies are exposed to increased or new hazards compared with smokers who continue to smoke conventional cigarettes.

The rationale for the pre-clinical assessment program was based on the United States Food and Drug Administration (FDA)'s Center for Tobacco Product draft guidance (US Food & Drug Administration 2012). The program included evaluation of the toxicological properties of the materials used to make the product, evaluation of the aerosol through non-targeted chemical screening, measurement of droplet size to characterize the aerosol to which smokers are exposed, measurement of the yield of harmful and potentially harmful constituents, assessment of the genotoxicity and cytotoxicity of the aerosol compared with combustable reference cigarettes, and a 90-day inhalation study on rats.

The scope of this report is the systems toxicological arm of the 90-day rat inhalation study, and to report on the findings that were obtained from analyzing the omics endpoints.

4.3 Experimental Design

Male and female Sprague-Dawley rats were exposed to three concentrations of (mainstream) aerosol from THS2.2 and three concentrations of mainstream smoke from 3R4F generated under the Health Canada Intense (HCI) smoke protocol (Health Canada 1999), and a full assessment of the parameters listed in TG 413 was carried out. Groups of rats were necropsied either at the end of the 90-day inhalation period or after a 42-day post-inhalation period to assess the reversibility or delayed occurrence of the findings. The overall health and well-being of the rats, together with smoke uptake and response, were tracked throughout the study. Lung inflammation, histology of respiratory and non-respiratory organs, and blood composition were assessed at the end of the 90-day inhalation and 42-day recovery periods. To obtain an understanding of the toxicity pathways involved in the response to the exposure regimen, additional rats were exposed in parallel to the same conditions and harvested for the systems toxicology study designated OECD 'Plus'. The molecular endpoints comprised transcriptomic data from respiratory nasal epithelium, and lung and liver tissue after 90 days exposure and after the 42-day recovery period; proteomic data from lung and liver tissue after 90-days exposure and after the 42-day recovery period; and microRNA (miRNA) omic data only from lung tissue after 90-days exposure and after the 42-day recovery period.

4.4 Results and Main Conclusions

Cigarette smoke (CS) from 3R4F reference cigarettes and the THS2.2 aerosol were reproducibly generated and delivered to the animals. Exposure was confirmed by the test atmosphere-concentration dependent amounts of nicotine metabolites recovered in the urine of the rats. The inhalation toxicity observed in animals exposed to THS2.2 aerosol was much lower than in the rats exposed to 3R4F CS. The biological effects caused by the THS2.2 aerosol were limited mainly to hematological changes, blood chemistries, and changes in liver weights, which have been shown to be related to nicotine exposure. Lung inflammation and changes in the histopathology of respiratory tract organs in THS2.2 aerosol-exposed rats were much lower than those in 3R4F CS-exposed rats.

In the respiratory nasal epithelium of male rats, a high number of differentially expressed genes (DEGs) was found when the rats were exposed for 90 days to the 3R4F CS compared to sham exposed rats. A total of 3900 DEGs were detected when the highest dose of 3R4F CS (23 µg/L nicotine) was applied. In contrast, at a 5% false-discovery rate (fdr), no DEGs were found after exposure to THS2.2 aerosol, not even when a dose of 50 µg/L nicotine was applied.

A clear concentration-dependent response was observed in the lung tissue of 3R4F CS-exposed male rats, with 788 DEGs being detected with the highest dose of 3R4F (23 µg/L nicotine) (5% fdr). The same 3R4F CS concentration induced about 1300 DEGs in the lung tissue of female

rats compared to sham exposed female rats. At the same nicotine concentration (23 µg/L) THS2.2 aerosol exposure resulted in 0 DEGs for male rats and in 1 DEGs for female rats. THS2.2 aerosol at 23 µg nicotine/L also did not induce any DEGs in lung tissue of male rats and resulted in 17 DEGs in the lung tissue of female rats.

Causal network analyses showed that, in both respiratory nasal epithelium and lung tissue and both sexes, 3R4F CS exposure caused perturbation of several biological processes including inflammatory, cell stress, cell fate, tissue repair and angiogenesis, and cellular proliferation processes. While perturbation of these processes was also seen in these tissues after THS2.2 exposure, the perturbation amplitudes at matched nicotine concentrations were much lower for THS2.2 aerosol than 3R4F CS – and even for the maximum THS2.2 aersol concentration (50 µg/L) remained well below the effects induced by the maximum 3R4F CS concentration (23 µg/L).

90 days of exposure to 3R4F CS also induced a clear change in the lung proteome compared to sham exposed rats. Exposure to THS2.2 aerosol, at corresponding nicotine concentrations, resulted in less differentially abundant proteins. Functional enrichment analyses further supported this difference at the level of biological processes.

Liver tissue of 3R4F-exposed male rats showed a limited but dose-dependent effect. In fact, 36 DEGs were observed upon 3R4F CS exposure only at higest nicotine concentration (23 µg/L). In contrast, only 2 DEGs were found in the THS2.2 group at similar nicotine concentration. Of note, only 10 DEGs were observed in the high THS2.2 group (50 µg/L nicotine).

In female rats the overall response in the liver was higher compared to male rats. Thus, at matching nicotine concentration (23 µg/L), 167 genes were differentially regulated upon 3R4F CS exposure and 93 in the THS2.2 group. Surprisingly, at lowest nicotine concentration (15 µg/L), the number of DEGs was higher in the THS2.2 group compared with the 3R4F one (78 vs 15).Overall, the data obtained in this 90-day inhalation study performed according to OECD TG 413 (OECD 2009b), including the systems toxicology arm, demonstrate that THS2.2 causes very low – and much reduced compared to 3R4F CS – toxicity in the respiratory and not respiratory tissues. These results, combined with the in vitro toxicology data that was collected, indicate that human adults using THS2.2 will not be exposed to additional hazards to those presented by smoking conventional cigarettes

5 INTRODUCTION

This 90-day inhalation study aimed to characterize adverse toxicological effects following repeated daily inhalation exposure to aerosol from our heat-not-burn THS2.2 product and to compare this with the effects of reference cigarette 3R4F exposure in male and female rats. This nose-only inhalation study was conducted with Sprague-Dawley rats according to the Organisation for Economic Co-operation and Development (OECD) Test Guideline 413 (TG 413) (OECD 2009a). The THS2.2 aerosol was generated according to the Health Canada Intensive Smoking Protocol (Health_Canada 1999) on 30-port carousel Smoking Device Docking Station Smoking Machines. The 3R4F cigarettes were smoked on SM 2000 Smoking Machines using the same smoking regimen. The test atmosphere was diluted to the target nicotine concentrations (8, 15, and 23 µg/L for 3R4F; and 15, 23, and 50 µg/L for THS2.2) with filtered and conditioned air and delivered into flow-pass nose-only exposure chambers. The duration of the exposure was 90 days using a 5 days per week, 6 hour per day exposure regimen. A time adaptation phase was included in week 1 with 7 days exposure to acclimatize the rats to the target aerosol concentration. Following the 90 days exposure period, the OECD 90-day groups of animals were dissected, while the OECD recovery groups of animals were dissected after an additional 42 day of non-exposed recovery period.

To monitor aerosol exposure, on-line monitoring of carbon monoxide as well as sample collection for analysis, and determination of nicotine concentration and total particulate matter from the breathing zone of the exposure chambers were performed. The uptake of aerosol was confirmed by blood carboxyhemoglobin (HbCO) measurements, and quantification of main nicotine metabolites and biomarkers of exposure in urine samples of the exposed rats. The respiratory response of the rats during exposure was monitored using head-out plethysmography. Throughout the study, the general health condition of the rats was assessed including in-life observations, body weight, ophthalmoscopy, and food consumption. At dissection, pathology (post-mortem gross necropsy and histopathological assessment of both respiratory and non-respiratory organs and tissues), blood analysis (complete blood counts, clotting potential, and blood chemistry), and bronchoalveolar lavage fluid (BALF) analyses were performed in accordance with OECD TG 413 (OECD 2009a).

The study was run in an AVA-licensed (Agri-Food and Veterinary Authority of Singapore) facility, where care and use of the rats was in accordance with the National Advisory Committee for Laboratory Animal Research Guideline 2004 (NACLAR 2004).

In addition to the standard toxicological assessment, a systems toxicology assessment was carried out (designated as ‘OECD Plus’), in which additional animals were concomitantly exposed in accordance with those in the above-mentioned 90-day inhalation study; i.e., using the same exposure chambers and using the same exposure regimen. As for the standard toxicological assessment, rats were exposed for 90 days, or were kept for the 42-day post-inhalation period. Additional groups of rats were exposed for 28 days, but these results are not in scope of this report. This systems toxicology arm was performed to gain a mechanistic understanding at the molecular level, and

comprised additional molecular endpoints that were added in order to complement the classical OECD approach with a systems toxicology approach (i.e., OECD plus).

The concept of combining the conventional toxicological assessment approach with additional molecular biological endpoints is presented in Figure 1. In brief, in the toxicity study, additional animals were used to obtain "omics" endpoints. These animals underwent the same exposure conditions at the same time as the animals in the OECD study. Organs (e.g., lung) of interest were collected from the exposed animals. Omics data were generated and computationally analyzed. The omics data were obtained from the exact same tissue sites where a phenotypic change was expected to occur based on the corresponding classical endpoints. This approach aims to facilitate the identification of the toxicity mechanisms at the cell, tissue, and organ levels (Sturla 2014).

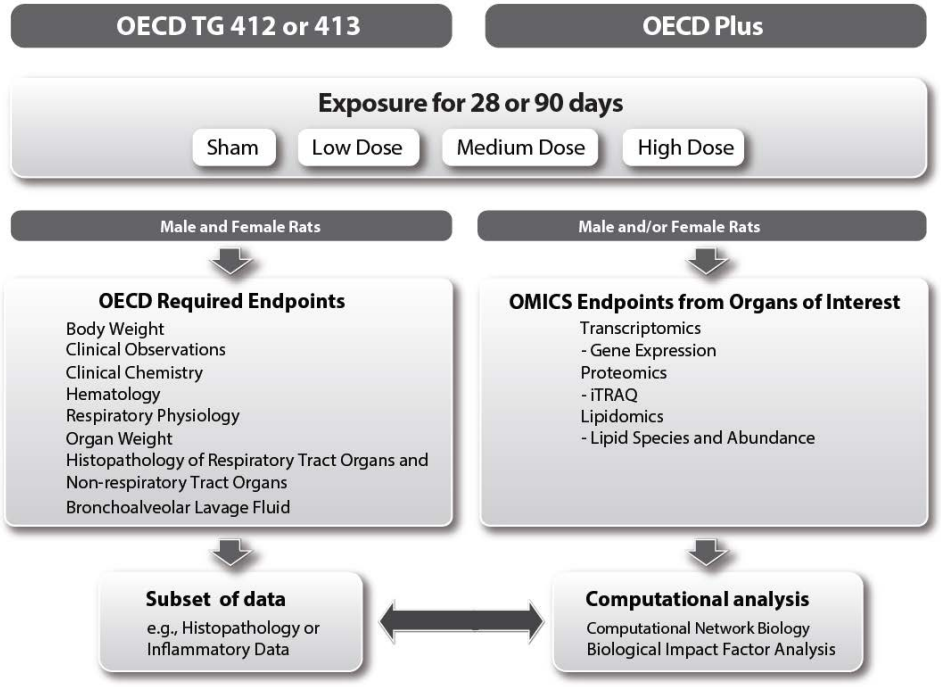


Figure 1: Outline of the OECD and OECD Plus study designs

Comparison of the endpoints in the classical OECD and OECD Plus concept.

The parts of the study up to the dissection are described in the overall inhalation study plan:

"15006 THS SP".

The animal dissection and sample collection are described in detail in the Omics study plans:

"15006_150082 THS SP" and "15006_150083 THS SP".

6 OBJECTIVES

The primary objective was to generate omics data from the tissues that were collected in the OECD study (15006 THS SR Part 1).

The secondary objective was to provide a systems toxicology picture of diverse organs of the rats in order to propose a mechanistic explanation of the data obtained from classical endpoints in the OECD part.

The overall objective was to compare the effects of THS2.2 versus 3R4F exposure at the molecular level (i.e., transcriptomics, miRNAomics, and proteomics levels).

7 EXPERIMENTAL DESIGN

7.1 Test and Reference Items

Test items and storage conditions are described in detail in the main study plan "15006 THS SP" in section 4.1. Briefly, reference cigarettes 3R4Fs were purchased from the University of Kentucky (Davies 2003). The test item was the THS2.2 sensorial medium (batch number B-00829). THS2.2 consists of a tobacco stick, a holder containing a battery, electronics, and a calibrated heating element designed to generate a temperature profile of 320–350°C to ensure constant delivery of nicotine. THS2.2 was produced by Philip Morris International, Neuchâtel, Switzerland. 3R4F and THS2.2 were conditioned according to ISO 3308 (International Organization for Standardization 2000) prior to use. Both 3R4F and THS2.2 were smoked at three target nicotine concentrations. The test atmospheres were characterized as detailed in the main study plan.

7.2 Test System

Experiments were conducted on male and female Sprague Dawley rats bred under specific pathogen-free conditions purchased from Charles River Laboratories, Inc., USA (at 8-weeks old). For more details see the main study plan, section 4.2: "15006 THS SP".

Test and reference items were administered by nose-only inhalation to ensure maximum reproducibility.

7.3 Experimental Groups and Dosing

All experimental groups included in the main study are described in the main study plan, section 4.3: "15006 THS SP".

Only some of the groups were subjected to the analyses included in this particular omics study (OECD+). Details of the study groups are summarized in Table 2.

Table 2: **Exposure groups selected for the exploratoryomics OECD Plus study.**

		90 d	90+42 d	28 d
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**	**M/F**
Sham	0	6/6	5/5	6/0
3R4F low	8	6/6	6/6	6/0
3R4F med	15	6/6	6/6	6/0
3R4F high	23	6/6	6/6	
THS2.2 low	15	6/6	6/6	6/0
THS2.2 med	23	6/6	6/6	6/0
THS2.2 high	50	6/6		
Total		42/42	35/35	30/0

A total of 184 rats were used. The animals were distributed in seven groups, two genders, three time points for male and two time points for female, and with five to six biological replicates in each group. Note that the groups of the 28 day time point are not in scope for this report.

F = female, M = male.

7.4 Sample Generation

The animal dissection and sample collection methods are described in detail in the study plan: "15006 THS SP".

The tissues listed in Table 3 were collected for the purposes indicated. Only the tissues in bold were processed further and are reported on this report. (The reasons are given in 7.5.)

Table 3: **Collected fluids and tissues and their intended use**

Blood/tissue	Planned Endpoint
Whole blood	Transcriptomics
Serum	Lipidomics
Serum	Proteomics
Left lung	**Transcriptomics/Proteomics/miRNAomics**
Right lung	Lipidomics
Liver	**Transcriptomics** / Lipidomics 28-day time point
Kidney	Transcriptomics
Left RNE	**Transcriptomics**
Right RNE	Proteomics
Trachea and larynx	Transcriptomics

RNE = respiratory nasal epithelium.

7.5 Scope of Reporting

This report contains an evaluation of the transcriptomics data from the left lung, respiratory nasal epithelium (RNE), and liver tissue after 90-days exposure to 3R4F or THS2.2 and after an additional 42-days recovery period. The report also contains an evaluation of the proteomics data from the left lung after 90 days exposure to 3R4F or THS2.2 and after the 42-days recovery period. The right RNE was collected but not tested for proteomics (Table 3).

For miRNA profiling, lung tissue of the 90-day time point was analyzed.

For lipidomics, the right lung and liver tissues were collected at a 28-day time point. However, the data for this time point are not within the scope of this report.

Whole blood for transcriptomics and serum for proteomics were collected. These samples have been stored in a freezer for future translational confirmation experiments.

The trachea and larynx tissues have also been stored for future use.

7.6 Lung Sample Cryoslicing

The Tissue Research Laboratory cryosectioned the left lung lobe into 40 µm-thick slices. Serial sections were collected alternately for transcriptomics, proteomics (iTraq analysis), and genomics analysis. (The sections for genomics were frozen for use in a different study plan.)

Details have been described in 15006_150082 THS SP.

The lung sections used for the transcriptomics analysis were placed in a 2-ml MagNA Lyser Green Beads (Roche CatNo 03 358 941 001) tube containing ceramic beads for homogenization in 600 µL RLT buffer (Qiagen CatNo 79216) + B-ME (Sigma), then vortexed for 30 sec and frozen on dry ice until RNA extraction (described in PMI_RD_WKI_001115). The final elution volume was 50 µl.

For proteomics, the lung sections were homogenized in lysis buffer and split into two aliquots (one for iTRAQ liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), and one was frozen and stored in a freezer). The iTRAQ aliquot was subjected to acetone precipitation by adding 4 x volume of pre-chilled acetone at −20°C for 2 hour, followed by centrifugation at 16,000 x g at 4°C. The precipitates were resuspended in 200 µL 0.5 M triethylammonium bicarbonate (TEAB) buffer containing 1 M urea and 0.1% SDS. The detailed procedure is described in WKI: PMI_RD_WKI_001043.

7.7 Respiratory Nasal Epithelium Processing for RNA Extraction

For transcriptomics, the left RNE samples were collected in 1.5-mL cryo tubes and frozen at −80°C. Samples were thawed, then 700 µL Qiazol buffer was added and the samples were vortexed for 30s. Total RNA extraction was performed using a miRNAeasy extraction kit (Qiagen, Hilden, Germany) according to the details described in PMI_RD_WKI_001117. Purified RNA was eluted in 30 µL RNase-free water.

7.8 Quantity Determination of Isolated RNA and Protein

The quantity of the isolated RNA was determined with the NanoDrop 1000 or 8000 spectrophotometer (Thermo Scientific, Waltham, MA, USA) or Spectramax Plus 384 (Molecular Devices, Sunnyvale, CA, USA). For the subsequent RNA amplification the protocol for high-throughput 3' in vitro transcription (HT 3′ IVT; Affymetrix, High Wycombe, UK) was followed. 100 ng RNA in total was required, and the starting concentration required was at least 33 ng/µL. If the concentration was too low, the sample was concentrated by vacuum centrifugation (Eppendorf SpeedVac). The RNA was dissolved in an appropriate volume of RNAse free water.

Protein concentrations of the samples were determined using the Bradford assay according to the supplier’s instructions. Procedure is described in WKI: PMI_RD_WKI_001103.

7.9 Quality Control of the Isolated RNA and Protein

RNA quality was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). The minimal acceptance criterion for the RNA samples that were analyzed for gene expression was RNA integrity score greater than 6.0 (WKI: PMI_RD_WKI_000978)

The quality control (QC) procedure to check the stability of the proteins before performing the experimental analyses consisted of two steps: (i) determination of protein concentrations and (ii) determination of sample integrity using LabChip technology according to WKI: PMI_RD_WKI_001131.

7.9.1 Affymetrix GeneChip processing

For the lung and RNE samples, amplification and labeling of RNA was done with the HT 3' IVT express kit (Affymetrix), as described in the GeneChip® HT 3' IVT express kit User Manual (Affymetrix).

For details, see PMI_RD_WKI_001118, PMI_RD_FOR_000763, PMI_RD_WKI_000813, PMI_RD_WKI_000816, and PMI_RD_FOR_000761.

Each sample was hybridized to an Affymetrix GeneChip® Rat Genome 230 2.0 Array.

7.9.2 Data processing for gene expression analysis

The Affymetrix genechip Rat Genome 230 2.0 Array was be used for hybridization, which simultaneously probes the expression of thousands of genes.

In accordance to PMI_RD_WKI_001228, PMI_RD_SOP_000346, raw CEL files was background-corrected, normalized, and summarized using frozen-Robust Microarray Analysis (fRMA) (McCall, Bolstad, & Irizarry, 2010). Background correction and quantile normalization was used to generate microarray expression values from all arrays passing quality control checks, which was performed using the custom CDF (Chip Description File) environment Rat2302_Rn_ENTREZG v19.0 (Dai et al., 2005). A log-intensities, normalized-unscaled standard error (NUSE) , relative log expression (RLE), median absolute value RLE (MARLE) and pseudo-images as well as raw images plot were generated for quality checks using R packages (AffyPLM; Bioconductor, Seattle, WA, USA) (Bolstad 2003a, Gautier 2004).

CEL files that fulfill at least one of the quality metric rules described below dropped for further analysis.

a) Pseudo-image displaying a spatial pattern covering approximately 10% of the pseudo-image.
b) Median NUSE >1.05
c) |Median RLE| >0.1

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

Subsequently, the RLE and NUSE-based metrics were recomputed, until no more CEL file is removed.

After the QC analysis, the following sample sizes were used for data analysis:

Table 4: Sample sizes used for the transcriptomics analysis

A) Respiratory nasal epithelium (RNE):

		90 d	90+42 d
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**
Sham	0	6/**2**	4/5
3R4F low	8	5/5	6/5
3R4F med	15	8*/6	5/6
3R4F high	23	5/4	5/4
THS2.2 low	15	7*/6	6/4
THS2.2 med	23	5/4	5/6
THS2.2 high	50	3/6	
Total		39/-$^{+}$	31/-$^{+}$

B) Lung tissue:

		90 d	90+42 d
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**
Sham	0	4/6	5/5
3R4F low	8	6/0	6/0
3R4F med	15	6/0	5/0
3R4F high	23	5/4	5/5
THS2.2 low	15	6/0	6/0
THS2.2 med	23	4/6	5/5
THS2.2 high	50	6/5	
Total		37/21	32/15

(Table continues)

C) Liver Tissue:

		90 d	90+42 d
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**
Sham	0	6/6	5/4
3R4F low	8	6/6	6/6
3R4F med	15	5/6	6/5
3R4F high	23	5/4	5/4
THS2.2 low	15	5/6	6/6
THS2.2 med	23	6/6	5/6
THS2.2 high	50	6/6	
Total		39/40	33/31

*Because the overall quality of the RNA was low, the RNE from reserve animals was also processed. [+] Female samples not included due to low number in 90d sham group (n = 2).

The table shows the numbers of replicates that where obtained after QC analysis of transcriptomics samples.

F = female, M = male.

For each treatment group and its associated control, a model for estimating the treatment effect was fitted with LIMMA (Smyth 2004) and adapted to the experiment statistical design. The p-values for each computed effect were adjusted across genes using the Benjamini-Hochberg false discovery rate method (fdr). The results were displayed as volcano plots (x-axis representing the estimated effect and the y-axis representing the –log10(fdr) for each gene). Differentially expressed genes (DEGs) were defined as the set of genes for which the fdr is below 0.05.

The QC analysis of the RNE samples from the female rats in the sham group showed that only two replicates were suitable for further processing, therefore no data analysis was performed for any of the female RNE samples.

7.10 iTRAQ LC MS/MS Procedure

7.10.1 Sample lysis and labeling for quantitative proteomics

Rat lung samples were homogenized and sonicated in tissue lysis buffer (BioRad, Hercules, CA, USA) in random order and the proteins were precipitated with acetone. The precipitate was resuspended in 0.5 M triethylammonium bicarbonate (TEAB, Sigma-Aldrich, St. Louis, MO, USA), 1 M urea (Sigma-Aldrich) and 0.1% SDS (Sigma-Aldrich). A 50 µg aliquot from each sample was labeled using the iTRAQ 8-plex kit according to the manufacturer's instructions (AB Sciex,

Framingham, MA, USA). Trypsin (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 tryptic-digested samples were labeled with the reporter-ion tags for the different treatment groups for the three defined iTRAQ multiplexing sets (7.10.2, Table 5). In addition to the individual samples, one common reference mix for each iTRAQ analysis set was prepared as an optional normalization standard. These reference mixes contained 50 µg of the protein extracts from each sample in each iTRAQ analysis set and these were also labeled with the iTRAQ reporter-ion tags. All the labeled samples in each iTRAQ set (and the corresponding reference mix) were pooled and dried in a SpeedVac. Samples were desalted first on 0.5-mL bed volume detergent removal columns (Pierce, Rockford, IL, USA) and then on 1-cc C18 reversed-phase SepPak columns (Waters, Milford, MA, USA) according to the manufacturers' manuals. Samples were dried in a SpeedVac and resuspended in nanoLC buffer A (5% acetonitrile (Sigma-Aldrich), 0.2% formic acid (Sigma-Aldrich)).

7.10.2 Definition of iTRAQ analysis sets

Three multiplexed iTRAQ analysis sets were prepared to analyze the samples from both genders and both time points (90 d and 90+42 d) (Table 5).

iTRAQ set 1 contained the female samples for the 90 d and 90+42 d time points. iTRAQ set 2 and iTRAQ set 3 contained the male sample groups for the 90 d and 90+42 d time points, respectively. Missing samples were replaced by the reference mix. These placeholder samples were removed during data analysis and processing.

7.10.3 Mass spectrometry for identification and quantification

Samples were analyzed in random order using an Easy nanoLC 1000 instrument (Thermo Scientific) connected online to a Q-Exactive mass-analyzer (Thermo Scientific). Peptides were separated on a 50-cm long C18RP RSLC Easyspray column (2-µm particle size; Thermo Scientific) at a flow rate of 200 nL/min with a 200 min gradient from nanoLC buffer A (5% acetonitrile, 0.2% formic acid) to 40% acetonitrile, 0.2% formic acid. Each sample was injected twice with two different analysis methods: a fast method and a sensitive method, as described by Kelstrup et al. (Kelstrup 2012). Both mass spectrometry runs were searched together as merged mass-lists against the *Rattus norvegicus* reference proteome (UniProt Proteome ID: UP000002494), version October 2013, canonical isoforms only) using Proteome Discoverer version 1.4.0.288 software (Thermo Scientific). SequestHT (implemented in Proteome Discoverer) was used as the search tool and iTRAQ-reporter ion intensities were determined from the Proteome Discoverer software. The Percolator node of the Proteome Discoverer software 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 by the Proteome Discoverer software. Global variance stabilizing normalization (VSN) was performed with the Bioconductor package in R (Huber

2002, Hultin-Rosenberg 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 2013). To detect differentially expressed proteins, a linear model was fitted for each exposure condition and the corresponding sham group, and p-values from a moderated t-statistic were calculated based on the empirical Bayes approach (Gentleman 2004a). Then the BH FDR method was used to correct for multiple testing effects. Proteins with an adjusted p-value <0.05 were considered as differentially expressed.

7.10.4 Number of replicates per group and protein identification

The number of replicates effectively available for each group is summarized in Table 5.

Figure 2 shows a representative result from a QC analysis. One replicate set of the male samples for the 90+42 d time point was excluded due to low signal intensity. After QC no samples were excluded from the analysis of the female sets.

Table 5: Number of valid replicates of lung proteomics profiles

		iTRAQ set 1		iTRAQ sets 2/3	
Group Name	Target nicotine (µg/L)	Female		Male	
		90 d	90+42 d	90 d	90+42 d
Sham	0	6	5	6	4
3R4F low	8			6	5
3R4F med	15			6	4
3R4F high	23	4	5	5	4
THS2.2 low	15			6	5
THS2.2 med	23	6	6	5	5
THS2.2 high	50	5	0	6	0

For the 90 d male set (iTRAQ set 2), one 3R4F high and one THS2.2 med sample were not available. For the 90+42 d male set (iTRAQ set 3), one sham, one 3R4F med, and one 3R4F high sample were not available. For the female set (iTRAQ set 1), two 3R4F high (90 d), one 3R4F high (90+42 d), one THS2.2 high (90 d), and one sham (90+42 d) sample were not available. One replicate set for the 90+42 d time point of the male samples was excluded due to low signal intensity

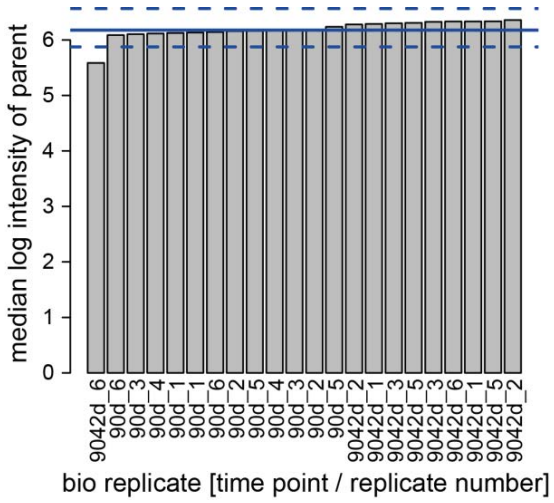


Figure 2: Median parent ion intensity of the reference mix sample in iTRAQ male samples

The median log ion intensity of the reference mix sample across all the analyzed replicates is shown. The solid line indicates the median, the dotted lines the boundaries of the 25%itle − 1.5 x interquantile range and 25%itle + 1.5 x interquantile range. Note that replicate 6 for the 90+42 d time point had a lower than expected intensity and was not included in the analysis.

In total, 3854 proteins were identified in the female rat samples and 4627 proteins were identified in the male rat samples. For the identification of differentially expressed proteins compared with the sham exposure group only proteins quantified in at least 2/3 of one compared group were considered. The number of proteins considered after filtering for each comparison is given in Table 6.

Table 6: Number of proteins in the lung samples considered for differential expression analysis.

		iTRAQ set 1		iTRAQ sets 2/3	
Group Name	**Target nicotine (µg/L)**	**Female**		**Male**	
		90 d	**90+42 d**	**90 d**	**90+42 d**
3R4F low	8			3088	2355
3R4F med	15			3089	2436
3R4F high	23	2621	2457	3086	2436
THS2.2 low	15			3088	2354
THS2.2 med	23	2589	2589	3086	2354
THS2.2 high	50	2590		3088	

The sham group was used as the reference.

7.11 miRNA Microarray

7.11.1 miRNA processing

MiRNA was labeled with the FlashTag™ Biotin HSR Labeling Kit as described in PMI_RD_WKI_001123. MiRNA analysis was performed on an Affymetrix GeneChip® miRNA 3.0 array starting from 200 ng of RNA.

7.11.2 miRNA array data processing and quality control

The raw data were preprocessed through a standard pipeline. Briefly, raw data CEL files were read using the read.celfiles function of the oligo packages in the Bioconductor suite of microarray analysis tools available for the R statistical software environment (Carvalho 2010), (Gentleman 2004b, R Core Team 2014). The QC was based on the arrayQualityMetrics package andexamined four distinct metrics: Euclidean distances between arrays in raw and normalized data heat maps, NUSE plots, and intensity distribution box plots (Kauffmann 2009). Arrays that were identified as outliers in at least two of the metrics were discarded, and the QC metrics were recalculated on the remaining arrays until all were accepted. This approach led to seven arrays being discarded in two iterations. Normalized probe-level data were obtained by applying robust multi-array normalization and summarized at the probe set level using the median polish method (Bolstad 2003b, Irizarry 2003). Only the miRNA probe sets that had significantly higher intensity values than their GC- and sequence length-matched background probes (p-value cut-off = 0.01) were considered as “present”, based on the manufacturer’s instructions (see “Detection” section in Affymetrix® miRNA QCTool users guide). If a miRNA probe set was “present” in more than half the samples of at least one sample group, it was retained for the downstream analysis, if not it was discarded together with all the probe sets measured on the array that belonged to other species. This preprocessing yielded 321 “present” miRNA probe sets.

After the QC analysis the sample sizes that were used for data analysis and evaluation are shown in Table 7.

Table 7: Number of valid replicates of lung miRNA profiles

Lung tissue:

A)

Delivered CEL files		**90 d**	**90+42 d**	**28 d**
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**	**M/F**
Sham	**0**	**5/6**	**5/5**	0/
3R4F low	**8**	**6**/0	**6**/0	0/
3R4F med	**15**	**6**/0	**6**/**0**	0/
3R4F high	**23**	**5/4**	**5/5**	
THS2.2 low	**15**	**6**/0	**6**/0	0/
THS2.2 med	**23**	**6/6**	**6/6**	0/
THS2.2 high	**50**	**6/6**		
Total (112)		**40/22**	**34/16**	0/

F = female, M = male.

B)

QC-passed CEL files		**90 d**	**90+42 d**
Group Name	**Target nicotine (µg/L)**	**M/F**	**M/F**
Sham	**0**	**4/6**	**5/5**
3R4F low	**8**	**6/**	**5/**
3R4F med	**15**	**5/**	**5/**
3R4F high	**23**	**5/4**	**4/4**
THS2.2 low	**15**	**6/**	**6/**
THS2.2 med	**23**	**6/6**	**6/6**
THS2.2 high	**50**	**5/6**	
Total (105)		**37/22**	**31/15**

F = female, M = male.

A) Number of Cel files generated; B) Number of Cel files that passed the QC and which were used for further analysis.

8 STATISTICAL METHODS AND COMPUTATIONAL METHODS

8.1 Sample Randomization Protocol

To compare (a) the effect of THS2.2 vs 3R4F, (b) the effect of the recovery period, and (c) the effect of sex, all the samples were randomized and processed together when possible. For the cryoslicing, the 90 d and 90+42 d samples were randomized in two batches.

For the transcriptomics randomization, potential cryoslicing batch effects, RNA extraction batch effects, and the effects of different Affymetrix fluidics stations were all taken into account. The randomization plan was provided by the bioinformatics team. For the iTRAQ analysis set, the samples were randomized according to the plan provided by the bioinformatics team.

8.2 Computational Methods

8.2.1 Gene/protein-set analysis

Gene-set analysis (GSA) of transcriptomics data is performed using the Piano package in the R statistical environment (Väremo 2013). Pathway maps are obtained from the KEGG database (Kanehisa 2014) and exported as R objects into the package graphite (Sales 2012). Gene-set enrichment is assessed using over-representation analysis (ORA) and gene-set analysis. For the ORA, a one-tailed Fisher's exact test is employed. For the GSA, the fold-change is used as the gene-level statistic, the mean is used as the set statistic and sample permutation (Q2) is used to assess statistical significance (Ackermann 2009). For each test, the p-value is adjusted using the Benjamini-Hochberg procedure. The gene-set enrichment are the gene-set statistics mentioned before.

8.2.2 Network Perturbation Amplitude

The collection of causal biological networks used in the study(s) was the rat network suite CBN v1.3. These networks range between a few dozen and two hundred nodes in size.

Network scoring exploits the network backbone nodes that are connected to downstream mRNA-abundance-nodes based on their known relationship reported in the literature. These signed relationships can be the backbone node increasing or decreasing the abundance of certain mRNAs and are used to infer the activation of a backbone node using transcriptomics data. Since not all backbone nodes have downstream mRNA nodes, the network models need to be prepared for scoring to improve the specificity and relevancy of the overall NPA score. Only backbone nodes that are on a directed path that starts and ends with a node that has downstream mRNA nodes are considered. After removing the nodes that do not satisfy the above, the largest connected component is kept. Finally the edges "causesNoChange" are disregarded for the scoring.

The Network Perturbation Amplitude (NPA) methodology (Hoeng 2014, Martin 2012) aims at contextualizing the high dimensional transcriptomics data by combining gene expression (log_2) fold-changes, β, into fewer differential node values (one value for each node of the network), *f*. The differential nodes values are determined by a fitting procedure that infers the values that best satisfy the directionality of the causal relationships (positive or negative signs) contained in the network model, while being constrained by the experimental data (the gene log_2-fold-changes, which are described as downstream effects of the network itself):

$$f = L_3^{-1} L_2^T \beta$$

where L is the signed weighted Laplacian of the network and the extra edges and nodes describing the downstream effects (to gene expression nodes); L_3 is the sub-matrix of L for the nodes in the network; and L_2 is the sub-matrix corresponding to the edges connecting the network nodes to the downstream gene expression nodes. The differential node values are in turn summarized as a single positive number, referred to as the amplitude of perturbation ("NPA scores").

$$NPA = \frac{1}{|E|} \sum_{e\ in\ E} (f(e_0) - \sigma(e) f(e_1))^2$$

where E is the set of edges in the network; |E| is its size; and e_0 and e_1 denote the start and the end, respectively, of an edge, e. The sum computing the NPA score can be expressed as $f^T Q f$, where Q is the signed Laplacian of the network when all of the edge signs have been reversed.

All the details of the methodology is described in a previous publication (Martin 2014).

For the NPA scores, a confidence interval accounting for the experimental variation and the associated p-value are computed. In addition, companion statistics are derived to inform on the specificity of the NPA score with respect to the biology described in the subnetwork model and are shown as *O and K* if their p-values are below the significance level (0.05). A network is considered to be significantly impacted if all of the three values (the p-value for experimental variation, *O, and K* statistics) are below 0.05.

8.2.3 Biological Impact Factor

The methodology Biological Impact Factor (BIF) (Martin 2014, Martin 2012) provides a unified and coherent framework for investigating mechanistic effects at each level of granularity of the biological processes represented in the network models. This methodology enables the derivation of the relative BIF, which is a measure of the overall biological impact across all network models. The BIF allows an assessment of the exposures in an objective, systematic, and quantifiable manner, computing a systems-wide and pan-mechanistic biological impact measure for a given substance, mixture, or test item.

The network models are organized into five major families, representing global biological processes: cell stress, cell proliferation, inflammation, tissue repair, and cell fate. Each of them comprised of various subnetwork models, representing the specific processes within a given network family.

An aggregation of the NPA should satisfy the following criteria:

Credibility: The BIF should be an intuitive and logical aggregation scheme.

Fairness: All network families are equally important.

Understandable: Each decision or step has a defined rationale.

To fulfil the first criteria, network scores are aggregated via a weighted sum (as every scoring in NPA is additive). In addition, only networks that have the three statistics p-values below 0.05 (*O, *, K*) will be considered (denoted by N_i^*). The BIF is defined as:

$$BIF = \sum_{\mathcal{N}_i^*} w_i NPA(\mathcal{N}_i) = \sum_{\mathcal{N}_i^*} w_i \sum_{e \in E(\mathcal{N}_i)} (f(e_0) + \sigma(e) f(e_1))^2$$

Where, f denotes the function on the network nodes describing the differential node values; $E(N_i^*)$ denotes the network edges; and e_0 and e_1 denote the start and the end, respectively, of an edge, e .

For the second criteria, a "fair" BIF would represent the equality of the biological processes (networks) in a given system (e.g., 2D normal human bronchial epithelial cells, 3D organotypic tissue cultures, or lung tissue samples from in-vivo studies). If all subnetworks models were perfectly fitted by the gene expression data, the contribution of each subnetwork to the BIF is assumed to be equal. Therefore, the methodology considers that all specific processes represented in the subnetwork models, are equally involved and important for a given network family (e.g., cell stress). A perfect fit of the network will maximize the NPA value while being smooth over the network. Using the notation above, a “perfect” network response (differential node values) is both maximizing the f^TQf with $||f||_2 = 1$ and minimizing the f^TL_3f (the NPA criteria (Hoeng 2014)). Additionally, there is a constraint that L_3f is vanishing for the network nodes that do not have direct downstream gene expression nodes (Hoeng 2014, Martin 2012), as those are expected to be in the image of L^T_2. Such nodes are denoted as *NH*. Therefore, the following optimization problem will be solved:

$$argmax_{f\ s.t.(L_3 f)|_{NH}=0} \frac{f^T Q f}{f^T L_3 f}$$

To rebalance the network families, the average over the subnetworks will be considered.

For each edge $e = (e_0, \sigma(e), e_1)$, the number of occurrence of an edges, *e*, in the set of networks used in the BIF calculation is denoted by o(e). Therefore, the maximum occurrence-corrected amplitude of a network N_i can be defined as:

$$NPA_{max}(\mathcal{N}_i) \doteq \sum_{e \in E(\mathcal{N}_i)} \frac{1}{o(e)} (f_{max}(e_0) + \sigma(e) f_{max}(e_1))^2$$

Finally, to normalize the NPAs, a correction factor of $|E(N_i)|/NPA_{max}(N_i)$ is applied. Considering all of these factors, the BIF is defined as the weighted sum (F denotes the subnetwork families), as follows:

$$\sum_{\mathcal{F}} \frac{1}{|\mathcal{F}|} \sum_{\mathcal{N}_i^* \in \mathcal{F}} \frac{1}{NPA_{max}(\mathcal{N}_i)} \sum_{e \in E(\mathcal{N}_i)} \frac{1}{o(e)} (f(e_0) + \sigma(e) f(e_1))^2$$

Therefore, the contribution of each subnetwork F is:

$$\frac{\frac{1}{|\mathcal{F}|} \sum_{\mathcal{N}_i^* \in \mathcal{F}} \frac{1}{NPA_{max}(\mathcal{N}_i)} \sum_{e \in E(\mathcal{N}_i)} \frac{1}{o(e)} (f(e_0) + \sigma(e) f(e_1))^2}{BIF_{SDP}}$$

9 RESULTS

9.1 Body Weight

Body-weight trends before exposure, during the first week of adaptation, and progression during the 90-day exposure and 42-day post-exposure recovery periods were very similar to those in the OECD groups (see "15006 THS SR Part 1"). Mean body weights showed the following trends:

Male rats after the 90-day exposure period:

- Significantly lower mean body weights ($p < 0.05$) were found in the 3R4F- and THS2.2-exposed animals compared with the sham group. Statistically significant weight differences between the sham and 3R4F high groups were observed from study days 25 to 81, and the weights were lower in all three 3R4F groups from study days 53 to 81. In the THS2.2 low group, the weights were lower from study days 39 to 92 and in THS2.2 med group the weights were lower from study days 46 to 92 compared with the sham group.

- Mean body weight in 3R4F high was lower than in THS2.2 high ($p < 0.05$) during study days 53–88.

Male rats after the 42-day recovery period:

All animals gained weight, the mean body weights was approaching that of the sham group.

Female rats after the 90-day exposure period:

- Weight loss in the sham group was higher than in the 3R4F- and THS2.2-exposed groups during the first week of dose adaptation; therefore, the sham group had the lowest mean body weight throughout the exposure period. The weight loss in the sham group is likely due to stress from the tube restrain during nose-only exposure.

- Higher mean body weight in THS2.2 med group compared to the sham group ($p < 0.05$) was observed from study days 18 to 92. We considered this result largely technical because of lower mean body weight in the sham animals.

- At equivalent nicotine doses, mean body weight in 3R4F high was lower than in the THS2.2 med group ($p < 0.05$) during study days 25 to 92.

Female rats after the 42-day recovery period:

- In all the the 3R4F- and THS2.2-exposed groups, the mean body weights of all the animals approached that of the sham groups.

- Mean body weight of the 3R4F high group was lower than that of the THS2.2 med group ($p < 0.05$).

9.2 RNE Transcriptome at the 90-day Exposure Time Point and after the 42-day Recovery Period in male and female rats

9.2.1 Differentially expressed genes in RNE after 90 days of exposure to 3R4F and THS2.2 or after the 42-day recovery period

The amplitude and statistical significance of gene expression changes in RNE of 3R4F- and THS2.2-exposed rats after 90 days of exposure, or after the 42-day recovery period were compared with the sham group and visualized as volcano plots (Figure 3). The analysis could only be performed for the male rat samples, because after quality control only two replicates of the female sham group were suitable for further processing (see 7.9.2).

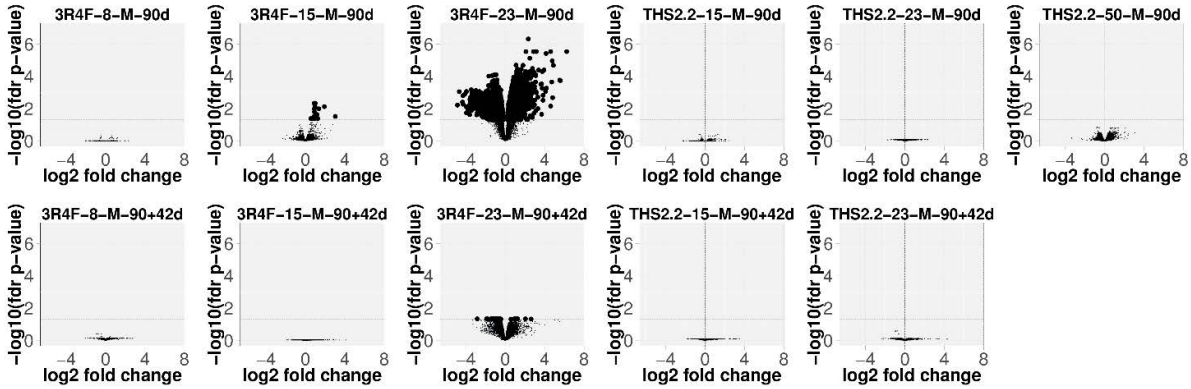


Figure 3: Volcano plots representing the expression profiles of differentially expressed genes in RNE samples

The gene expression change for each gene, calculated as the log2 fold change, is plotted on the x-axis and the statistical significance, proportional to the −log10 of the FDR p-value, is plotted on the y-axis. Black dots indicate genes that are statistically significantly up- or down-regulated (right and left quadrants, respectively) compared with those in sham at the same time point. The exposure vs. sham group comparisons are labeled according to "test item – nicotine concentration – Male/Female – time point". For example, 3R4F-23-M-90d represents the group of male rats exposed to 3R4F at a nicotine concentration of 23 µg/L for the 90d time point and THS2.2-15-M-90+42d represents the group of male rats exposed to THS2.2 at a nicotine concentration of 15 µg/L for the 90+42d time point.

In the RNE of 3R4F-exposed male rats about 3,900 differential expressed genes (DEGs) (fdr-adjusted p-value < 0.05) were detected after 90 days of exposure to the highest concentration of 3R4F (i.e 23 µg/L nicotine) (Figure 3). With the same threshold, no DEGs were found in the 3R4F low group, and only 16 DEGs were found in the 3R4F medium group. The majority of DEGs reverted to levels that were not statistically different from those in the sham group after the 42 days of recovery.

With the same threshold, no DEGs were found in the RNE of male rats in the THS2.2 groups, including the THS2.2 high group (50 µg/L nicotine).

An overview of the numbers of DEGs after 3R4F and THS2.2 exposure for 90 days and recovery for 42 days is displayed in Figure 4.

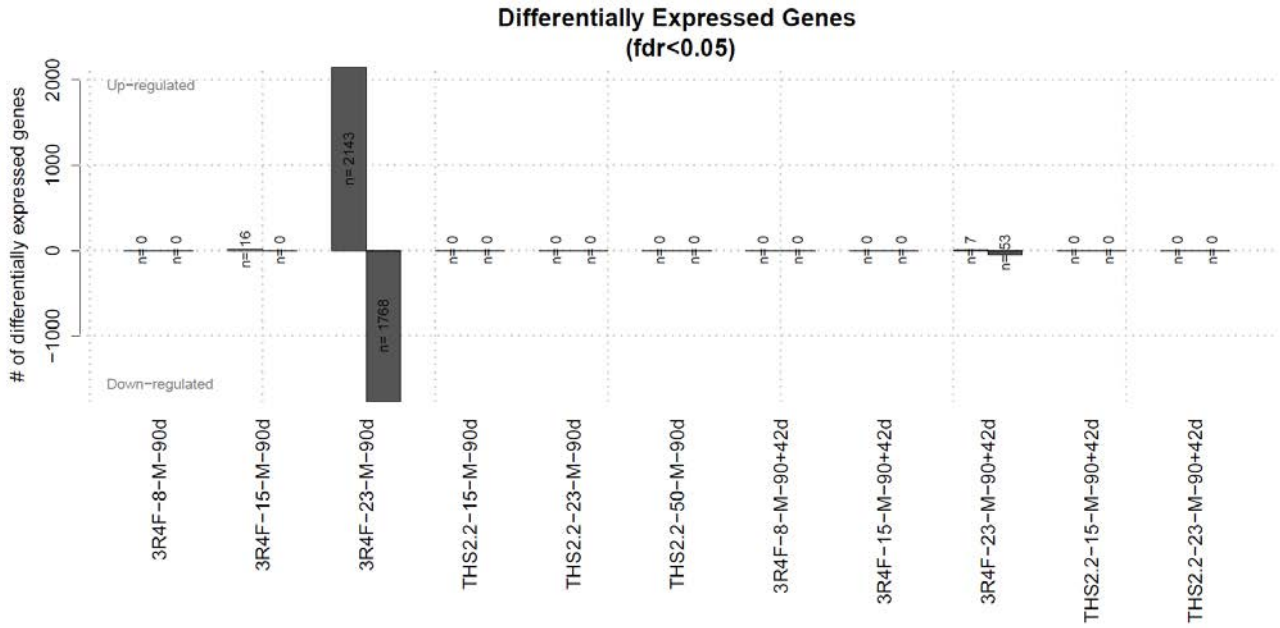


Figure 4: Significantly differentially expressed genes in the RNE of male rats

Bar plots indicate the number of DEGs. See caption of Figure 3 for labeling of the group comparisons.

An overview of the distribution and overlap of DEGs among the different exposure groups is shown as a heatmap in Figure 5.

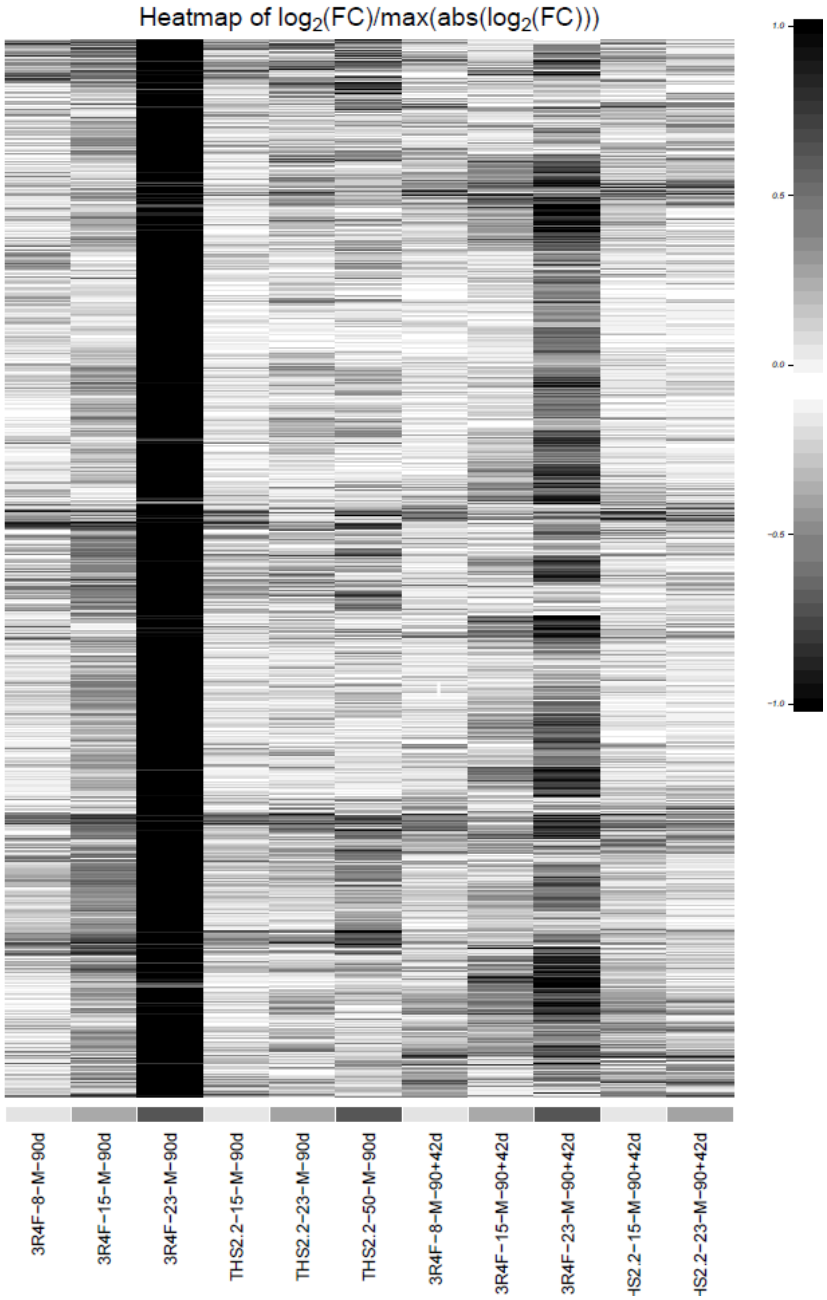


Figure 5: Heatmap of changes in gene expression in RNE

The heatmap represents the log2(fold-change)/max(abs(log2(fold-change))) for all the genes that were differentially expressed in at least one of the comparisons. The grey scaling is based on the absolute values (range -1 to 1). FC = fold change. See caption of Figure 3 for labeling of the group comparisons.

The heatmap shows that differential gene expression in the RNE of male rats increased steeply with increasing 3R4F concentrations. THS2.2 at equivalent nicotine concentrations did not show a strong dose effect.

9.2.2 Biological network perturbation amplitude analysis in RNE after 90 days of exposure to 3R4F and THS2.2, and after a 42-day recovery period

To obtain an overview of which biological networks were impacted by the different exposures, the perturbations of computational networks were calculated. The networks that were investigated were capturing the biological processes cell fate (CFA), tissue repair and angiogenesis (TRA), inflammatory processes (IPN), cell stress (CST), and cellular proliferation (CPR). The biological processes that contributed to network perturbation are shown as starplots in Figure 6.

In the RNE of male rats, the five tested networks appeared to be impacted after exposure of 90 days to the reference cigarettes 3R4F at 23 µg/L nicotine with the inflammatory process network being the most affected. In contrast, only very minor network perturbations were found in the RNE of male rats in all three THS2.2-exposure groups (Figure 6).

The five biological processes were separated into their constituent subprocesses to obtain more detailed mechanistic insights. The subprocesses that showed a significant perturbation in at least one of the exposure groups are presented in Figure 7.

In the RNE from male rats, no subnetworks were impacted by THS2.2 exposure that were not impacted by 3R4F exposure. The most perturbed pathways within the tested THS2.2 concentrations were xenobiotic metabolism response and NFE2L2 signaling, which in amplitude, however, remained well below the effect of 3R4F CS at 23 µg nicotine/L.

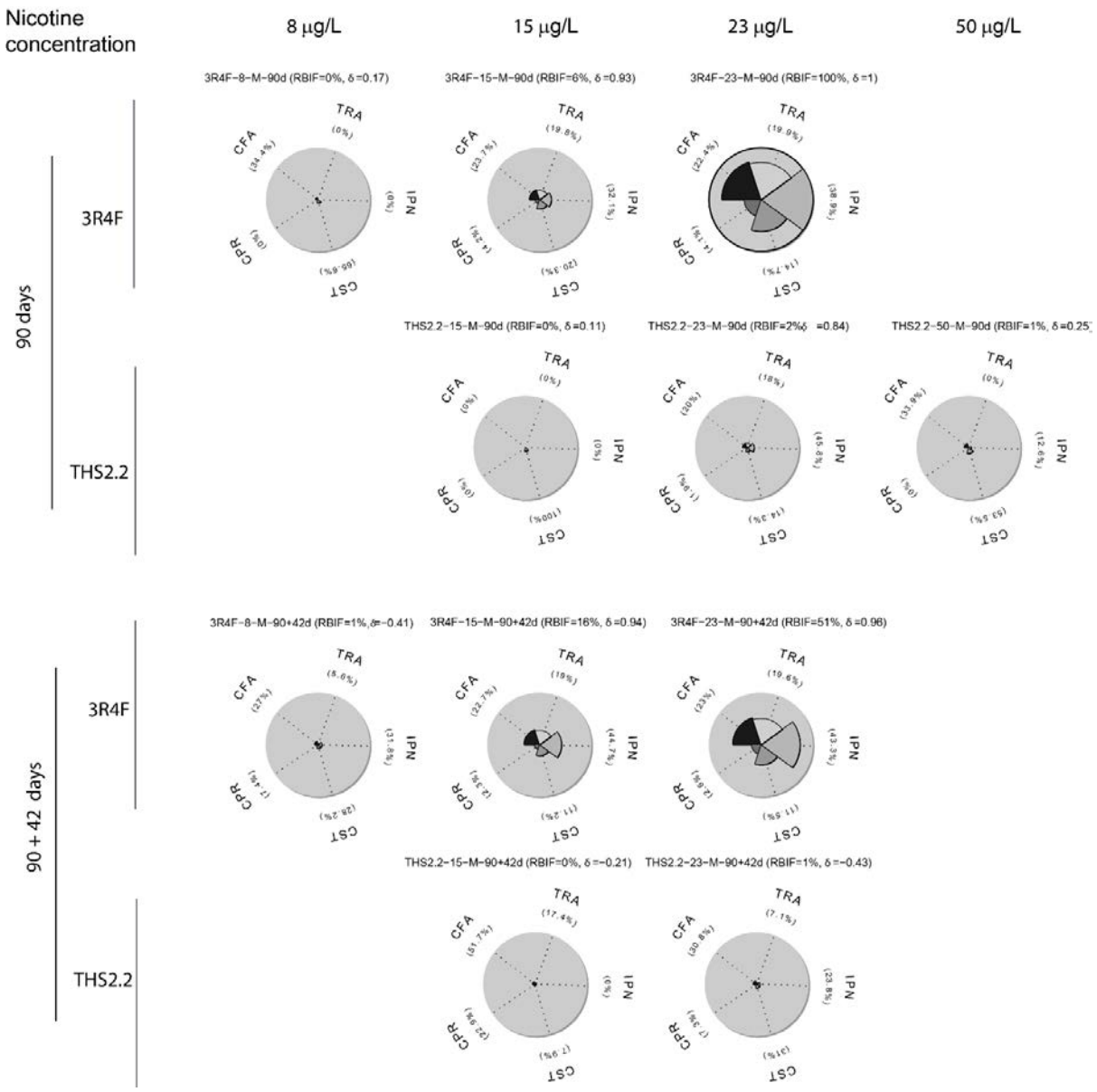


Figure 6: Biological impact of the different exposures in the RNE of male rats displayed as starplots

Perturbation analysis of five biological networks was conducted for each exposure group. The surface area of each slice is proportional to the contribution of a particular treatment to each network perturbation (shown as %). The values are adjusted by the overall relative biological impact factor (BIF) for the treatment compared with the reference such that the sum of the slice areas for each treatment equals the relative BIF for the treatment. The starplots represent the distribution of the sum of contributions for each network across all treatment groups. CFA = cell fate; TRA = tissue repair and angiogenesis; IPN = inflammatory processes; CST = cell stress; and CPR = cellular proliferation. See caption of Figure 3 for labeling of the group comparisons.

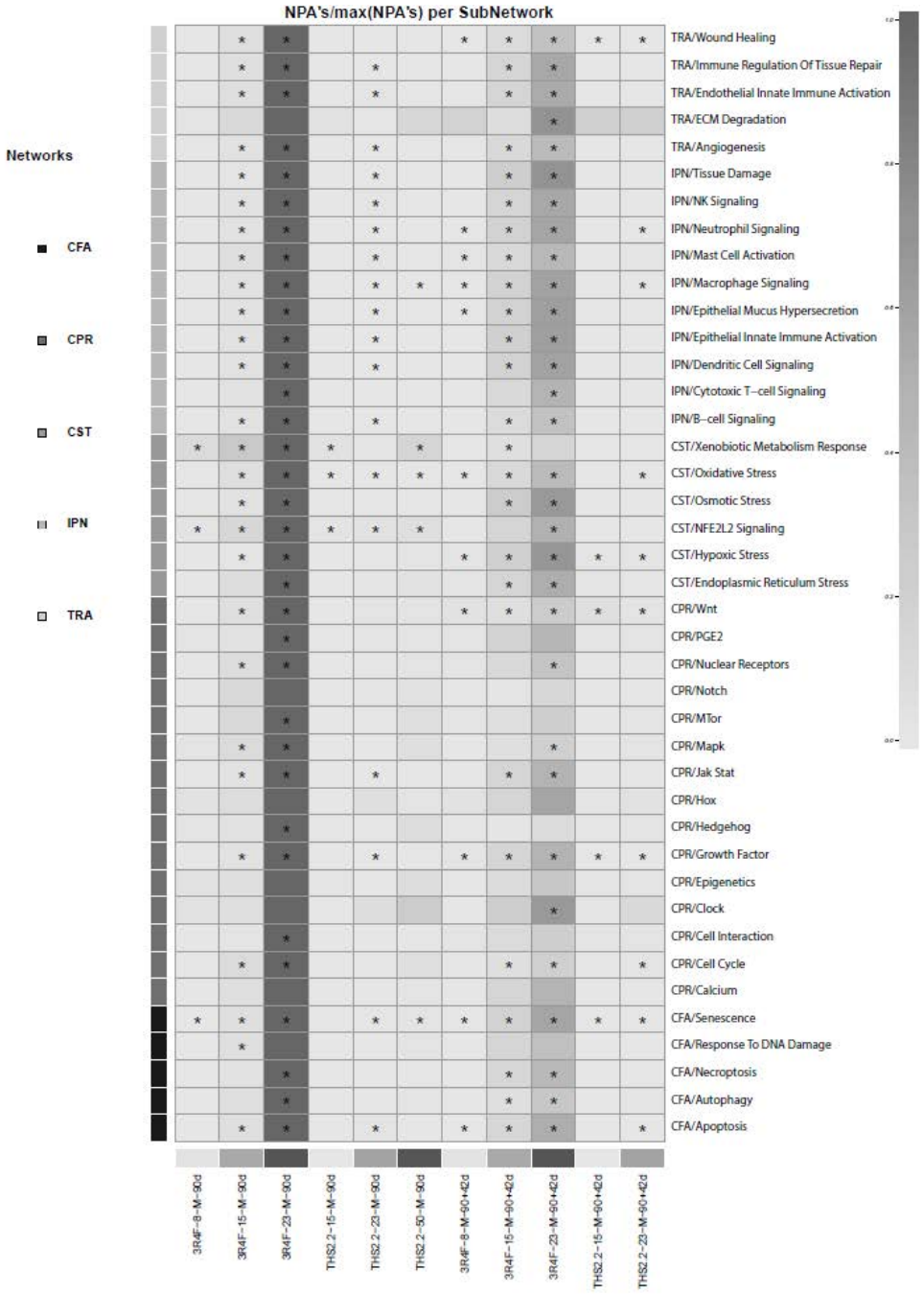


Figure 7: Heatmap of subnetwork perturbations in the RNE of male rats

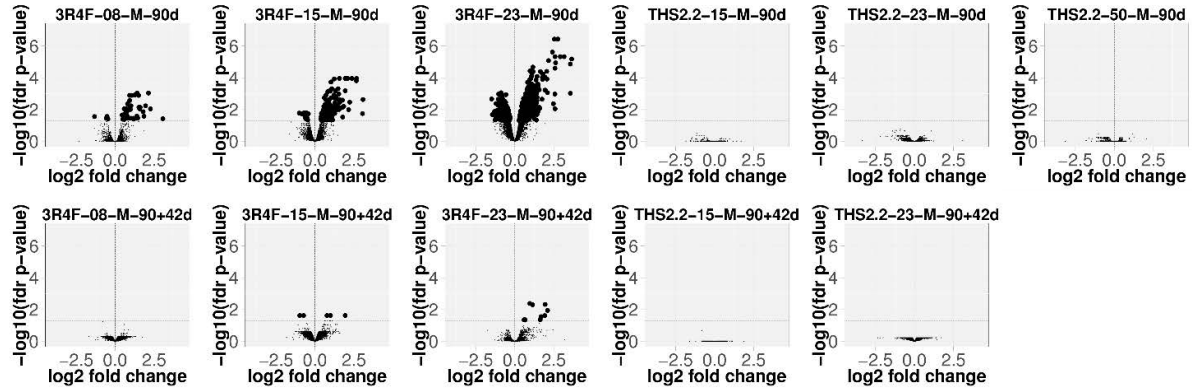
All subnetworks relevant to RNE are shown. A network was considered perturbed if, in addition to the statistical significance of the network perturbation amplitude (NPA) score with respect to the experimental variation, the two companion statistics (O and K) derived to determine the specificity of the NPA score with respect to the biology described in the network are significant. NPA score ranges from 0 to 1. * indicates O and K statistic p-values below 0.05 and significant different with respect to experimental variation. See caption of Figure 3 for labeling of the group comparisons.

9.3 Lung Transcriptome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats

9.3.1 Differentially expressed genes in the lung transcriptome after 90 days of exposure to 3R4F or THS2.2, or after a 42-day recovery period

The amplitude and statistical significance of gene expression changes in lung tissue for 3R4F and THS2.2-exposed rats after 90 days of exposure or after a 42-day recovery period were compared with the sham group and visualized as volcano plots (Figure 8).

A)



B)

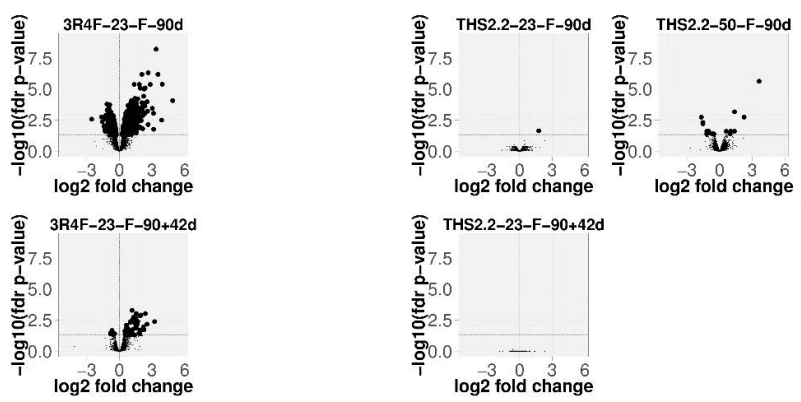


Figure 8: Volcano plots showing the systems response profiles in lung transcriptome

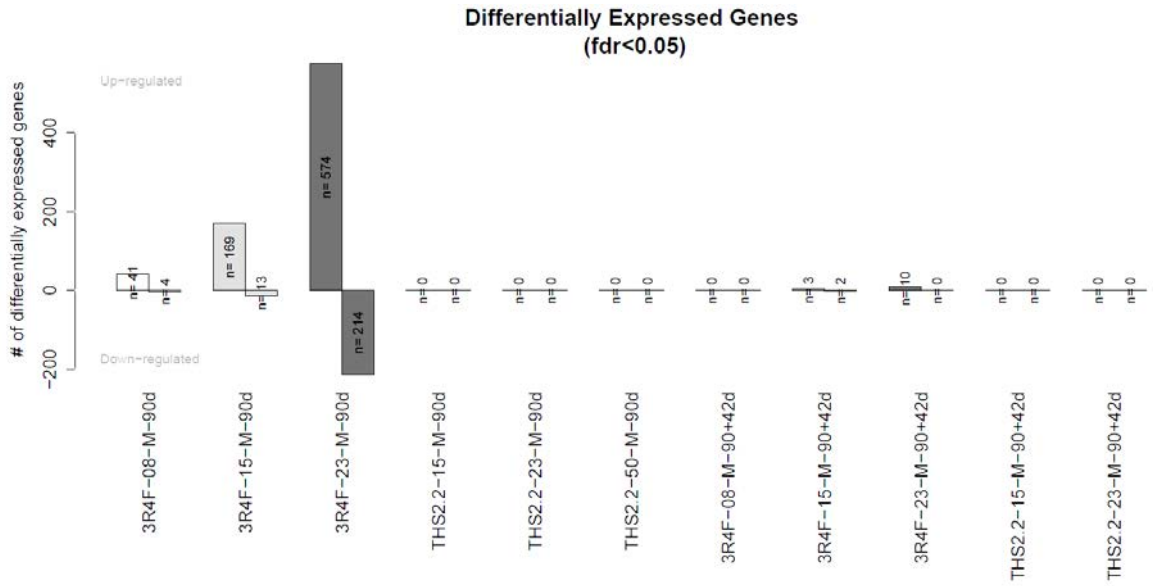
The expression change for each gene, calculated as the log2 fold change, was plotted on the x-axis and the statistical significance, proportional to −log10 of the FDR p-value, was plotted on the y-axis. Dark grey dots indicate genes that are statistically significantly up- or down-regulated (right and left quadrants, respectively) compared with sham at the same time point. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

A dose-dependent effect was found in the lung tissue of 3R4F-exposed male rats. At an adjusted p-value < 0.05 threshold, a total of 788 DEGs were found in the 3R4F high group (23 µg/l nicotine) compared to sham exposed rats. In contrast, at the same threshold, no DEGs were found in any of the three THS2.2-exposed groups (Figure 9 A).

A total of 1,302 DEGs were found in the lung tissue of female rats exposed for 90 days to 3R4F at a concentration of 23 µg/l nicotine. At the same nicotine concentration, one DEG was found in the THS2.2-exposed rats, and 17 DEGs were found after a 90 days exposure to the high concentration of THS2.2 (50 µg/L nicotine).

After the 42-day recovery period from the 90-day 3R4F exposure, more than 90% of the DEGs in the male and female lung tissues were not significantly different from sham levels anymore. After the 42-day of recovery period from the 90-day THS2.2 exposure, none of the DEGs in the male and female lung tissues was significantly different from sham levels anymore. The numbers of up- and down-regulated genes are displayed in Figure 9.

A)



(Figure continues)

B)

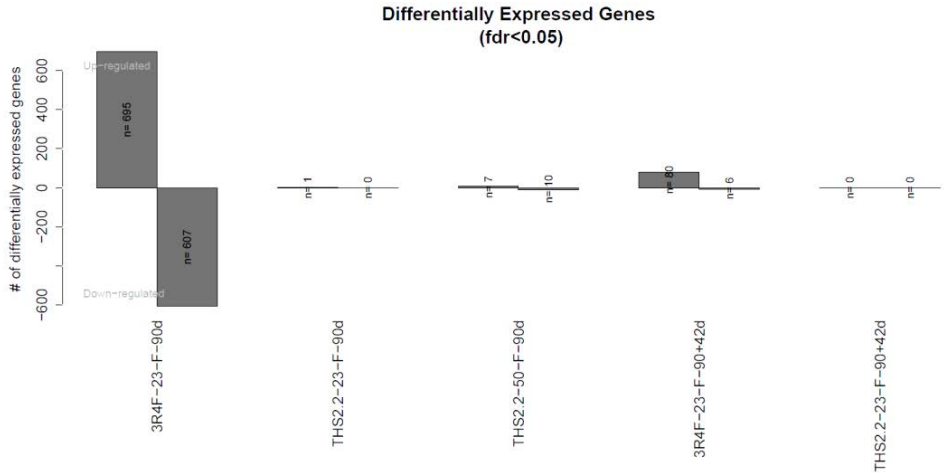
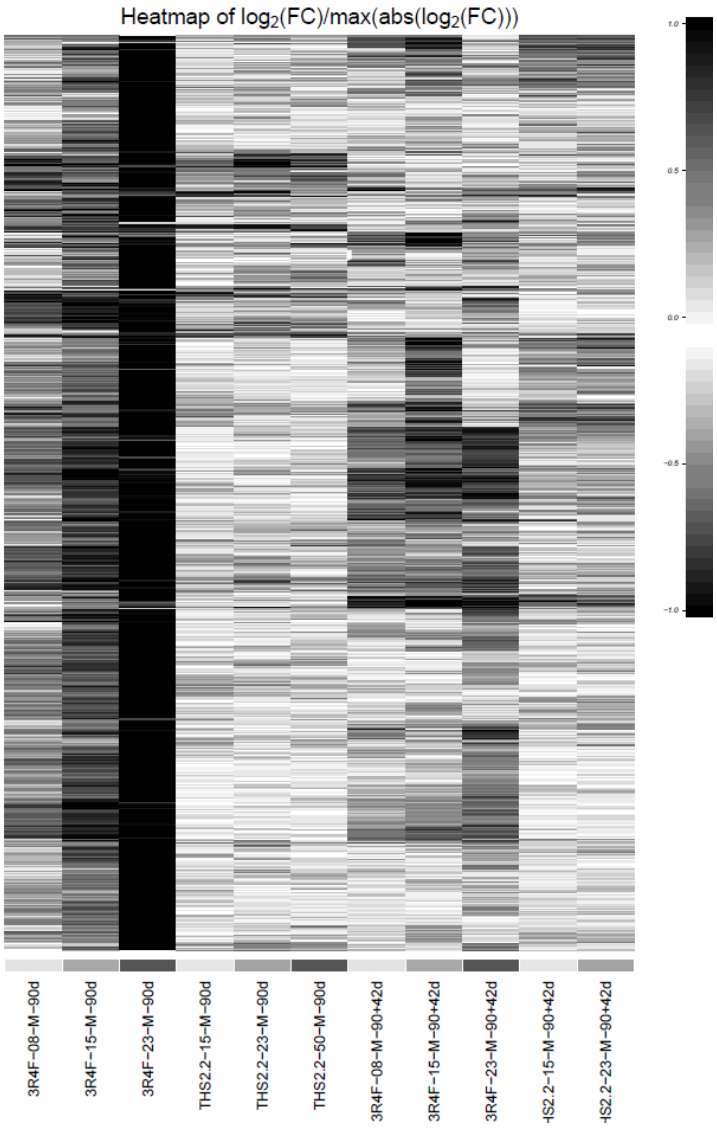


Figure 9: Significantly differentially expressed genes in lung tissue of male and female rats.

Bar plots indicate the number of DEGs with FDR <0.05. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

Heatmaps were generated to obtain an overview of the response profiles of the DEGs among the different exposure groups (Figure 10).

A)



(Figure continues)

B)

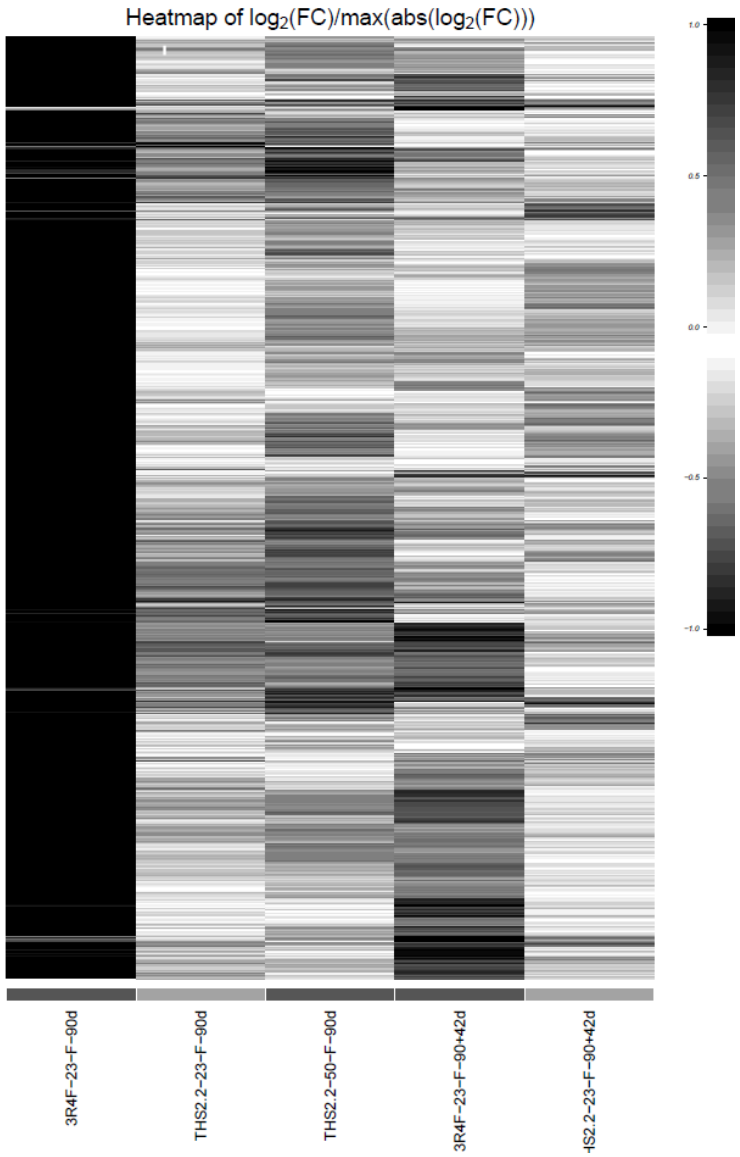


Figure 10: Heatmap of changes in gene expression in lung tissue

The heatmaps represent the log2(fold-change)/max(abs(log2(fold-change))) for all the genes that were differentially expressed in at least one of the comparisons.. The grey scaling is based on the absolute values (range from -1 to 1). A) Male rats, B) Female rats. FC = fold change. See caption of Figure 3 for labeling of the group comparisons.

The heatmaps show that the overall fold change of the DEGs was relatively lower in the THS2.2-exposed groups compared with the 3R4F-exposed groups.

Of the 17 DEGs detected in the lung of female rats after exposure to the highest concentration of THS2.2, 14 were also differentially expressed in the 3R4F-exposed groups.

The genes that were significantly differentially expressed only after THS2.2 exposure are listed in Table 8. Note that, generally, these genes show the same trend upon 3R4F exposure.

Table 8: "Unique" genes that were differentially expressed only in lung of female rats after THS2.2 exposure.

Gene Name	3R4F high - 23- 90 d	THS2.2 med-23-90 d	THS2.2 high-50 - 90d	3R4F high 23-90+42 d	THS2.2 med-23 90+42 d
Ifi27l2b	-0.52256	-0.56221	-1.00 *	-0.18299	-0.28167
Isg15	-1.0175	-0.67045	-1.181 *	-0.43809	-0.26863
Mx1	-1.23024	-0.94784	-1.678 *	-0.50426	-0.23964

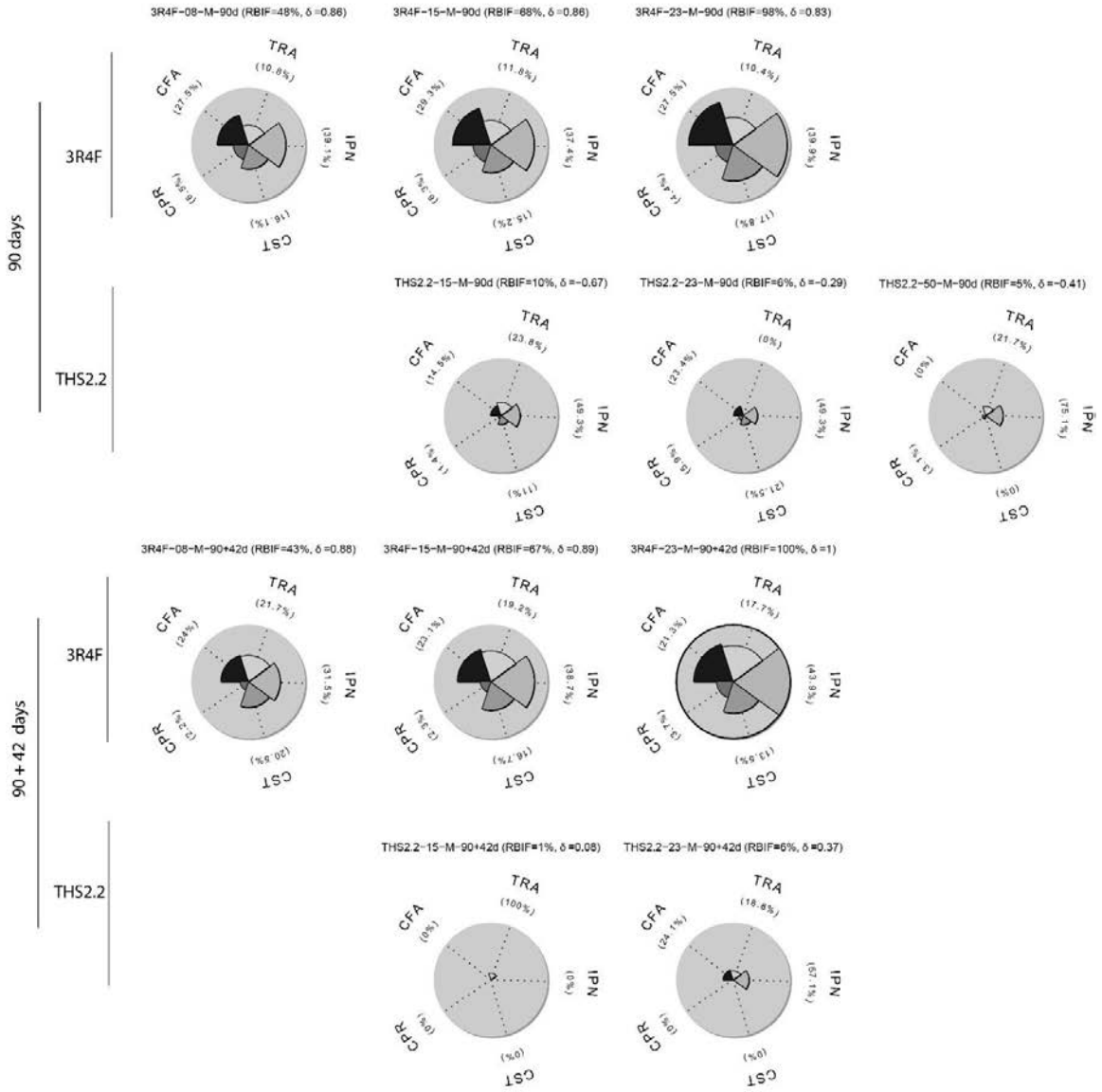
These DEGs were found only in the lung tissue of female rats exposed to THS2.2 aerosol but not those exposed to 3R4F CS. Unique refers to those three DEGs that had a FDR <0.05 only in the THS2.2 groups of female rats and not in any of the other groups tested. The log2 fold-changes for the comparisons with the respective sham groups are shown, statistical significance (FDR < 0.05) is indicated with an asterix ("*").

9.3.2 Biological network perturbation amplitude analysis in RNE after 90-days exposure to 3R4F and THS2.2 or after a 42-day recovery period

To determine which biological networks were impacted by the different exposures, the perturbations of computational networks were calculated. The networks that were investigated were cell fate (CFA), tissue repair and angiogenesis (TRA), inflammatory processes (IPN), cell stress (CST), and cellular proliferation (CPR).

The biological processes that contributed to network perturbation are shown as starplots in Figure 11.

A)



(Figure continues)

B)

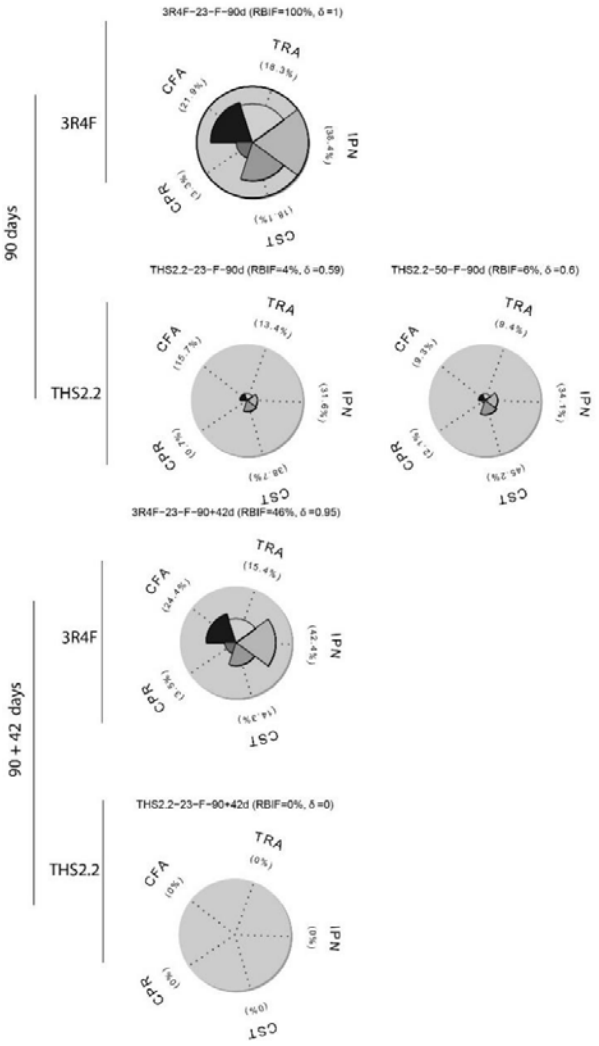


Figure 11: Biological impact of the different exposures in lung tissue of male and female rats displayed as starplots

Perturbation analysis of five biological networks was conducted for each exposure group. The surface area of each slice is proportional to the contribution of a particular treatment to each network perturbation (shown as %). The values are adjusted by the overall relative biological impact factor (BIF) for the treatment compared with the reference such that the sum of the slice areas for each treatment equals the relative BIF for the treatment. The starplots represent the distribution of the sum of contributions for each network across all treatment groups. CFA = cell fate; TRA = tissue repair and angiogenesis; IPN = inflammatory processes; CST = cell stress; and CPR = cellular proliferation. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

In the lung tissue of male rats, all the tested networks appeared to be impacted after exposure of 90 days to 3R4F at the three tested nicotine concentrations (8, 15, and –23 µg/L). The inflammatory process network gave the highest contribution to overall perturbation score. After 90 days of exposure to the low THS2.2 concentration (15 µg/L nicotine) all networks except the proliferation network were perturbed, but with much lower amplitudes than for 3R4F CS exposure at matching nicotine concentration. After 90 days exposure to the medium concentration of THS2.2 (23 µg/L nicotine), perturbations of the cell stress network, cell fate, and inflammatory processes were found, again with much reduced amplitudes than for 3R4F CS exposure at the same nicotine concentration. After 90 days exposure to the high concentration of THS2.2 (50 µg/L nicotine) perturbation of the inflammatory process network was accompanied by only a small perturbation of the tissue repair and angiogenesis network.

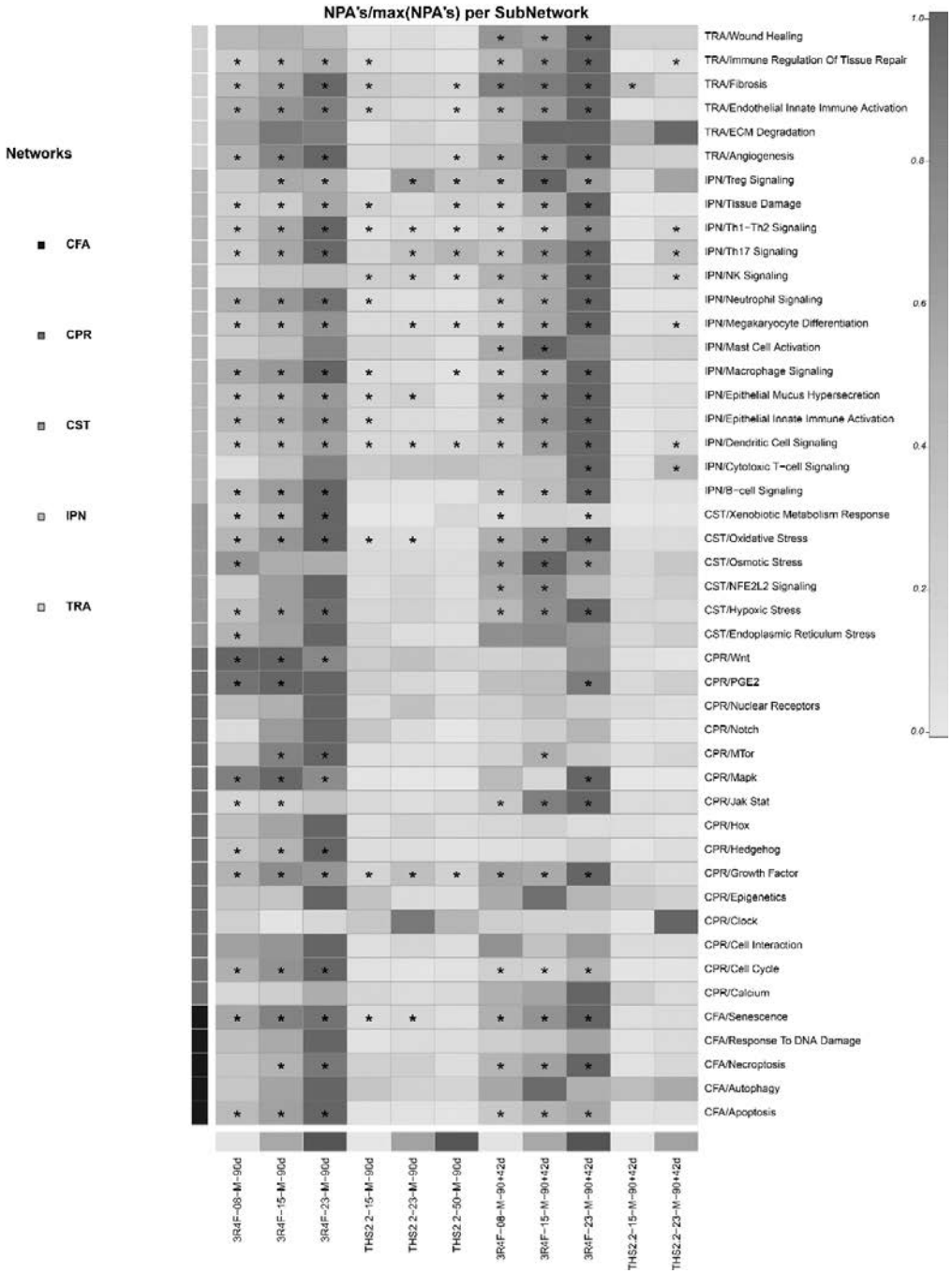
After the 42 days recovery period, no recovery was detected in the lung tissue of male rats in the starplot representations (Figure 11). The perturbation of the inflammatory process network after THS2.2 exposure also seemed not to revert in the 42 days recovery period. Of note, in this kind of analysis, all the genes were used regardless of whether they were above or below the significance threshold (see 8.2). The rationale behind this approach was that non-significant but differentially regulated genes could create a significant effect if they acted in the same pathways. Further it should be noted that in these representations of the network perturbation no directionality can be seen.

In the lung tissue of female rats after 90 days of exposure to 3R4F CS, the inflammatory process network again had the highest contribution, followed by cell fate, tissue repair and angiogenesis, and cell stress networks. After 90 days of exposure to THS2.2, the cell stress and the inflammatory process networks also showed perturbation. However, the amplitude of the perturbation was lower than that for the 3R4F exposure.

After 42 days of recovery from 3R4F exposure, the pattern of perturbed networks remained the same, but the amplitude of the perturbations was lower. After 42 days of recovery from THS2.2 exposure, no network perturbation was detected, indicating full recovery.

The five biological processes were separated into their constituent subprocesses to obtain more detailed mechanistic insights. The subnetworks that showed a significant perturbation in at least one of the exposure groups are presented in Figure 12.

A)



(Figure continues)

B)

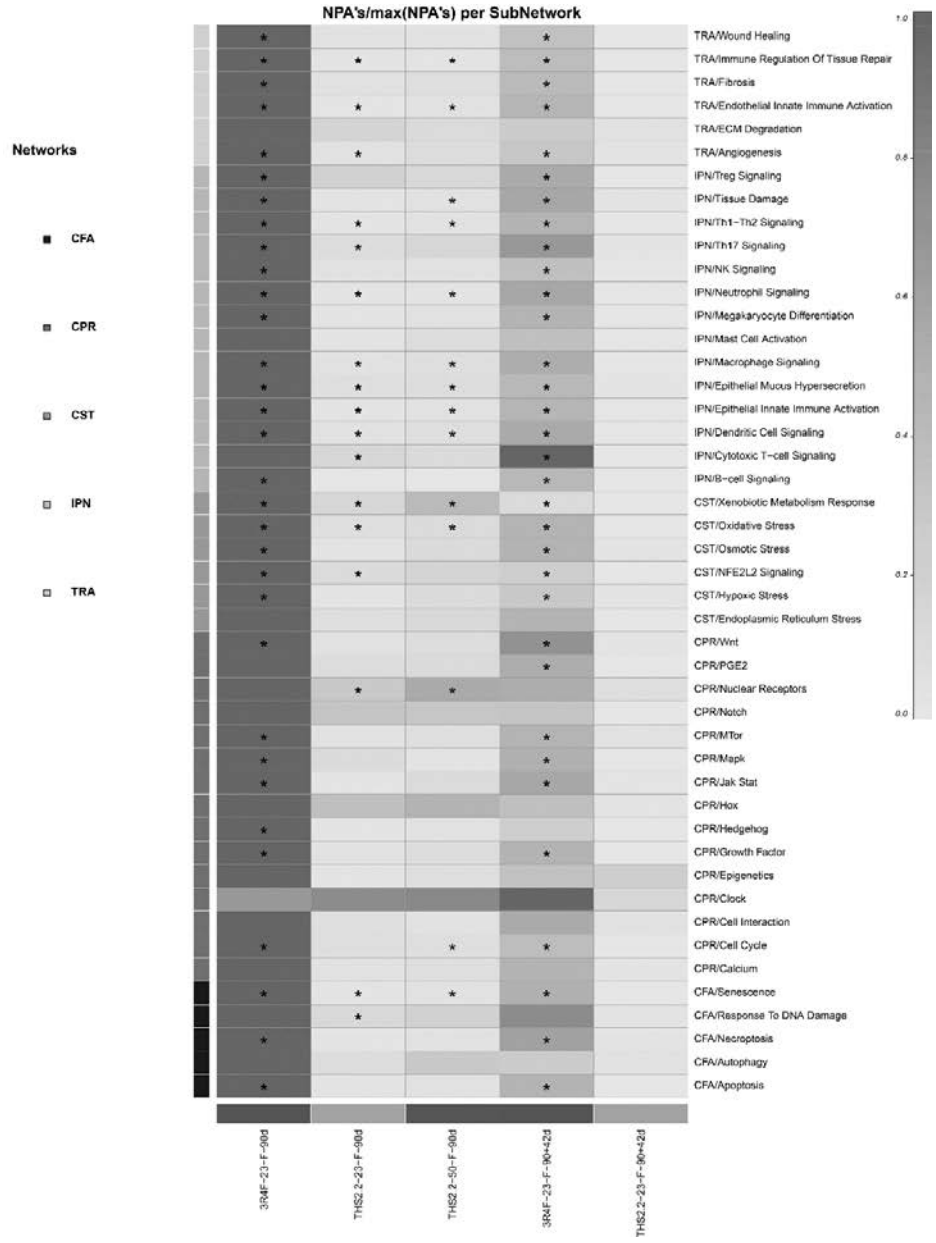


Figure 12: Heatmap of subnetwork perturbations in lung tissue of male and female rats

All subnetworks relevant to lung tissue are shown. A network is considered as perturbed if, in addition to the statistical significance of the network perturbation amplitude (NPA) score with respect to the experimental variation, the two companion statistics (O and K), derived to determine the specificity of the NPA score with respect to the biology described in the network, are significant. * indicates O and K statistic p-values below 0.05 and statistical significance with respect to the experimental variation**.** See caption of Figure 3 for labeling of the group comparisons. (A) Males, (B) Females

In general, the subnetworks in the lung tissue of male rats after 3R4F exposure were more strongly perturbed than after THS2.2 exposure, as indicated by the shading in Figure 12. The Natural killer (NK) cell subnetwork within the inflammatory process network was impacted only after 90 days of THS2.2 exposure and not after 90 days of 3R4F exposure. Interestingly, this network was perturbed in the 3R4F 90 days +42 days recovery group.

Similarly, all the subnetworks in the lung tissue of female rats were more perturbed after 3R4F exposure than after THS2.2 exposure. The nuclear receptor signaling subnetwork within the cell proliferation network was significantly perturbed only after THS2.2 exposure (23 and 50 µg/L) and not after 3R4F exposure.

9.4 Lung Proteome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats

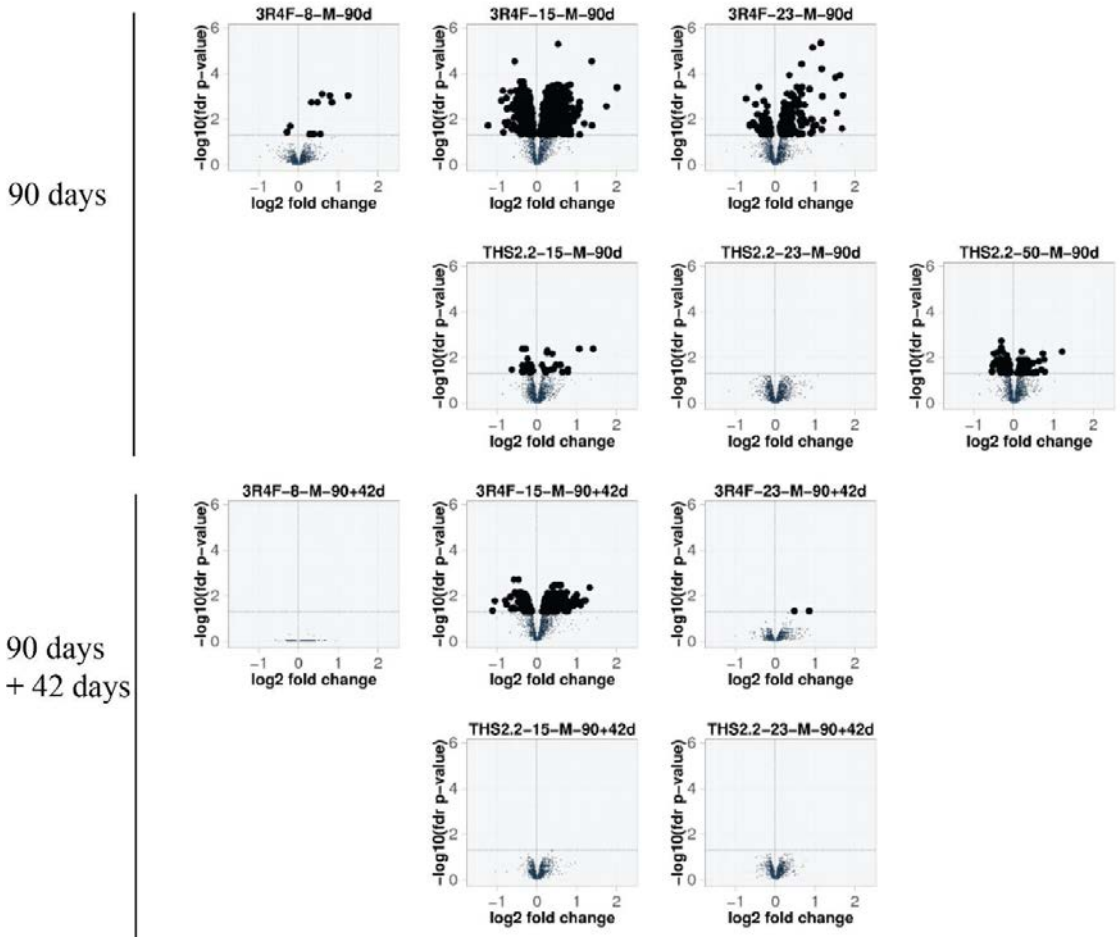
An iTRAQ LC-MS/MS-based quantitative proteomics approach was used to assess how the different exposure conditions quantitatively affected the lung proteome. This isobaric tag-based method allows multiplexed analysis of up to eight samples by making use of up to eight different peptide-reporter labeling reagents. Here, three multiplexed iTRAQ labeling sets were defined: iTRAQ set 1, which focused on the exposure and recovery effects for the female lung samples; and iTRAQ sets 2 and 3, which contained the male lung sample groups for the 90 d and 90+42 d time points, respectively (Table 5).

[Note this data set is the first iTRAQ analysis performed for rat samples at PMI and should be evaluated in this context and the context of other related endpoints; e.g., the transcriptomics analysis of the lung tissue.]

9.4.1 Differentially expressed proteins in lung tissue

To evaluate the global impact on the proteome of each exposure condition compared with sham exposure, the amplitude and statistical significance of the differentially expressed proteins were visualized by volcano plots (Figure 13). The numbers of differentially expressed proteins identified across the exposure conditions are summarized in Figure 14.

A)



(Figure continues)

B)

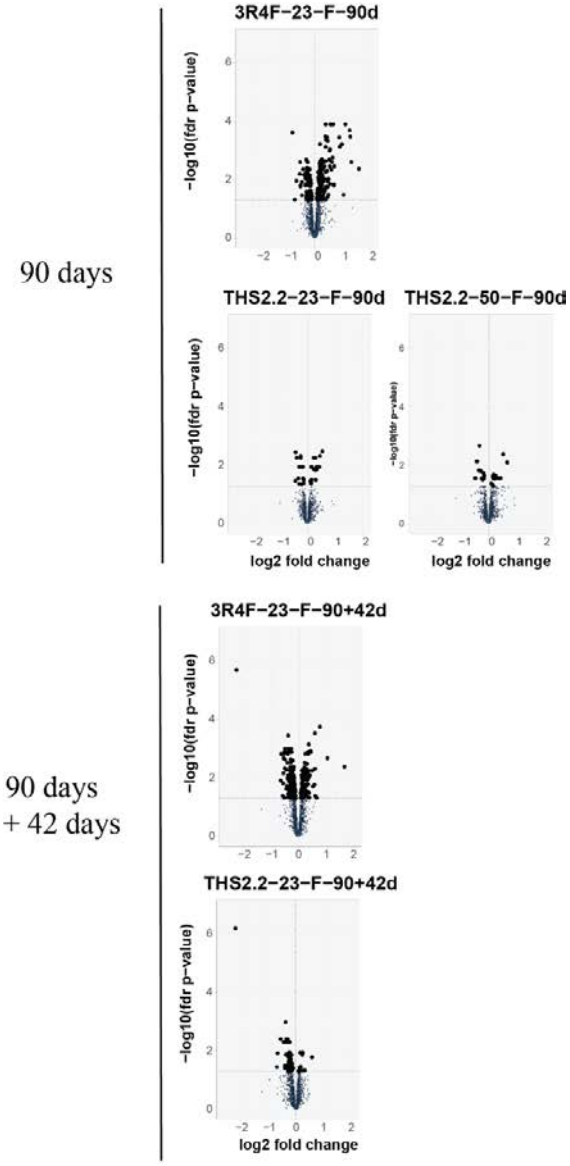


Figure 13: Volcano plots representing proteome response profiles in lung

The protein expression change for each protein, calculated as the $\log_2$ fold change, is plotted on the x-axis and the statistical significance, proportional to the $-\log_{10}$ of the adjusted p-value, is plotted on the y-axis. Bold, black dots indicate proteins that are statistically significantly up- or down-regulated (right and left quadrants, respectively) compared with the corresponding sham group at each time point (BH adjusted p-value <0.05). A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

A proteome response was detected in the lung tissue of 3R4F low, medium, and high (i.e., 8, 15, 23 µg/L) exposed male rats after 90 days of exposure. Notably, no consistent dose effect was observed. However, while, in total, fewer differentially expressed proteins were observed with the 3R4F high dose than with the 3R4F medium dose, several affected proteins showed high fold changes (Figure 13A). In contrast, the response to the 90-day THS2.2 exposures resulted in fewer significantly differentially expressed proteins than was observed upon 3R4F exposure at corresponding nicotine concentrations.

After the 42-day recovery phase, nearly all the differentially expressed proteins in the 3R4F high group recovered to sham levels. However, the 3R4F medium group showed only a partial recovery of the differential protein expression response. In addition, after 42 days of recovery following exposure to the THS2.2 low and medium doses (15 and 23 µg/L nicotine), nearly all differently expressed proteins reverted to the sham level (Figure 13A).

The total number of significantly up- or down-regulated proteins for all treatment groups at the 90-day and 90+42-day time points are summarized in Figure 14. As also evident from the volcano plots, there was a clear decrease in the number of differentially expressed proteins for all concentrations of 3R4F when the 90-day exposure period was compared with the additional 42-day recovery period. At the 90-day exposure time point, 12 differentially expressed proteins were identified in the 3R4F low (8 µg/L nicotine) group and 521 differential expressed proteins were identified in 3R4F medium (15 µg/L nicotine) group. However, in the 3R4F high (23 µg/L nicotine) group, the number of differentially expressed proteins decreased to 183, but still showed a strong increase compared with the 3R4F low group.

There was a clear decrease in the number of differentially expressed proteins in the THS2.2 (all treatment concentration) groups compared with the numbers in the 3R4F (all treatment concentration) groups at both time points (90 d and 90+42 d). Thus, THS2.2 demonstrated a consistently lower effect on the differential expression of proteins in the lung proteome compared with 3R4F.

A) **B)**

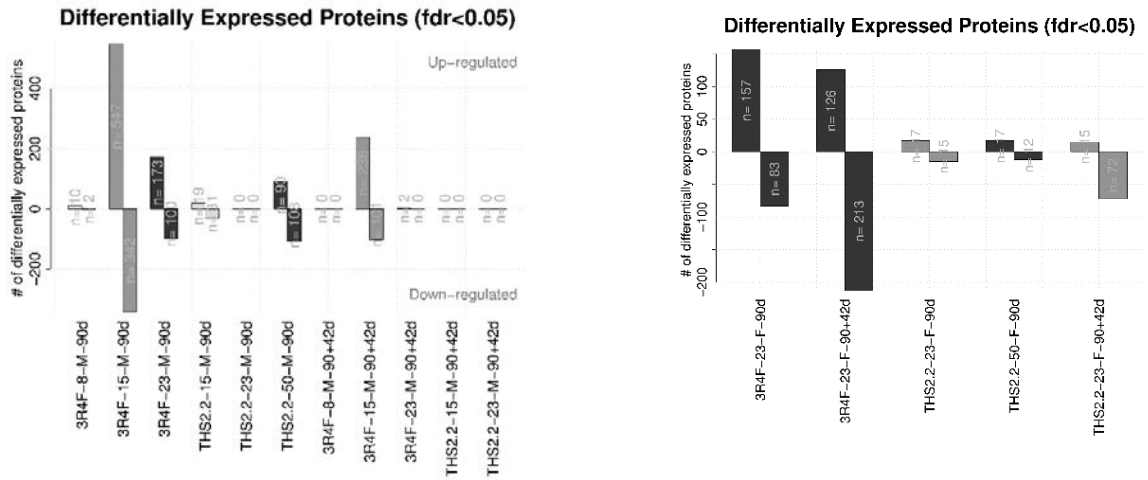


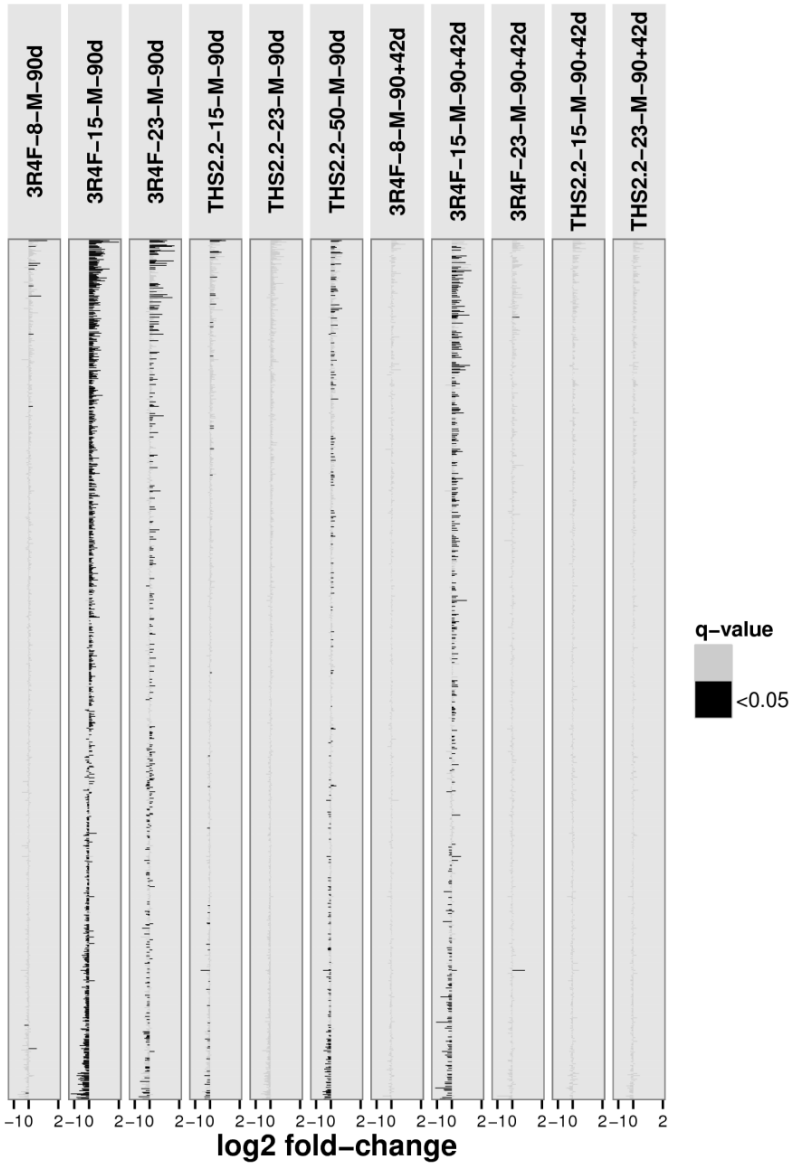
Figure 14: Significantly differentially expressed proteins in lung of male and female rats

A) Male rats at the 90-day and 90+42-day time points, B) Female rats at the 90-day and 90+42-day time points. The comparisons were made against the corresponding sham group. The numbers of up- and down-regulated proteins are shown within the bars. The BH-adjusted FDR was <0.5%. See caption of Figure 3 for labeling of the group comparisons.

In overall agreement with the male results, the volcano plots for the female data showed also a reduced response to THS2.2 exposure of the lung tissue proteome (Figure 13 B). While the female rats showed a similar response as the male rats to the high 3R4F concentration (judged by the number of differentially expressed proteins), no recovery effect was observed for the females after the 42-day recovery period (Figure 14 B). Similarly, in the THS2.2 medium group (23 µg/L nicotine) the female rats showed a higher than expected number of down-regulated proteins compared with the sham group at the 90+42-day time point. Although no clear biological or technical explanation for these results could be found, the separation of the 90+42-day sham group in the principal component analysis (PCA) score plot (see section 7.10.4) suggested that expression effects specific to the 90+42-day sham group rather than the other 90+42-day exposure groups may explain this effect. In this context, note that this study was the first at PMI in which the iTRAQ approach was used for rat samples and it is suggested to evaluate this data in this context and in the context of the other measured endpoints of this study, especially the transcriptomics analysis.

To compare the exposure response more directly across conditions, the fold-change and statistical significance for each differentially expressed protein under all the conditions are shown in a vertical bar chart (Figure 15). This visualization revealed the overall similarity of the exposure response in the three 3R4F groups and showed that for the majority of proteins the directionally was the same. A reduced exposure response was seen in the three THS2.2 groups compared with the response in the 3R4F groups.

A)



(Figure continues)

B)

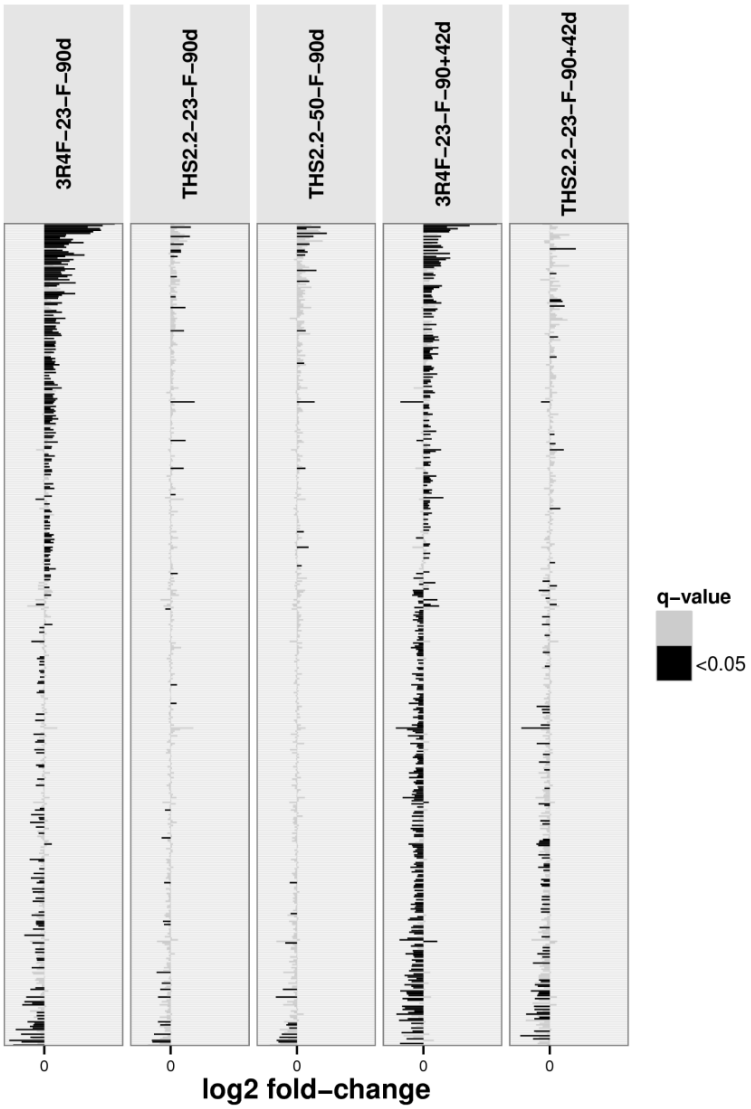


Figure 15: Fold-change and significance of the differentially expressed proteins across the exposure conditions

The bar plots show the differential expression of proteins significantly affected by any of the exposure conditions. Treatment groups were compared with the corresponding sham group. The length of the bars indicates the $\log_2$ fold-change (right = up-regulated, left = down-regulated). Only significantly differentially expressed proteins (BH-adjusted p-value <0.05, q-value) in any of the comparisons are shown. (A) Male rats, (B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

Table 9: "Unique" proteins that were differentially expressed only after THS2.2 treatment

A)

Symbol	Uniprot ID	3R4F-8-M-90d	3R4F-15-M-90d	3R4F-23-M-90d	THS2.2-15-M-90d	THS2.2-23-M-90d	THS2.2-50-M-90d	3R4F-8-M-90+42d	3R4F-15-M-90+42d	3R4F-23-M-90+42d	THS2.2-15-M-90+42d	THS2.2-23-M-90+42d
Ca1	B0BNN3	-0.51	-0.04	-0.24	-0.54	-0.49	-0.55 *	0.16	-0.06	-0.06	0.18	-0.28
Itih1	B2RYM3	-0.22	-0.22	-0.22	-0.19	-0.16	-0.31 *	0.49	0.15	0.17	0.23	0.1
Dhx15	D3ZD97	0.11	0.08	0.01	0.01	0.06	0.12 *	-0.06	0.03	-0.11	-0.02	-0.1
Rit1	D3ZJH4	0.13	0.3	-0.19	0.4	0.2	0.49 *	NA	NA	NA	NA	NA
Phldb1	D3ZNS1	0.04	0.28	0.1	0.14	0.22	0.32 *	NA	NA	NA	NA	NA
Trove2	D3ZRN5	0.03	0.14	0.07	0.12	0.17	0.29 *	NA	NA	NA	NA	NA
Napg	D4A0E2	0.07	0.11	-0.02	0.13	0.16	0.16 *	-0.01	0.08	-0.04	0	0.03
Smtn	D4ABA5	0.04	-0.01	0.27	0.16	0.08	0.34 *	NA	NA	NA	NA	NA
Fnbp1	F1LZF2	-0.04	0.08	0.13	0.3	0.3	0.18 *	NA	-0.15	-0.13	NA	NA
Prkca	F1M2P8	-0.05	0.19	-0.1	0.08	0.2	0.21 *	NA	NA	NA	NA	NA
C7	F1M7F8	-0.07	0.04	0.09	-0.15	-0.06	-0.17 *	0.04	0.3	0.01	0.07	0.02
Isoc1	F2Z3T7	-0.02	0.09	0.07	0.11	0.16	0.19 *	0.08	-0.13	0.09	0.11	0.15
Gna11	G3V6Q6	0.12	-0.1	0.16	0.18	0.17	0.23 *	0.01	-0.08	0	-0.08	0.03
Pdcd6	G3V7W1	0.08	0.01	0.09	0.08	0.07	0.16 *	-0.01	0.03	0.06	-0.03	-0.04
Pabpn1	G3V7Z8	0.08	0.1	-0.12	-0.16 *	-0.14	-0.16 *	0	-0.1	-0.04	-0.07	-0.02
Usp4	M0R851	-0.04	0.17	-0.08	0.03	0.16	0.19 *	NA	NA	NA	NA	NA
Eng	M0RA19	0.04	0.09	0.04	0.03	0.1	0.16 *	0.13	0.13	0.11	0.17	0.19
Pdia3	P11598	0	-0.08	0.11	-0.06	-0.08	-0.13 *	-0.01	-0.09	0.05	-0.03	-0.04
A1i3	P14046	-0.25	-0.2	-0.25	-0.35	-0.3	-0.45 *	0.12	0.02	0.01	0.11	-0.15
Lamp1	P14562	-0.02	-0.06	0.08	-0.25 *	-0.2	-0.23 *	-0.04	-0.01	0.04	-0.17	-0.12
Nes	P21263	0.23	0.15	-0.07	0.31	0.3	0.35 *	-0.11	-0.43	-0.05	0.4	0.32
Ahsg	P24090	-0.35	-0.22	-0.2	-0.21	-0.23	-0.33 *	0.08	-0.03	-0.02	0.07	-0.08
Renbp	P51607	-0.03	-0.21	-0.26	-0.3	-0.23	-0.36 *	NA	NA	NA	NA	NA
Pdlim1	P52944	0.05	0.02	0.06	0.01	0.05	0.08 *	0.01	-0.01	-0.01	0.03	0.01
Arl1	P61212	-0.07	-0.27	0.07	-0.07	-0.02	-0.22 *	-0.05	-0.07	0.03	-0.05	-0.06
Eef1a2	P62632	-0.06	-0.24	-0.37	-0.37	-0.24	-0.41 *	NA	NA	NA	NA	NA
Ube2i	P63281	0.01	-0.09	-0.08	-0.12	0.02	-0.21 *	0.04	-0.03	-0.06	-0.04	0.07
Blmh	P70645	-0.02	-0.14	-0.02	-0.06	-0.08	-0.12 *	-0.01	-0.1	0	-0.04	-0.08
Cox7c	P80432	0.04	-0.08	0.1	-0.18 *	-0.13	-0.19 *	0.06	0.07	0.05	-0.07	-0.02
Arse	Q32KK0	-0.34	0.69	-0.18	0.29	0.15	0.76 *	-0.06	0.46	-0.02	0.27	0.2
Pspc1	Q4KLH4	0.06	0.05	0.09	0.09	0.08	0.13 *	0	0.05	0.04	0.05	0.15
Thumpd1	Q5M943	0.06	0.05	-0.07	0.12	0.18	0.16 *	NA	NA	NA	NA	NA
Arhgap29	Q5PQJ5	-0.03	-0.04	-0.17	-0.02	-0.07	-0.19 *	NA	NA	NA	NA	NA
Nap1l4	Q5U2Z3	0	-0.02	-0.02	0.07	0.12	0.1 *	0.03	0.13	-0.01	0.11	0.08
Fip1l1	Q5U317	0.1	-0.04	0.24	0.86	0.51	0.54 *	NA	NA	NA	NA	NA
Sart1	Q5XIW8	0.11	0.22	-0.04	0.04	0.15	0.24 *	0.65	0.89	0.38	0.88	0.7
Nedd4	Q62940	0.05	-0.01	0.04	0.14	0.14	0.23 *	0	0.16	0.01	0.09	0.16
Pon3	Q68FP2	-0.01	-0.11	-0.14	-0.09	-0.06	-0.14 *	0	-0.01	0.06	-0.03	-0.03
Vps26a	Q6AY86	0.07	0.06	0.04	0.13	0.14	0.16 *	0.06	0.11	0.02	0.01	0.08
Tap2	Q6MGA3	-0.04	-0.04	-0.11	-0.07	-0.03	-0.14 *	NA	NA	NA	NA	NA
Bre	Q6P7Q1	-0.04	0.08	0.08	-0.11	-0.09	-0.24 *	0.01	0.17	0.08	-0.16	-0.04
Slc7a10	Q75T81	0.1	0.32	0.07	0.08	0.28	0.27 *	NA	NA	NA	NA	NA
Lin7c	Q792I0	0.03	0.07	0	0.09	0.13	0.11 *	-0.02	0.11	-0.02	0.05	0.05
Hsbp1	Q8K3X8	-0.01	-0.25	0.01	-0.07	-0.15	-0.2 *	0.05	-0.05	0.05	-0.03	0
LOC100909593	Q95574	-0.16	0.08	-0.01	-0.28	-0.55	-0.49 *	NA	NA	NA	NA	NA
Ilf3	Q9JIL3	0.11	0.08	-0.08	0.18 *	0.11	0.17 *	0.03	0.24	0	0.09	0.12
Uqcrb	B2RYS2	0.01	-0.09	-0.04	-0.15 *	-0.11	-0.09	-0.06	0.01	-0.08	0	0.01
Abhd12	Q6AYT7	-0.07	-0.1	-0.07	-0.23 *	-0.04	-0.16	NA	NA	NA	NA	NA
Serping1	Q6P734	0.03	0.04	-0.06	-0.15 *	-0.02	0.07	0.07	-0.02	0.07	-0.07	-0.03

(Table continues)

B)

Symbol	Uniprot ID	3R4F-23-F-90d	THS2.2-23-F-90d	THS2.2-50-F-90d	3R4F-23-F-90+42d	THS2.2-23-F-90+42d
Aqp1	P29975	0.07	0.13	0.16 *	-0.07	-0.07
Acta1	P68136	-0.22	-0.11	-0.15 *	-0.01	-0.06
Hsd17b4	P97852	0.08	0.05	0.11 *	-0.06	-0.06
Ddi2	Q4V8J5	0.13	0.21	0.28 *	0.13	0.21
Ddx39a	Q5U216	0.07	0.3 *	0.19 *	-0.18	-0.05
Fermt3	B2GVB9	0.1	0.09	0.12	-0.1	-0.16 *
Carkd	D3ZLK9	-0.11	-0.08	-0.14	-0.14	-0.16 *
LOC100365995	D3ZQV0	-0.01	-0.04	0.11	NA	-0.3 *
Elane	D4A488	-0.03	-0.02	0.06	-0.03	-0.12 *
Bub3	D4A567	0.02	-0.07	-0.07	-0.11	-0.17 *
Uqcrfs1	P20788	0.02	-0.02	-0.01	-0.07	-0.11 *
Cd9	P40241	-0.14	-0.05	-0.06	0.12	0.17 *
Lgals5	P47967	-0.08	-0.07	-0.02	-0.45	-0.68 *
Anxa6	D4ABR6	0.05	-0.02	-0.02	-0.01	-0.01
Cox7c	P80432	0.1	-0.07	-0.11	-0.09	-0.24 *
Cd47	P97829	-0.04	0.02	-0.01	0.06	0.12 *
Tnnc1	Q4PP99	-0.24	-0.25	-0.15	-0.11	-0.24 *
Sun1	Q5U2W0	0.02	0.06	0.15	0.27	0.34 *
Rcn2	Q62703	-0.05	-0.02	-0.09	-0.1	-0.12 *
Olfml1	Q66H86	-0.04	-0.07	-0.05	-0.09	-0.16 *
Pccb	P07633	0.05	NA	NA	NA	NA
Efemp1	Q6AXN2	0.07	-0.02	-0.02	-0.19	-0.25 *
Uqcrq	Q7TQ16	-0.19	-0.09	-0.03	NA	-0.39 *
Sgpl1	Q8CHN6	0.29	0.23	0.19	NA	0.6 *
Pabpc1	Q9EPH8	0.08	0.03	0.01	-0.06	-0.09 *
C7	F1M7F8	-0.07	-0.21 *	-0.11	0.02	-0.04
Mpst	P97532	0.1	0.11 *	0.09	0	-0.04
Dcn	Q01129	-0.28	-0.32 *	-0.31	0.15	0.03
Mapk14	Q56A33	0.11	0.34 *	0.33	-0.05	0.1
Sumo3	Q5XIF4	-0.08	-0.12 *	-0.12	0.15	0.11
A1m	Q63041	-0.18	-0.2 *	-0.2	0.12	-0.34

Log_2 fold-changes are shown. Statistically significant differential expression is indicated with a "*" (BH-adjusted p-value <0.05). These differentially expressed proteins were found only after THS2.2 exposure, but not after 3R4F exposure. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

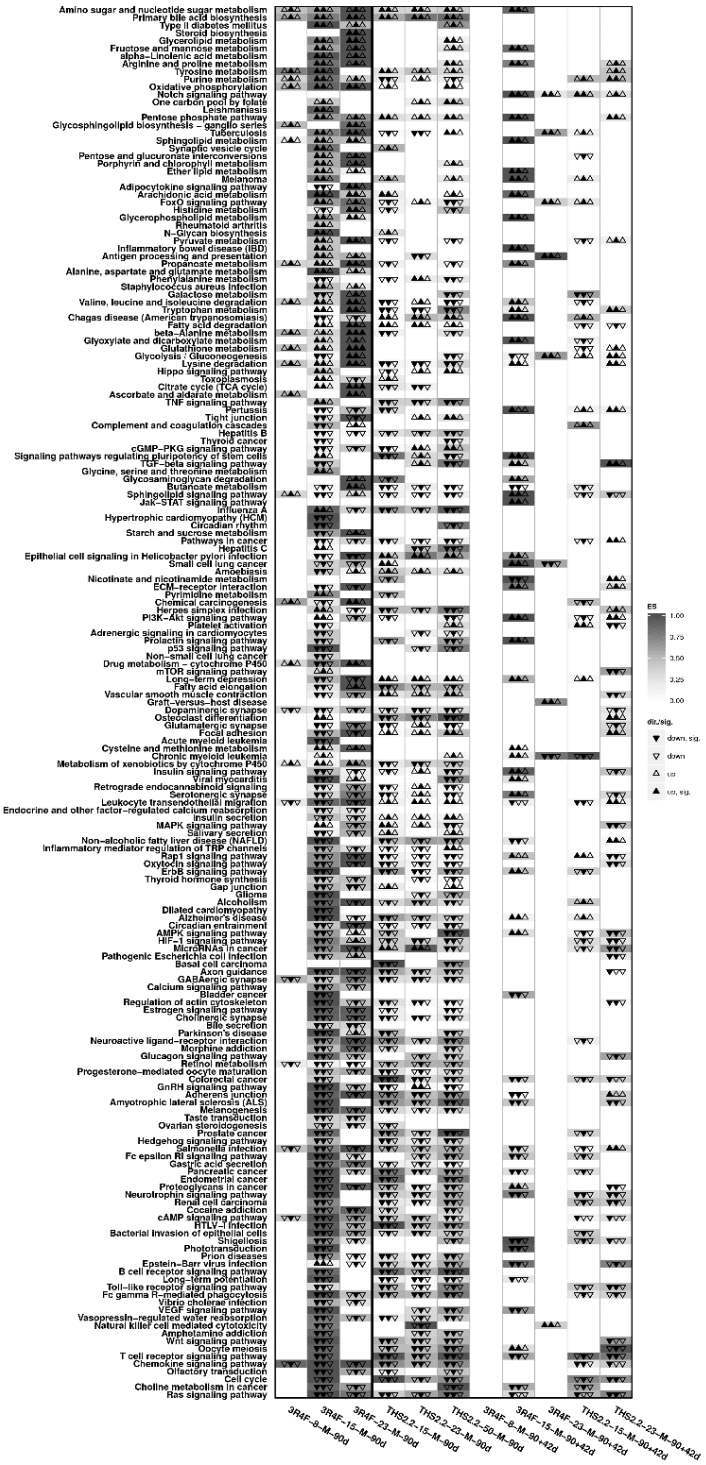
In the male rats, 48 significantly differentially expressed proteins were found after THS2.2 exposure that were not significantly differentially expressed after 3R4F exposure (Table 9A). In the female rats, 30 significantly differentially expressed proteins were found after THS2.2 treatment that were not found to be differentially expressed after 3R4F treatment (Table 9B).

In the male rats most of the "unique" proteins were found in the 90-day exposed groups, whereas in the female rats, more "unique" proteins" were found in the 90+42 days recovery group than in the 90-day exposed groups. This is consistent with a specific sham, 90+42-day effect (see above).

9.4.2 Set enrichment analyses to determine the effects of the exposure conditions on groups of proteins

The analysis on the level of individual proteins was complemented by a protein set enrichment analysis (Figure 16). For this, protein sets from the KEGG compendium (Kanehisa 2014) that were significantly affected by the tested exposure conditions were identified. The KEGG compendium is a comprehensive database of functionally related gene/protein sets. For biological interpretation it is important to note that not all protein/gene sets in the diverse KEGG collection are necessarily directly relevant for our biological context.

A)



(Figure continues)

B)

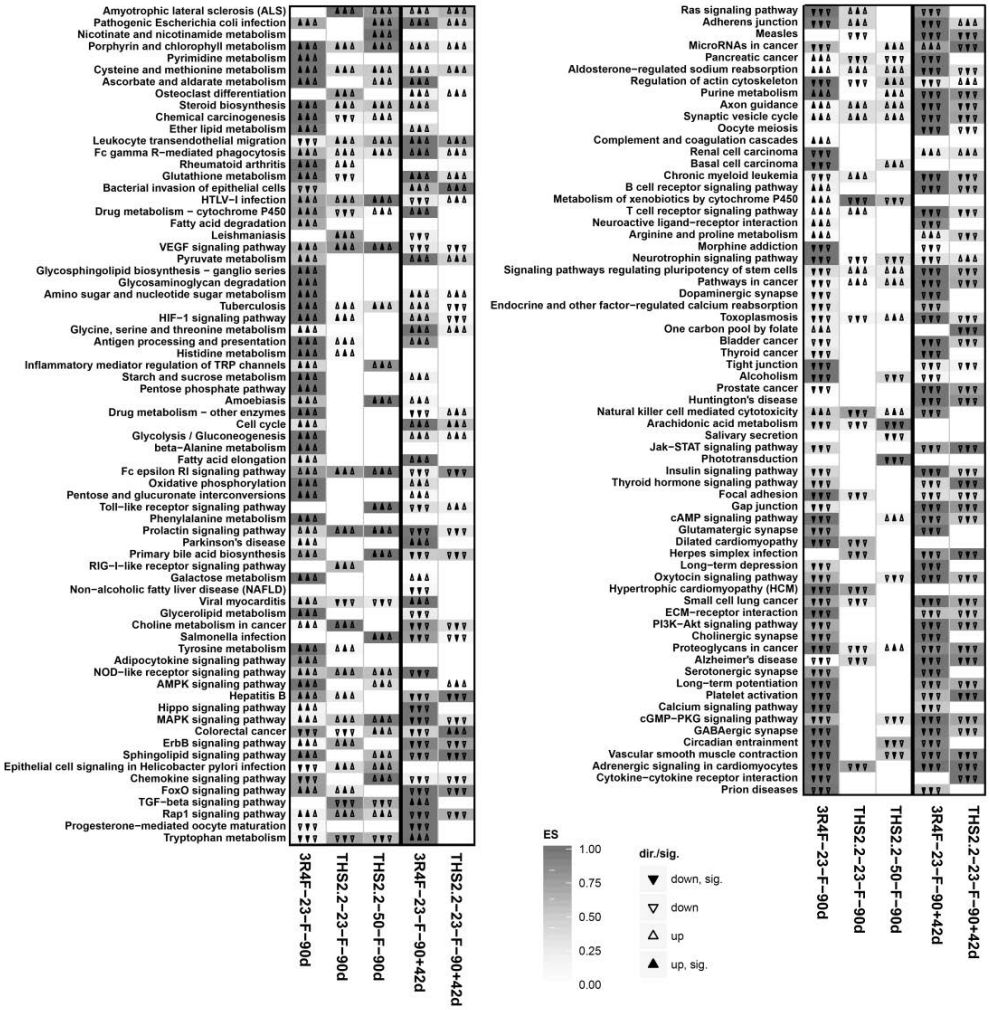


Figure 16: Enriched pathways for the differentially expressed proteins under the different exposure conditions

The enrichment was assessed by an over-representation analysis (ORA) and a gene-set analysis (GSA). For the GSA, the fold-change compared with the sham group was used as the protein-level statistic, the mean was used as the set statistic, and sample permutation (Q2) was used to assess the statistical significance. The enrichment results were normalized per pathway and the normalized GSA enrichment score was mapped to the grey-scale. Significance is indicated by a filled triangle (BH-adjusted p-value <0.05), with the direction is given by the orientation of the triangle. In each column, the first triangle represents the overall adjusted Fisher p-value of the GSA and ORA analyses, the second triangle represents the GSA, and the third triangle represents the ORA results. A blank space indicates non-significant associations. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons. Please zoom in pdf version for more details.

Exposure to 3R4F at 15 and 23 µg/l nicotine concentrations induced an overall strong protein set response in male rats. THS2.2 exposure at 15, 23 µg/L and 50 µg/L nicotine also induced an enrichment response in many protein sets, amongst which two were unique (i.e., not enriched in the 3R4F-exposed groups). In the 3R4F-exposed groups, fewer affected protein sets were identified for the 90+42-day recovery period.

Exposure to 3R4F at 23 µg/l nicotine induced a strong protein set response in female rats. Fewer affected proteins sets were found after THS2.2 exposure to 23 and 50 µg/L nicotine.

In the recovery phase, many significantly enriched protein sets were found in female rats in both the 3R4F and THS2.2 groups, which again can likely be explained by the specific effect of the sham, 90+42-day group (see above).

9.5 Lung miRNAome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats

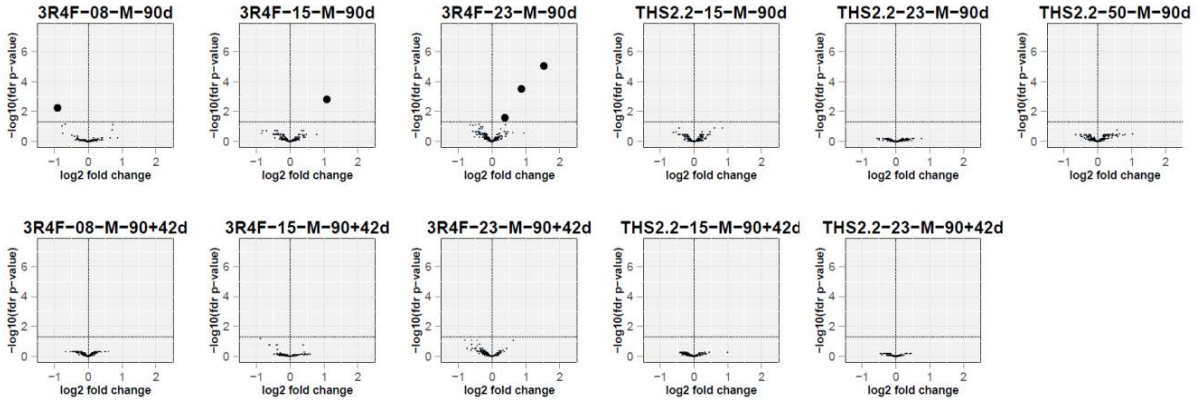
9.5.1 Differentially expressed miRNAs in lung tissue

The amplitude (expressed as log2 fold change, x-axis) and statistical significance (expressed as −log10 of the FDR) of miRNA expression changes in lung for the 3R4F- and THS2.2-exposed rats after 90 days of exposure or after a 42-day recovery period were compared with the sham group and visualized as volcano plots (Figure 17).

In the lung tissue of male rats three miRNAs were found to be differentially expressed after 90-days exposure to the highest concentration of 3R4F (23 µg/L nicotine). No differentially expressed miRNAs were found in the tissue of male rats after 90-days exposure to all three THS2.2 doses. No differentially expressed miRNAs were found after the 42-days recovery period in the 3R4F- and THS2.2-exposed groups.

In the lung tissue of female rats four miRNAs were found to be differentially upregulated after 90-days exposure to 3R4F (23 µg/L nicotine). In the THS2.2 group exposed to 23 µg/L nicotine, two differential expressed miRNAs were found, and in the THS2.2 group exposed to 50 µg/L) nicotine four differentially expressed miRNAs were found. A total of eight miRNAs were found to be differentially expressed in lung tissue of female rats after the 42-days recovery period, six of them were down-regulated. No differentially expressed miRNAs were found in the recovery period after THS2.2 exposure in the lung tissue of the female rats.

A)



B)

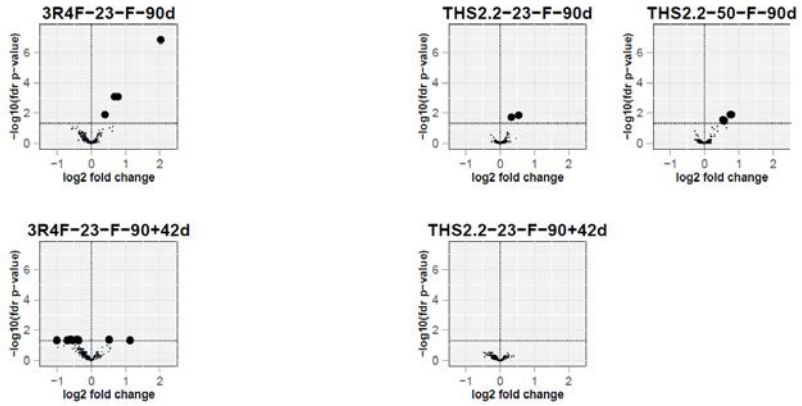


Figure 17: Volcano plots representing the expression profiles of differentially expressed miRNA in lung samples

The expression change for each miRNA, calculated as the $\log_2$ fold change, is plotted on the x-axis and the statistical significance, proportional to the $-\log_{10}$ of the BH adjusted p-value, is plotted on the y-axis. Dark grey dots indicate genes that are statistically significantly up- or down-regulated (right and left quadrants, respectively) compared with those in sham at the same time point. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

A heatmap was generated to obtain an overview on the overlap of the differential expressed miRNA in each group (Figure 18).

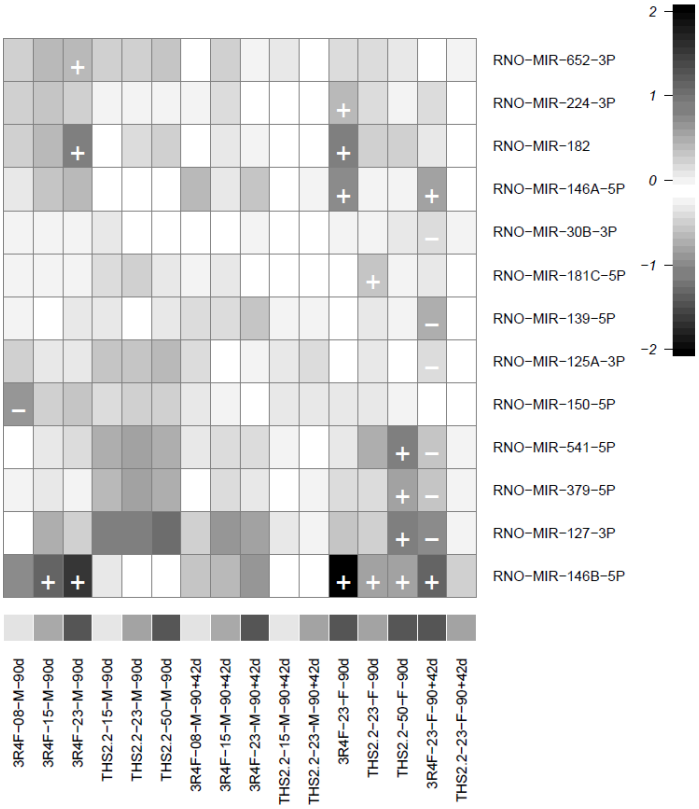


Figure 18: Heatmap of changes in miRNA expression in lung tissue

The heatmap represents the log2(fold-change) for all the miRNA that were significantly differentially expressed in at least one of the comparisons. The grey scaling is based on the absolute values. See caption of Figure 3 for labeling of the group comparisons.

For example, rno-miR-146b was found to be significantly differentially expressed in male and female rats after 3R4F exposure. After THS2.2 exposure, rno-miR-146b was significantly differentially expressed only in female rats, but the fold change was lower than it was after the 3R4F exposure. The expression of rno-miR-146b decreased after the 42-days recovery phase in female rats in the 3R4F and THS2.2 groups. rno-miR-146b was not found to be significantly expressed after the 42-days recovery phase in male rats in the 3R4F group.

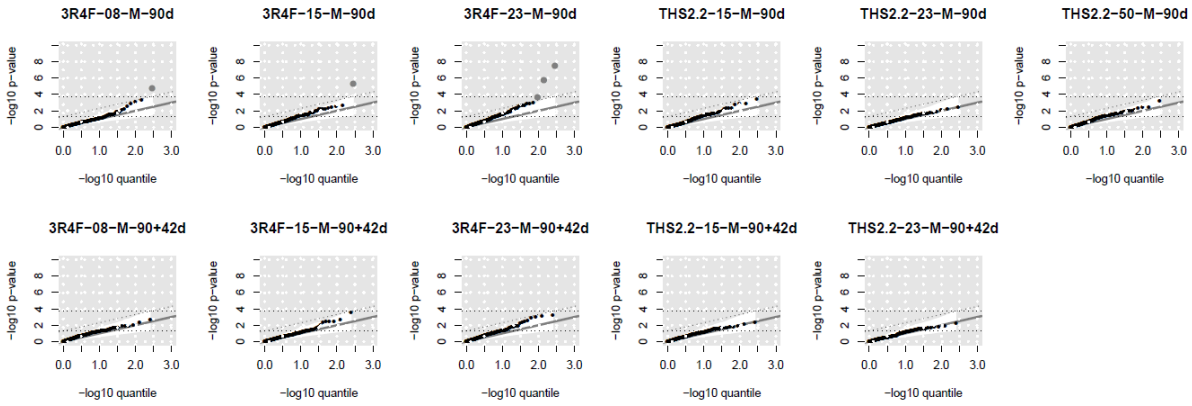
Three miRNAs (rno-miR-541, rno-miR-379, and rno-miR-127) were significant upregulated in the high THS group 50 µg/L nicotine in female lung tissue. Up-regulation was also seen in the male

THS2.2 groups but the fold-change was not significant. These three miRNAs are significantly down-regulated in the 3R4F recovery phase in female rats.

To confirm the cases where no differentially expressed miRNAs were found, a quantile-quantile (QQ) plot was used to provide additional support to the absence of a detectable signal by explicitly comparing the obtained results with the results of randomly permuted expression matrices.

The obtained p-values, measured by the statistics of the miRNA differential expressions in exposed versus sham groups, were ranked and plotted against their corresponding quantile, as shown in Figure 19. If the data indeed follow a random distribution, then the points on the QQ plot should fall approximately on a straight line. Thus, the QQ plots allow the effect of the exposures to be assessed globally.

A)



B)

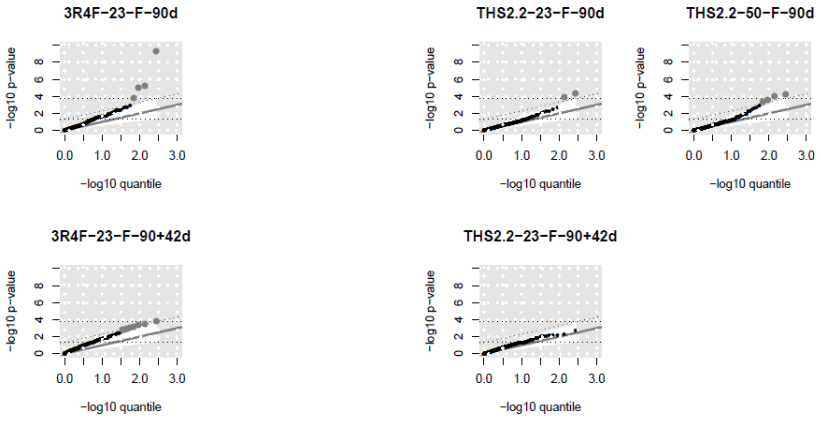


Figure 19: Quantile-quantile plots (QQ plot) of miRNA differential expression

For each comparison, all the detected miRNAs were ranked according to decreasing raw p-values. The −log10 values of the corresponding quantiles are plotted on the x-axis, and the −log10 values of their raw p-values are plotted on the y-axis. Deviations above the diagonal, which represents the expectation for a random expression matrix, indicate a globally significant signal was detected in the comparisons. See caption of Figure 3 for labeling of the group comparisons.

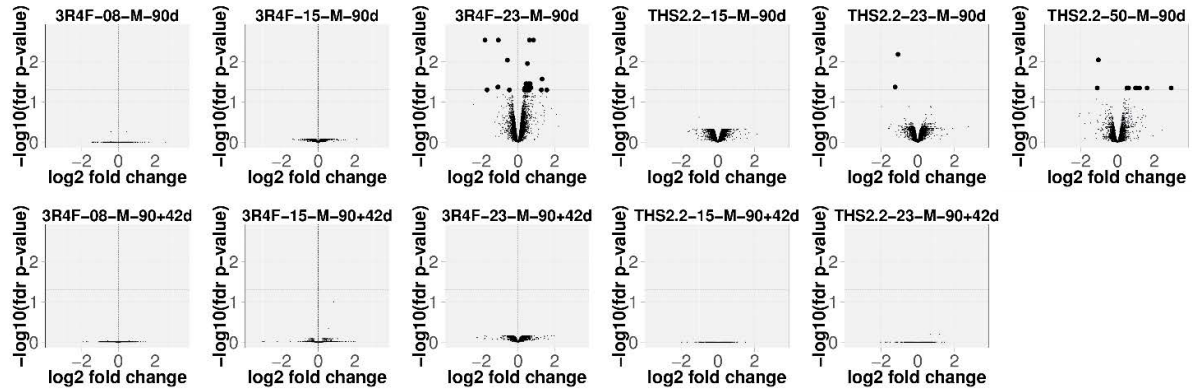
.

9.6 Liver Transcriptome at the 90-day Exposure Time Point and after a 42-day Recovery Period in Male and Female rats

9.6.1 Differentially expressed genes in liver tissue after 90 days of exposure to 3R4F or THS2.2, or after a 42-day recovery period

The amplitude and statistical significance of gene expression changes in liver after 90 days of exposure to 3R4F CS or THS2.2 aerosol, and after a 42-day recovery period were compared with the sham group and visualized as volcano plots (Figure 20).

A)



B)

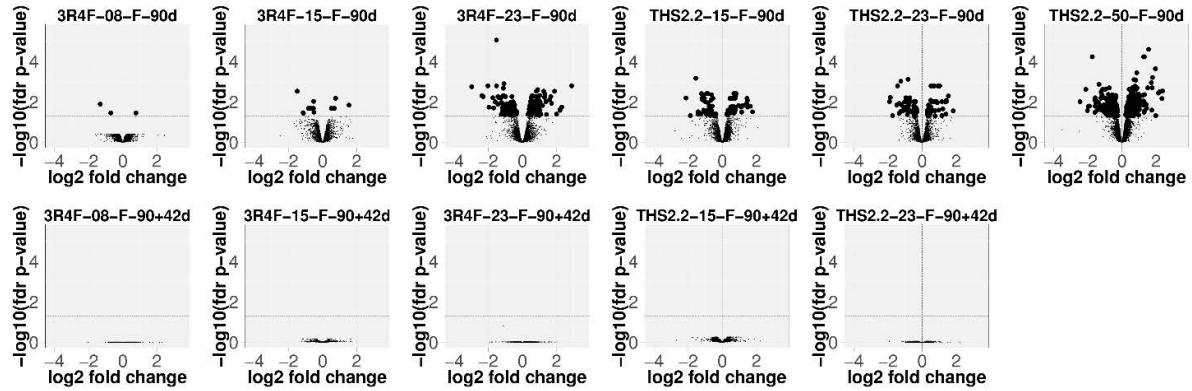


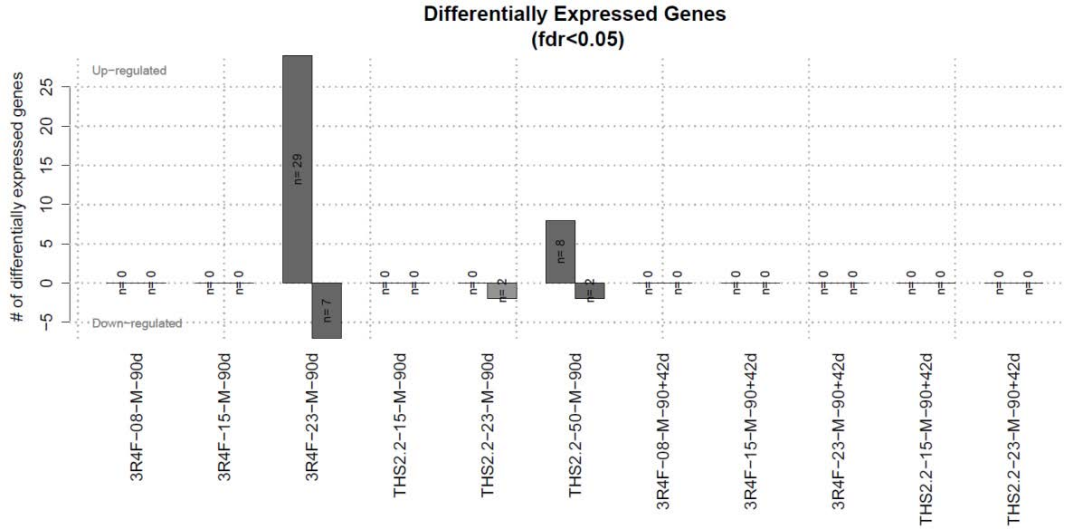
Figure 20: Volcano plots showing the systems response profiles in liver transcriptome

The expression changes for each gene, calculated as the log2 fold change, was plotted on the x-axis and the statistical significance, proportional to −log10of the FDR p-value, was plotted on the y-axis. Dark grey dots indicate genes that are statistically significantly up- and down-regulated (right and left quadrants, respectively) compared with sham at the same time point. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

A limited but dose-dependent effect was found in the liver of 3R4F-exposed male rats. The THS2.2-exposed male rats also showed a dose-dependent response, but fewer genes were involved compared with the 3R4F group. Indeed, in the high 3R4F dose (23 µg/L nicotine), a total of 36 genes were differentially expressed. In contrast, only two DEGs were found in the THS2.2 group at the same nicotine concentration (23 µg/L nicotine), and only 10 DEGs were observed in the high THS2.2 group (50 µg/L nicotine) (Figure 20A).

A dose-dependent effect was also found in the liver of 3R4F-exposed female rats. A total of 167 genes were differentially regulated in the high 3R4F group (23 µg/L nicotine). A total of 78 DEGs were found in the low THS2.2 group (15 µg/L nicotine) while only 15 DEGs were found at the same concentration in the 3R4F group. Contrarily, only 93 DEGs were found in the medium THS2.2 group (23 µg/L nicotine) while 167 DEGs found at the same concentration in the 3R4F group (Figure 20 B).

A)



(Figure continues)

B)

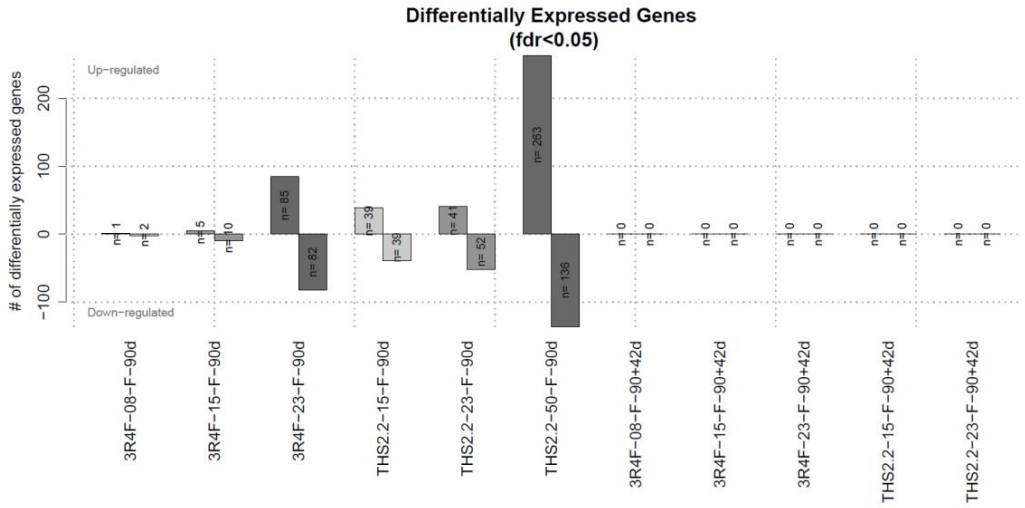
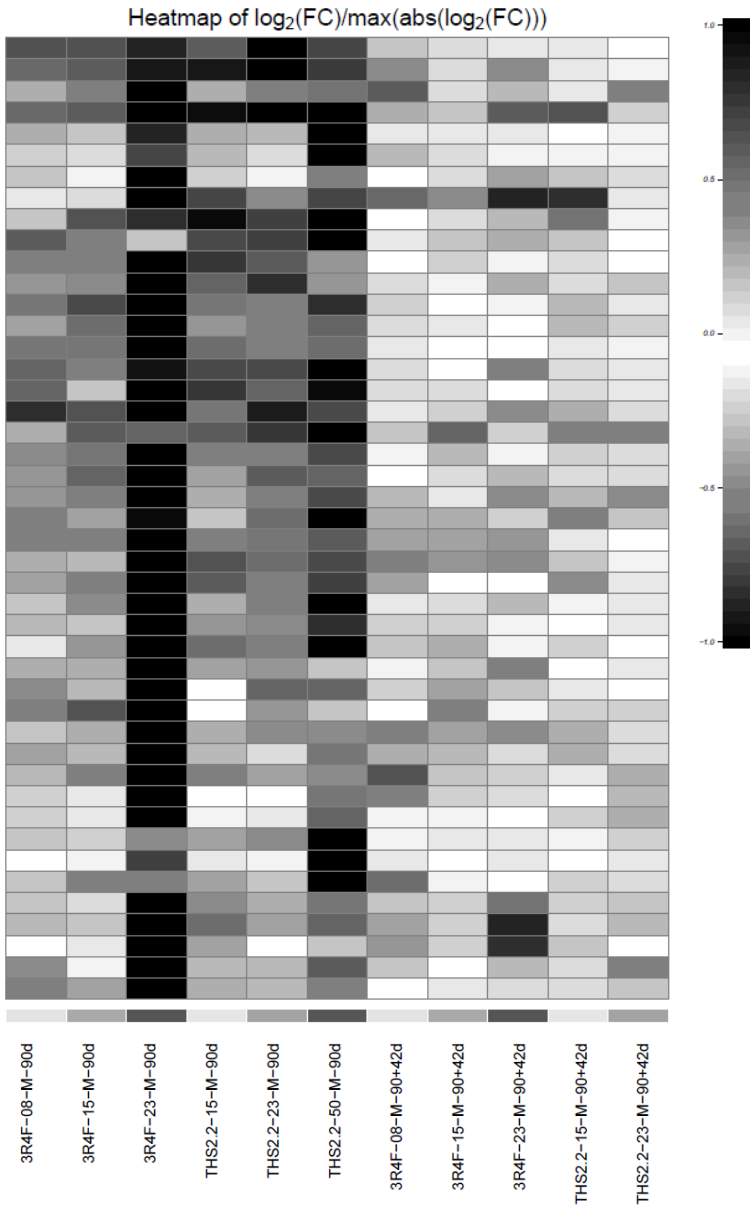


Figure 20: Significantly differentially expressed genes in liver tissue of male and female rats

Bar plots indicate the number of DEGs with FDR <0.05. A) Male rats, B) Female rats. See caption of Figure 3 for labeling of the group comparisons.

A heatmap was generated to obtain an overview how the differential gene expression was distributed over the different exposure groups, and to visualize the overlap of the DEGs (Figure 21).

A)



(Figure continues)

B)

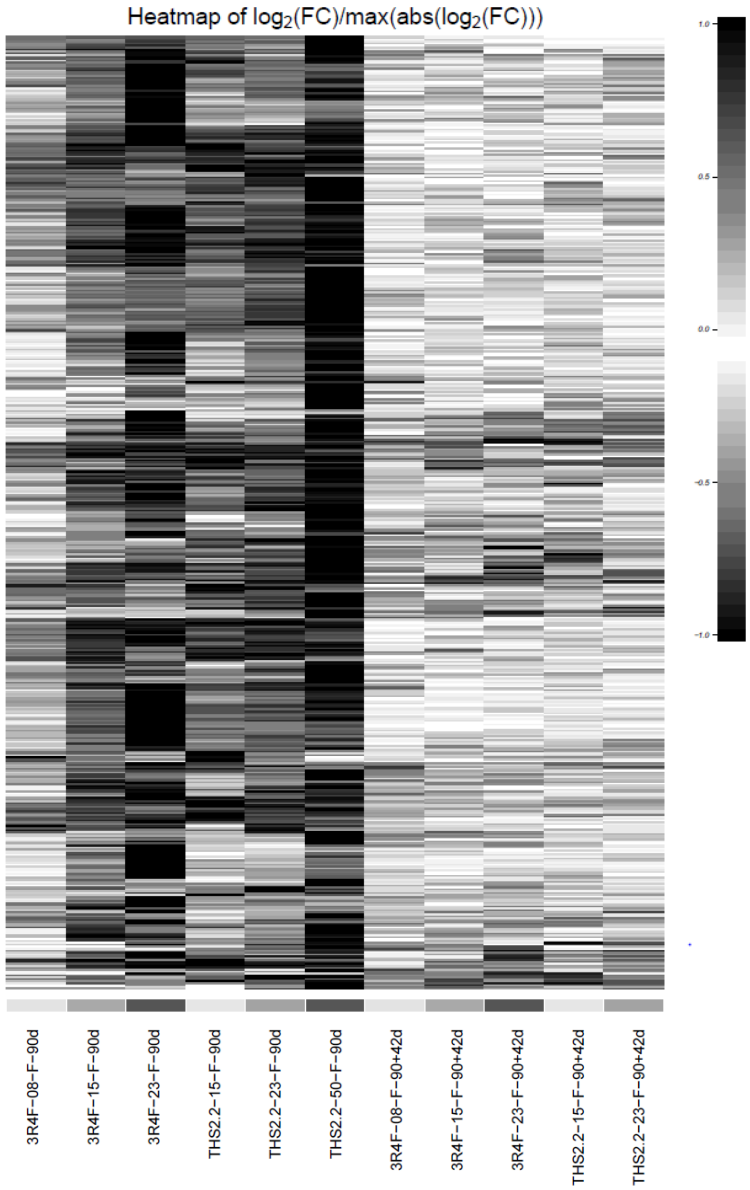


Figure 21: Heatmap of changes in gene expression in liver tissue

The heatmaps represent the log2(fold-change)/max(abs(log2(fold-change))) for all the genes that were differentially expressed in at least one of the comparisons. The grey scaling is based on the absolute values (range -1 to 1). A) Male rats, B) Female rats. FC = fold change. See caption of Figure 3 for labeling of the group comparisons.

In both male and female rats the heatmaps show a reduction in the fold change of the expressed genes in the medium exppsed THS2.2 group (23 µg/L nicotine) compared with the 3R4F-23-M/F-90d group (Figure 21). In both the 3R4F and THS2.2 groups exposed to 15 µg/L of nicotine, the overall fold change of expressed genes was similar. Similar reductions of gene fold changes were observed in all the 3R4F- and THS2.2-groups after the 42-days recovery period.

The genes that were significantly expressed only after THS2.2 exposure, but not after 3R4F exposure, are listed in Table 10.

Table 10: "Unique" genes that were differentially expressed genes in liver only after THS2.2 exposure.

A)

	Differentially Expressed Genes in Male Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Abcc3	0.21	0.50	0.49	0.32	0.19	1.2 *	0.63	0.00	-0.08	0.13	0.11
Akr1b7	-0.09	0.05	2.09	0.17	-0.38	2.92 *	0.15	-0.06	0.06	-0.15	0.11
Fads2	0.34	0.23	0.08	0.39	0.42	0.60 *	-0.09	-0.17	-0.22	-0.16	-0.01
Fbxo3	-0.10	-0.08	-0.08	-0.04	-0.14	-0.17 *	-0.31	-0.22	-0.33	-0.18	-0.28
Hao2	-0.26	-0.22	-0.81	-0.36	-0.22	-1.10 *	-0.34	-0.20	0.02	-0.11	0.01
Lgals2	0.22	0.57	0.53	0.58	0.75	0.98 *	0.17	0.56	0.10	0.44	0.43
LOC103690059	0.23	1.02	1.26	1.53	1.12	1.61 *	-0.14	-0.31	0.33	-0.95	-0.19
LOC691853	0.28	0.20	0.46	0.35	0.34	0.51 *	-0.09	-0.04	-0.26	-0.10	-0.09
Sgk2	0.14	0.11	0.39	0.27	0.36	1.04 *	-0.01	-0.14	-0.18	-0.14	0.13

(Table continues)

B)

Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group						
Cyp1a1	Prpsap1	Cyp2c13	Lactb2	Nsdhl	Mpped2	LOC685619
G6pd	Plekhb1	Dnajc21	Mybl1	Ranbp6	Tp53inp2	Cabyr
Igfbp3	Synj2bp	Cxcl10	Mfsd2a	Chuk	Ncapd2	Tprg1l
Junb	Slc13a3	Ppap2c	Isg15	Hps6	Cspp1	LOC691807
Me1	Cldn1	Serpina4	Slc30a6	Pcgf6	Tox	LOC691853
Mog	Rsad2	Mafk	Dhrs7	Osmr	Tmem55a	Fam177a1
Pfkfb1	Csnk1g2	Lzts3	Serpina6	Pmvk	Ndufaf4	Arxes2
RT1-N3	Aacs	Ppif	Mknk2	Manba	Snapc3	Anapc16
Ccl2	Bcam	Gm2a	Top1mt	Rap1gds1	Gpn2	Cdkl5
Slc10a1	Avil	Atf5	Hykk	Kbtbd4	Mthfr	Snx21
Tgfa	Igfals	Mx2	Fam46a	Rrbp1	Rtcb	LOC100363236
Ptn	Gna12	Cyp2b3	Scap	Dolk	Slc41a2	Cystm1
Xrcc6	Mvd	Camk2n1	Crb3	Crat	Creld2	LOC100911353
Rbp1	Mvk	Rogdi	Ankrd44	Stard8	Nmnat3	LOC102550246
Apoa1	Serpina7	Pgp	Tspan6	Cnot4	Ube2f	LOC102551341
Hsd11b1	Slc7a7	Adprm	Mbnl3	Dennd2a	Plin4	LOC102553866
Abcc2	Nde1	Tlcd1	Tmem97	Usp18	Pir	LOC103694321
Ephx1	Crot	Ptplb	Tmem98	Xkr8	Mid2	
Cpt2	Irs3	Tpst2	Lig3	Yars	Grap	
Cyp51	Enpp2	Hmcn1	Stxbp4	Tbc1d2	Rnaseh2a	
Lgals9	Nrtn	Slamf9	Snx11	Coro2a	Ssh3	
Ube2i	Plcl1	Eprs	Abca8a	Wdr78	Pigv	
G6pc	Spa17	Sowahb	Ube2o	Tfap2e	Slc30a3	
Igfbp2	Pex11a	Pdhb	Engase	Sptb	Nfe2	
Hmgcr	Col5a2	Tinf2	Igf2bp2	Otub2	Gjb2	
Igfbp1	Slc44a1	Dct	St3gal6	Tecpr2	RT1-CE10	
Cpt1b	Prkcdbp	Isyna1	Fam20c	Tmcc3	Repin1	
Dcn	Polg	Eci2	Zan	Irak3	Xk	
Capns1	Vamp7	Gpr158	Oasl	Tob2	Ern1	
Inhba	Cxadr	Nedd4l	Dcaf15	Sesn3	Fn3k	
Rcn2	Kif1c	Rprd1a	Diexf	Pih1d2	Mapkapk5	
Stmn3	Hspb8	RGD1311805	Osbp2	Morc4	Commd6	
Meox2	Ankh	Setd6	Comtd1	Ccdc51	Kxd1	
Chrne	Itga6	Aars	Abhd6	Abhd5	Idnk	
Hap1	Dusp1	Ypel3	Nek4	Rpl7l1	Thnsl1	
Msra	Erp29	Irf7	Unc5d	Nek6	Myo7b	
Ighmbp2	Cdh17	Actr1a	Iars	Igfbp4	Mrps11	
Fdft1	Sh3bp5	RT1-S3	Akr1c19	Plekhh3	Eid1	
Ndst1	Ebp	Scube3	Rbfa	Popdc2	Plekhd1	
Hmgcs1	Txnip	Supv3l1	Lama3	Sdsl	RGD1566220	
Pdlim1	Inadl	Man1a1	Kifc3	Noc4l	RGD1562114	
Btg3	Chek1	Pcdh18	Def8	Rgs16	Lrrc14	
Car2	A1bg	Mab21l3	Pdzd8	Tfdp1	Arhgap42	
Car3	Rcan2	Dph5	Cnksr3	Mocos	Bfsp2	
Timm17a	Abcc3	RGD1309730	Lrrk1	Prrg2	Stradb	
Adam15	Msmo1	Secisbp2l	Uvrag	Zfp143	Rnf14	
Cpq	Samsn1	Ttpal	Egln1	Srpk1	Tmem69	
Slc29a1	Cbwd1	Fam210b	Cdr2	Srsf3	LOC678766	
Lrat	Hdgfrp2	Psmd5	Dctn5	Mynn	Fer1l5	
Dhcr7	Rnf39	Dusp16	Slc22a18	Agl	Sin3b	
Timm8b	Cyp2c22	Terf1	Mrpl49	Creb3l1	LOC685078	

A) DEGs found only in the liver tissue of male rats after THS2.2 exposure and not in the liver tissue of female rats after the same treatment. (B) DEGs found only in the liver tissue of female rats after THS2.2 exposure and not in the liver tissue of male rats after the same treatment. All these DEGs had a FDR <0.05 only in one the THS2.2 groups and not in any of the 3R4F exposed groups. For the DEGS found only in female rats, only the names of the genes are listed. The full table including the fold changes and significance compared with the references can be viewed in Supplementary Table 1.

10 DISCUSSION AND CONCLUSION

The standard OECD study designs according to test guidelines 412 or 413 are providing information on the hazards that arises from repeated exposure by the inhalation route. These reveal effects on pulmonary inflammation and other biological endpoints, including histopathological endpoints (e.g.,(Kogel 2014, Werley 2008b)). But toxicological markers are often only detectable after substantial damage has already occurred and, thus, can lack sensitivity (Ellinger-Ziegelbauer 2011). The OECD also promotes the regulatory use of alternative testing methods that reduce, refine, and replace traditional animal testing. As a foundation of alternative testing methods, the mechanisms and modes of toxicity need to be understood (OECD 2013).

To obtain additional insights into the molecular mechanisms underlying the toxicity of THS2.2 and to increase sensitivity, we have included an additional study arm into the classical OECD inhalation study design (the "PLUS" arm). In addition to classical endpoints, we routinely also investigate the exposure effects on the transcriptome and proteome using omics approaches. Such combined approaches of omics and conventional toxicology are effective for mechanistic investigations and identification of putative biomarkers (Ellinger-Ziegelbauer 2011, Moffat 2015, Oberemm 2005).

Within the "PLUS" arm, we included additional animals that underwent the same exposure conditions as the animals in the regular OECD study arm. Organs of interest, including respiratory nasal epithelium (RNE), lung tissue, and liver tissue, were collected. With this, omics data from the tissue sites, for which a phenotypic change was expected to occur, were generated and computationally analyzed (Table 11).

In the nasal respiratory epithelium of 3R4F-exposed (23 µg/L nicotine) male rats more than 3,900 genes were differentially expressed after 90 days of exposure. In contrast, after exposure to the high dose of THS2.2M (50 µg/L nicotine) for 90 days, no genes were differentially expressed in the RNE of male rats.

To further elucidate the underlying biology, we conducted a causal biological network analysis. This analysis is based on a collection of modular, evolving, manually curated, biological expression language (BEL)-scripted causal network models that describe signaling pathways relevant in diseased and non-diseased pulmonary and vascular tissues (Boué 2015). We have assessed five process categories describing the mechanisms for cell fate, tissue repair and angiogenesis, inflammatory process network, cell stress, and cellular proliferation. All these biological processes are well-known to be perturbed by cigarette smoke exposure. Indeed, all five process categories were impacted by 3R4F CS in the RNE of male rats. In contrast, THS2.2M aerosol exposure impacted these processes much less.

In the lung tissue of 3R4F exposed male rats, a dose-dependent effect was found. The highest 3R4F dose showed up to 788 differential expressed genes. In contrast, no differential expressed gene was

found for THS2.2 aerosol exposure (even at the highest concentration of 50 µg/L nicotine). In the lung tissue of 3R4F exposed female rats, at a concentration of 23 µg/l nicotine, 1,302 differential expressed genes were identified. At the same nicotine concentration, only one differentialy expressed gene was identified when the rats were exposed to THS2.2 aerosol. For the highest THS2.2 aerosol concentration of 50 µg/L nicotine, 17 genes were differentially expressed after the 90-days exposure – still well below the effect of 3R4F CS at approximately one-half of the THS2.2 nicotine concentration (23 µg/l nicotine).

The causal network analysis showed that all five tested biological network processes (cell fate, tissue repair and angiogenesis, inflammatory processes, cell stress, and cellular proliferation) were impacted in male lung tissue after the 90-day exposure to 3R4F CS at all tested concentrations (8-23 ug/L nictotine). The inflammatory process network contributed most to the biological impact factor. After 90 days exposure to the highest concentration of THS2.2 (50 µg/L nicotine) a perturbation of the inflammatory process network was accompanied by a small perturbation of the tissue repair and angiogenesis network, but the perturbation amplitudes were much lower than after exposure to 3R4F CS. In the lung tissue of female rats after 90 days of exposure to 3R4F the inflammatory process network also had the highest contribution, followed by perturbation of cell fate, tissue repair and angiogenesis, and cell stress. Exposure to THS2.2 showed perturbation of the cell stress and the inflammatory process networks, but the amplitude of the perturbation was clearly lower than for reference cigarette exposure.

Consistent with the transcriptomics data, the proteomics data also supported a strong reduction in differentially expressed proteins for the THS2.2 exposed groups compared with the 3R4F exposure groups. Furthermore, evaluation of gene sets from the KEGG compendium, suggested a reduction in biological effects of THS2.2 aerosol compared to 3R4F CS.

Liver tissue (for the same exposure conditions) showed lower systemic effects on gene expression in term of total number of DEGs. Neverthless, the observations were in line with effects observed in the respitarory tissues. The liver of male rats exposed to 3R4F CS with the highest nicotine concentration (23 ug/L) showed 36 DEGs compared to 2 DEGs observed upon THS2.2 aerosol exposure. In the liver of female rats the total number of DEGs was higher than in the liver of male rats. The liver samples of female rats exposed to 3R4F CS showed 167 DEGs, whereas THS2.2 aerosol exposure resulted in 93 DEGs (both 23 ug/L). Interestingly, at lower nicotine concentration (15 ug/L) THS2.2 aerosol generated an higher DEGs number compared to 3R4F (78 vs 15).

In summary, we found throughout the omics analysis that THS2.2 aerosol, although applied at a higher nicotine concentration, did cause a less strong effect measured as differential fold changes for transcriptomics and proteomics, or measured as reduced quantitative impact on pathways for transcriptomics and proteomics (Table 11). Overall, the data suggest that human adults smoking THS2.2 will not be exposed to additional hazards compared with those from smoking conventional cigarettes.

Table 11: Overarching summary of results

	Group x Tissue (data type)		RNE			Lung				Liver	
			DEG		RBIF	DEG		DEP	RBIF	DEG	
90d	3R4F low (8 µg nicotine/l)	male	0	↑	0%	41	↑	10	48%	0	↑
			0	↓		4	↓	2		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	1	↑
			NA	↓		NA	↓	NA		2	↓
	3R4F med (15 µg nicotine/l)	male	16	↑	6%	169	↑	547	68%	0	↑
			0	↓		13	↓	342		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	5	↑
			NA	↓		NA	↓	NA		10	↓
	3R4F high (23 µg nicotine/l)	male	2143	↑	100%	574	↑	173	98%	29	↑
			1768	↓		214	↓	100		7	↓
		female	NA	↑	NA	695	↑	157	100%	85	↑
			NA	↓		607	↓	83		82	↓
	THS2.2 low (15 µg nicotine/l)	male	0	↑	0%	0	↑	19	10%	0	↑
			0	↓		0	↓	31		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	39	↑
			NA	↓		NA	↓	NA		39	↓
	THS2.2 med (23 µg nicotine/l)	male	0	↑	2%	0	↑	0	6%	0	↑
			0	↓		0	↓	0		2	↓
		female	NA	↑	NA	1	↑	17	4%	41	↑
			NA	↓		0	↓	15		52	↓
	THS2.2 high (50 µg nicotine/l)	male	0	↑	1%	0	↑	90	5%	8	↑
			0	↓		0	↓	108		2	↓
		female	NA	↑	NA	7	↑	17	6%	263	↑
			NA	↓		10	↓	12		136	↓
90+42d	3R4F low (8 µg nicotine/l)	male	0	↑	1%	0	↑	0	43%	0	↑
			0	↓		0	↓	0		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	0	↑
			NA	↓		NA	↓	NA		0	↓
	3R4F med (15 µg nicotine/l)	male	0	↑	16%	3	↑	238	67%	0	↑
			0	↓		2	↓	101		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	0	↑
			NA	↓		NA	↓	NA		0	↓
	3R4F high (23 µg nicotine/l)	male	7	↑	51%	10	↑	2	100%	0	↑
			53	↓		0	↓	0		0	↓
		female	NA	↑	NA	80	↑	126	46%	0	↑
			NA	↓		6	↓	213		0	↓
	THS2.2 low (15 µg nicotine/l)	male	0	↑	0%	0	↑	0	1%	0	↑
			0	↓		0	↓	0		0	↓
		female	NA	↑	NA	NA	↑	NA	NA	0	↑
			NA	↓		NA	↓	NA		0	↓
	THS2.2 med (23 µg nicotine/l)	male	0	↑	1%	0	↑	0	6%	0	↑
			0	↓		0	↓	0		0	↓
		female	NA	↑	NA	0	↑	15	0%	0	↑
			NA	↓		0	↓	72		0	↓

The table summarizes the number of differential expressed genes (DEG), the number of diffential expressed proteins (DEP), and the value of the relative biological impact factor (RBIF). Then number of DEG and DEP were given in the bar charts in the respective tissue chapter. The RBIF was given on the top of the starplots indicating the level of network perturbation. The highest perturbation was set to 100%.

11 ARCHIVING

After completion of the study, the study plan with any amendments, the test and reference item retention samples, all raw data, the report with any amendments, and all further study-related records needed to reconstruct the study will be archived. They will be retained for at least 10 years according to the applicable archiving procedures. Paper records will be archived in the archive at Philip Morris Products S.A., Research & Development, PMI Product Testing, Neuchâtel, Switzerland, and electronic records will be managed by the PMI Product Testing e-archivist on the central archiving server at Philip Morris S.A., Lausanne, Switzerland.

12 ABBREVIATIONS

AVA	Agri-Food & Veterinary Authority
BH	Benjamini-Hochberg
BIF	biological impact factor
CFA	cell fate
CST	cell stress
CPR	cellular proliferation
DEG	differentially expressed gene
FDA	Food and Drug Administration
FDR	false discovery rate
HCI	Health Canada Intense (smoking regimen)
IPN	inflammatory processes network
ISO	International Organization for Standardization
MRTP	modified risk tobacco product
NACLAR	National Advisory Committee for Laboratory Animal Research
NPA	network perturbation amplitude
OECD	Organisation for Economic Co-operation and Development
pMRTP	prototypic (or potential) modified risk tobacco product
REF	reference
RNE	respiratory nasal epithelium
THS	tobacco heating system
TRA	tissue repair and angiogenesis
WKI	Work Instructions

13 REFERENCES AND SOURCES

13.1 Literature

Ackermann M, Strimmer K. A general modular framework for gene set enrichment analysis. BMC bioinformatics. 2009;10(1):47.

Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003a;19(2):185-93. Epub 2003/01/23. PubMed PMID: 12538238.

Bolstad BM, Irizarry RA, Åstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003b;19(2):185-93.

Boué S, Talikka M, Westra JW, Hayes W, Di Fabio A, Park J, et al. Causal biological network database: a comprehensive platform of causal biological network models focused on the pulmonary and vascular systems. Database. 2015;2015:bav030.

Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010;26(19):2363-7.

Cutler P, Bell DJ, Birrell HC, Connelly JC, Connor SC, Holmes E, et al. An integrated proteomic approach to studying glomerular nephrotoxicity. Electrophoresis. 1999;20(18):3647-58. Epub 1999/12/28. PubMed PMID: 10612292.

Davies HM, Vaught A. The reference cigarette. Research cigarettes. 2003.

Ellinger-Ziegelbauer H, Adler M, Amberg A, Brandenburg A, Callanan JJ, Connor S, et al. The enhanced value of combining conventional and "omics" analyses in early assessment of drug-induced hepatobiliary injury. Toxicology and applied pharmacology. 2011;252(2):97-111. Epub 2010/10/05. PubMed PMID: 20888850.

Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20(3):307-15. Epub 2004/02/13. PubMed PMID: 14960456.

Gaworski CL, Oldham MJ, Wagner KA, Coggins CR, Patskan GJ. An evaluation of the toxicity of 95 ingredients added individually to experimental cigarettes: approach and methods. Inhalation toxicology. 2011a;23(sup1):1-12.

Gaworski CL, Schramke H, Diekmann J, Meisgen TJ, Tewes FJ, Veltel DJ, et al. Effect of filtration by activated charcoal on the toxicological activity of cigarette mainstream smoke from experimental cigarettes. Inhalation toxicology. 2009;21(8):688-704.

Gaworski CL, Wagner KA, Morton MJ, Oldham MJ. Insights from a multi-year program designed to test the impact of ingredients on mainstream cigarette smoke toxicity. Inhalation toxicology. 2011b;23(sup1):172-83.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome biology. 2004a;5(10):R80.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004b;5(10):R80. Epub 2004/10/06. PubMed PMID: 15461798; PubMed Central PMCID: PMC545600.

Health Canada. Official Method T-115, Determination of "Tar", Nicotine and Carbon Monoxide in Mainstream Tobacco Smoke. Ottawa: 1999.

Health_Canada. Health Canada T-115: Determination of "Tar", Nicotine and Carbon Monoxide in Mainstream Tobacco Smoke. 1999.

Herbrich SM, Cole RN, West Jr KP, Schulze K, Yager JD, Groopman JD, et al. Statistical Inference from Multiple iTRAQ Experiments without Using Common Reference Standards. Journal of proteome research. 2013;12(2):594-604.

Hoeng J, Talikka M, Martin F, Sewer A, Yang X, Iskandar A, et al. Case study: the role of mechanistic network models in systems toxicology. Drug discovery today. 2014;19(2):183-92. Epub 2013/08/13. PubMed PMID: 23933191.

Huber W, Von Heydebreck A, Sültmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics. 2002;18(suppl 1):S96-S104.

Hultin-Rosenberg L, Forshed J, Branca RM, Lehtio J, Johansson HJ. Defining, comparing, and improving iTRAQ quantification in mass spectrometry proteomics data. Molecular & cellular proteomics : MCP. 2013;12(7):2021-31. Epub 2013/03/09. PubMed PMID: 23471484; PubMed Central PMCID: PMCPMC3708183.

International Organization for Standardization. ISO 3308:2000 - Routine analytical cigarette-smoking machine -- Definitions and standard conditions. 2000.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic acids research. 2003;31(4):e15-e.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic acids research. 2014;42(D1):D199-D205.

Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. Bioinformatics. 2009;25(3):415-6.

Kelstrup CD, Young C, Lavallee R, Nielsen ML, Olsen JV. Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. Journal of proteome research. 2012;11(6):3487-97.

Kogel U, Schlage WK, Martin F, Xiang Y, Ansari S, Leroy P, et al. A 28-day rat inhalation study with an integrated molecular toxicology endpoint demonstrates reduced exposure effects for a prototypic modified risk tobacco product compared with conventional cigarettes. Food and Chemical Toxicology. 2014;68:204-17.

Martin F, Sewer A, Talikka M, Xiang Y, Hoeng J, Peitsch M. Quantification of biological network perturbations for mechanistic insight and diagnostics using two-layer causal models. BMC Bioinformatics. 2014;15(1):238. PubMed PMID: doi:10.1186/1471-2105-15-238.

Martin F, Thomson T, Sewer A, Drubin D, Mathis C, Weisensee D, et al. Assessment of network perturbation amplitudes by applying high-throughput data to causal biological networks. BMC Systems Biology. 2012;6(1):54. PubMed PMID: doi:10.1186/1752-0509-6-54.

Moennikes O, Vanscheeuwijck P, Friedrichs B, Anskeit E, Patskan G. Reduced toxicological activity of cigarette smoke by the addition of ammonia magnesium phosphate to the paper of an electrically heated cigarette: subchronic inhalation toxicology. Inhalation toxicology. 2008;20(7):647-63.

Moffat I, Chepelev NL, Labib S, Bourdon-Lacombe J, Kuo B, Buick JK, et al. Comparison of toxicogenomics and traditional approaches to inform mode of action and points of departure in human health risk assessment of benzo [a] pyrene in drinking water. Critical reviews in toxicology. 2015.

NACLAR. (National Advisory Committee for Laboratory Animal Research): Guidelines on the Care and Use of Animals for Scientific Purposes. 2004.

Oberemm A, Onyon L, Gundert-Remy U. How can toxicogenomics inform risk assessment? Toxicology and applied pharmacology. 2005;207(2):592-8.

OECD. OECD Test Guideline 413: Subchronic Inhalation Toxicity: 90-Day Study, revised Toxicity Guideline 2009a.

OECD. Guidance document on developing and assessing adverse outcome pathways. Series on Testing and Assessment No. 184. OECD Environment, Health and Safety Publications. 2013;ENV/JM/MONO(2013)6.

OECD OfEC-oaD. Test No. 413. Subchronic Inhalation Toxicity: 90-day Study. http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/OECD/OECD-TG413.pdf2009b.

Patskan G, Reininghaus W. Toxicological evaluation of an electrically heated cigarette. Part 1: Overview of technical concepts and summary of findings. Journal of Applied Toxicology. 2003;23(5):323-8.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0; 2014.

R Development Core Team. R: A Language and Environment for Statistical Computing. 2007.

Renne RA, Yoshimura H, Yoshino K, Lulham G, Minamisawa S, Tribukait A, et al. Effects of flavoring and casing ingredients on the toxicity of mainstream cigarette smoke in rats. Inhalation toxicology. 2006;18(9):685-706.

Sales G, Calura E, Cavalieri D, Romualdi C. g raphite-a Bioconductor package to convert pathway topology to gene network. BMC bioinformatics. 2012;13(1):1.

Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology. 2004;3:Article3. Epub 2006/05/02. PubMed PMID: 16646809.

Stavanja MS, Curtin GM, Ayres PH, Bombick ER, Borgerding MF, Morgan WT, et al. Safety assessment of diammonium phosphate and urea used in the manufacture of cigarettes. Experimental and Toxicologic Pathology. 2008;59(6):339-53.

Sturla SJ, Boobis AR, FitzGerald RE, Hoeng J, Kavlock RJ, Schirmer K, et al. Systems toxicology: from basic research to risk assessment. Chemical research in toxicology. 2014;27(3):314-29.

Terpstra PM, Teredesai A, Vanscheeuwijck PM, Verbeeck J, Schepers G, Radtke F, et al. Toxicological evaluation of an electrically heated cigarette. Part 4: Subchronic inhalation toxicology. Journal of applied toxicology. 2003;23(5):349-62. Epub 2003/09/17. PubMed PMID: 12975774.

Theophilus EH, Pence DH, Meckley DR, Shreve WK, Ayres PH, Bombick BR, et al. Toxicological evaluation of cigarettes with two banded cigarette paper technologies. Experimental and Toxicologic Pathology. 2007;59(1):17-27.

US Food & Drug Administration. Modified risk tobacco product applications–draft guidance for industry. 2012.

Vanscheeuwijck P, Teredesai A, Terpstra P, Verbeeck J, Kuhl P, Gerstenberg B, et al. Evaluation of the potential effects of ingredients added to cigarettes. Part 4: subchronic inhalation toxicity. Food and Chemical Toxicology. 2002;40(1):113-31.

Väremo L, Nielsen J, Nookaew I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic acids research. 2013;41(8):4378-91.

Werley MS, Freelin SA, Wrenn SE, Gerstenberg B, Roemer E, Schramke H, et al. Smoke chemistry, in vitro and in vivo toxicology evaluations of the electrically heated cigarette smoking system series K. Regulatory Toxicology and Pharmacology. 2008a;52(2):122-39.

Werley MS, Freelin SA, Wrenn SE, Gerstenberg B, Roemer E, Schramke H, et al. Smoke chemistry, in vitro and in vivo toxicology evaluations of the electrically heated cigarette smoking system series K. Regulatory toxicology and pharmacology : RTP. 2008b;52(2):122-39. PubMed PMID: 18590791.

13.2 Philip Morris International's List of Standard Operating Procedures (SOP) and Work Instructions (WKI).

The procedures and instructions that were followed in this study are listed below:

Isolation of total RNA including miRNA in lung tissue slices using RNeasy mini kit RLT protocol: PMI_RD_WKI_001115

Quality controls following RNA extractions: PMI_RD_WKI_000978

Proteins Extraction from Lung Tissues and Acetone Precipitaion: PMI_RD_WKI_001043

Isolation of total RNA including microRNAs from Tissue and Organotypic tissue inserts using miRNeasy mini kit protocol: PMI_RD_WKI_001117

Determination of Protein Concentration by using Bradford Assay: PMI_RD_WKI_001103

Quality Control of Protein samples using LabChip instrument: PMI_RD_WKI_001131

miRNA labeling using the FlashTag Biotin HSR kits: PMI_RD_WKI_001123

Normalization Macro prior to IVT Protocol: PMI_RD_WKI_001118

Tissue_IVT_Worksheet: PMI_RD_FOR_000763

Affymetrix-IVT2-Express BioMek: PMI_RD_WKI_000813

Affymetrix-3'array-Cartridge-Hybridization: PMI_RD_WKI_000816

Tissue_AGCC_Upload: PMI_RD_FOR_000761

Computational Practices for mRNA and miRNA analyses: PMI_RD_WKI_001228

Omics_Processing: PMI_RD_SOP_000346

13.3 Philip Morris International's Internal Documents

Philip Morris International. Unpublished on file data. Internal Study Plan: STUDY PLAN 150122 - OECD NIKE-GENOMIC ENDPOINTS. Study Number 150122. 2012.

15006 THS SP: Philip Morris International. Unpublished on file data. Internal Study Plan; "90 DAY OECD INHALATION STUDY ON SPRAGUE-DAWLEY RATS TO CHARACTERIZE TOXICITY OF MAINSTREAM AEROSOL FROM ZURICH FPD 4.2 CIGARETTES AND COMPARISON TO 3R4F MAINSTREAM AEROSOL TOXICITY". Study number 15006. 2013.

15006_150082 THS SP: Philip Morris International. Unpublished on file data. Internal Study Plan; "90 DAYS OECD (+) ZURICH STUDY GENOMICS APPROACH". Study Number 150082. 2013.

15006_150083 THS SP: Philip Morris International. Unpublished on file data. Internal Study Plan; "90 DAYS OECD (+) ZURICH STUDY PROTEOMICS APPROACH". Study Number 150083. 2013.

15006 THS SR Part 1: Philip Morris International. Unpublished on file data. Internal Study Report; "STUDY REPORT 90 DAY OECD REGULAR INHALATION STUDY – PART 1: SUMMARY OF RESULTS FOR OECD GROUP ANIMAL". Study Number 15006. 2014

14 SUPPLEMENTARY TABLES AND SUPPLEMENTARY FIGURES

14.1 Supplementary Tables

Supplementary Table 1. Unique" genes that were differentially expressed only in the liver of female rats after THS2.2 treatment.

All the listed DEGs had a FDR <0.05 only in one the THS2.2 groups and not in any of the 3R4F exposed groups.

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Cyp1a1	0.53	1.12	0.76	0.72	0.88	0.87 *	0.12	0.41	0.14	0.25	0.17
G6pd	0.55	0.6	0.78	1.14 *	1.41 *	1.3 *	-0.09	-0.15	-0.07	-0.37	-0.19
Igfbp3	0.14	0.15	0.62	0.35	0.26	0.62 *	0.15	0.3	0.1	0.31	0.38
Junb	-0.08	-0.01	0.18	-0.16	-0.42	-0.9 *	0.41	0.31	0.62	0.07	-0.42
Me1	0.34	0.2	0.17	0.57 *	1.03 *	1.01 *	-0.25	-0.36	-0.71	-0.48	-0.26
Mog	-0.4	-0.46	-0.28	-0.47	-0.24	-0.56 *	-0.04	-0.14	-0.31	-0.15	-0.15
Pfkfb1	0.7	1.07	0.77	1.11 *	1.38 *	1.21 *	-0.3	-0.05	-0.54	-0.52	-0.1
RT1-N3	-0.26	-0.54	-0.31	-0.18	-0.41	-0.63 *	-0.31	-0.12	0.03	0.02	-0.02
Ccl2	-0.04	-0.19	-0.52	-0.45	-0.55 *	-0.58 *	0.1	0.03	-0.15	-0.24	-0.03
Slc10a1	0.08	0.21	0.23	0.17	0.29	0.29 *	-0.08	0.06	0.11	0.06	0.01
Tgfa	0.06	0.23	0.38	0.14	0.25	0.33 *	-0.02	0.05	-0.01	-0.12	-0.16
Ptn	0.62	0.55	0.61	0.67 *	0.39	0.64	0.45	0.52	0.76	0.69	0.47
Xrcc6	0.06	0.19	0.1	0.31	0.29	0.44 *	-0.04	-0.11	-0.15	-0.24	-0.21
Rbp1	-0.05	-0.42	-0.4	-0.15	-0.37	-0.45 *	0.12	0.05	-0.04	-0.03	0.07
Apoa1	0.07	0.32	0.53	-0.08	0.1	0.62 *	-0.01	-0.25	-0.21	-0.23	-0.21
Hsd11b1	0.08	1.25	1.43	0.91	1.12	1.32 *	-0.3	-0.13	-0.43	-0.46	-0.62
Abcc2	0.18	0.13	0.03	0.18	0.11	0.34 *	0.12	0.08	-0.11	-0.1	-0.12
Ephx1	0.33	0.15	0.38	0.52	0.59 *	0.73 *	0.04	-0.11	-0.32	-0.3	-0.07
Cpt2	0.03	0.26	0.29	0.17	0.22	0.32 *	-0.16	-0.05	-0.33	-0.12	-0.05
Cyp51	0.01	0.21	0.19	0.15	0.19	0.39 *	-0.19	-0.11	-0.11	-0.03	-0.1
Lgals9	-0.16	-0.47	-0.46	-0.08	-0.16	-0.42 *	-0.18	-0.21	-0.13	-0.18	-0.17
Ube2i	0.24	0.33	0.45	0.56	0.39	0.54 *	0.04	0.07	0.13	-0.03	0
G6pc	-0.73	-1.26	-0.78	-0.4	-0.88	-1.59 *	0.51	0.48	0.42	0.06	0.25
Igfbp2	-0.83	-1.2	-0.64	-1.33	-1.17	-1.14 *	0.44	0.08	0.13	-0.08	-0.28
Hmgcr	0.27	0.53	0.52	0.66	0.62	0.82 *	-0.3	-0.01	-0.19	0.09	-0.31
Igfbp1	-0.36	-0.67	-1.1	-1.06	-0.69	-1.58 *	0.65	0.36	0.2	0.12	0.4
Cpt1b	-0.04	-0.08	0.34	0.02	0.14	0.51 *	0.21	-0.14	0.12	0.05	-0.02
Dcn	0.17	0.49	1.03	0.52	0.12	0.99 *	0.19	0.21	0.44	0.61	0.4
Capns1	-0.04	0.11	0.31	0.2	0.22	0.53 *	-0.18	-0.17	-0.24	-0.43	-0.13
Inhba	-0.81	-0.71	-0.59	-1.17 *	-0.53	-0.01	0.13	-0.11	-0.3	-0.27	-0.12
Rcn2	-0.04	-0.1	-0.21	-0.15	-0.19	-0.41 *	0.09	-0.11	0.15	0.09	-0.05
Stmn3	0.01	0.29	0.51	0.29	0.65	0.81 *	-0.03	0.12	0.3	-0.19	0.07
Meox2	0.02	-0.39	-0.73	0.24	-0.43	-1.22 *	0.41	-0.05	0.01	-0.19	0.22
Chrne	-0.33	-0.38	-0.5	-0.16	-0.47 *	-0.27	0.1	-0.12	-0.1	-0.03	0.02
Hap1	0.01	-0.25	-0.29	-0.28	-0.19	-0.53 *	-0.09	-0.22	0.16	-0.09	0.05
Msra	-0.46	-0.78	-0.75	-0.56	-0.59	-0.85 *	-0.08	-0.09	-0.36	-0.05	0
Ighmbp2	0.05	0.08	0.24	0.15	0.16	0.38 *	-0.14	-0.14	-0.38	-0.13	-0.11
Fdft1	-0.02	0.13	0.29	0.16	0.14	0.39 *	-0.02	0.07	-0.02	0.13	-0.07
Ndst1	0.15	0.17	0.39	0.31	0.21	0.32 *	-0.09	0.01	-0.18	-0.27	-0.02
Hmgcs1	-0.03	0.33	0.49	0.14	0.24	0.64 *	-0.22	-0.08	-0.04	0.05	-0.29
Pdlim1	0.23	0.46	0.61	0.44	0.62	0.66 *	-0.1	-0.18	-0.53	-0.35	-0.24
Btg3	-0.06	0.12	0	0.09	0.03	0.3 *	-0.21	-0.19	-0.01	-0.01	-0.05
Car2	0.33	0.38	0.24	0.19	0.75 *	0.16	-0.02	-0.08	-0.25	-0.01	-0.02
Car3	-0.38	-0.46	-0.63	-0.4	-0.6	-1.57 *	0.74	0.62	0.45	0.17	0.27
Timm17a	0.13	0.1	0.18	0.1	0.29	0.32 *	-0.3	-0.33	-0.23	-0.16	-0.25
Adam15	-0.25	-0.43	-0.4	-0.36	-0.5	-0.8 *	0.27	-0.05	0.01	-0.14	-0.12
Cpq	-0.1	-0.28	-0.13	-0.29	-0.43 *	-0.42 *	0.18	0.14	0.29	0.34	0.18
Slc29a1	-0.14	-0.36	-0.28	-0.09	-0.18	-0.39 *	0.22	-0.08	0.15	0.08	0.07
Lrat	0.37	0.31	0.44	0.59	0.62	0.7 *	0.04	0.39	0.27	0.48	0.45
Dhcr7	-0.18	0.11	0.19	-0.04	0.2	0.46 *	-0.22	-0.21	-0.1	-0.17	-0.15

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Timm8b	0.02	0.14	0.16	0.06	0.17	0.21 *	-0.08	-0.08	-0.06	-0.03	0.02
Prpsap1	0.2	0.15	0.27	0.35	0.18	0.33 *	0	-0.04	0.05	-0.02	0.1
Plekhb1	-0.56	-0.71	-0.42	-0.67 *	-0.8 *	-0.35	-0.14	-0.09	0.25	-0.15	0.05
Synj2bp	0.04	0.11	0.23	0.1	0.12	0.27 *	0.08	0.06	-0.02	0.09	0.06
Slc13a3	0.5	0.83	1.08	0.85	0.95	1.16 *	0.23	0.79	0.43	0.29	0.68
Cldn1	-1.07	-0.57	0.01	-1.41 *	-0.76	0.17	-0.12	-0.11	-0.08	-0.16	-0.03
Rsad2	-0.04	-0.33	0.07	-0.09	-0.09	-0.69 *	-0.14	-0.12	0.21	0.22	0.01
Csnk1g2	0.12	0.2	0.17	0.21	0.15	0.34 *	-0.05	-0.01	0.04	0.02	-0.08
Aacs	0.23	0.67	0.73	0.59	0.77	1.32 *	-0.62	-0.42	-0.32	-0.24	-0.53
Bcam	0.2	0.4	0.31	0.37	0.29	0.37 *	-0.06	-0.02	-0.14	-0.05	0.04
Avil	0.21	0.28	0.2	0.39	0.3	0.48 *	0.18	0.1	0.19	0.28	0.11
Igfals	0.15	0.14	0.26	0.19	0.24	0.35 *	-0.04	-0.11	-0.06	-0.17	-0.04
Gna12	-0.11	0.13	0.36	0.08	0.13	0.33 *	0.1	0.07	0.1	-0.08	0.13
Mvd	-0.02	0.17	0.36	0.14	0.25	0.53 *	-0.2	-0.14	-0.18	-0.22	-0.32
Mvk	0.08	0.29	0.27	0.26	0.31	0.52 *	-0.27	-0.19	-0.12	-0.29	-0.22
Serpina7	-0.95	-1.07	-0.9	-1.87 *	-1.86 *	-1.49	0.28	0.18	0.68	0.56	0.36
Slc7a7	-0.73	-1.05	-0.93	-0.93 *	-1.12 *	-0.88	0	-0.29	0.01	0.05	-0.06
Nde1	0.03	0.11	0.09	0.13	0.18	0.26 *	0.01	-0.03	-0.09	0.07	-0.01
Crot	0.1	1.04	1.4	0.46	1.16 *	1.41 *	-0.31	-0.34	-0.26	-0.16	0.14
Irs3	-0.26	-0.02	0.38	0.37	0.77	1.01 *	0.4	0.01	-0.12	-0.57	0.25
Enpp2	0.68	0.61	0.4	0.78 *	0.73 *	0.53 *	0.04	-0.23	-0.24	-0.55	-0.31
Nrtn	-0.23	-0.31	-0.46	-0.3	-0.46 *	-0.56 *	0.01	-0.03	-0.13	0.01	0.15
Plcl1	0.04	0.18	0	0.14	0.13	0.62 *	0.11	0.15	0.04	0.15	0.25
Spa17	0.27	0.73	0.43	0.31	0.37	0.8 *	0.23	-0.2	-0.18	-0.38	-0.27
Pex11a	0.04	0.39	0.84	0.07	0.27	0.76 *	-0.28	-0.19	-0.28	-0.16	0.09
Col5a2	0.4	0.29	0.55	0.45	0.28	0.79 *	-0.05	-0.22	-0.02	0	-0.09
Slc44a1	-0.11	-0.2	-0.06	-0.16	-0.27	-0.24 *	-0.05	-0.1	0.06	0.06	0.04
Prkcdbp	0.04	0.17	0.22	0.17	0.26	0.39 *	-0.13	-0.08	0.04	-0.1	-0.02
Polg	0	0.17	0.4	0.04	0.23	0.37 *	-0.03	-0.09	-0.04	-0.16	0
Vamp7	0.19	0.09	0.16	0.17	0.11	0.27 *	-0.12	0.03	-0.1	-0.09	0
Cxadr	0.05	0.4	0.5	0.17	0.34	0.61 *	-0.18	-0.16	-0.18	-0.21	-0.27
Kif1c	0.09	0.2	0.41	0.26	0.22	0.48 *	-0.22	-0.32	-0.2	-0.2	-0.04
Hspb8	0.03	0.02	-0.05	0.28	0.26	0.47 *	-0.1	0.05	-0.21	0.07	0.04
Ankh	0.2	0.24	0.2	0.47	0.52	0.79 *	-0.22	-0.34	-0.4	-0.15	-0.16
Itga6	0.31	0.38	0.65	0.28	0.6	0.46 *	0.14	-0.03	0.27	0.1	0.04
Dusp1	-0.28	-0.55	-0.62	-0.45	-0.54	-1.23 *	0.51	0.38	0.56	0.17	0.28
Erp29	-0.01	0.28	0.39	0.19	0.2	0.51 *	-0.07	-0.04	-0.19	-0.05	-0.11
Cdh17	-0.78	-1.3	-1.57	-1.41 *	-1.37	-0.85	0.31	-0.97	-1.01	-0.55	-0.53
Sh3bp5	-0.17	-0.17	-0.13	-0.02	-0.31	-0.4 *	0.12	0.17	0.09	0.11	0.25
Ebp	0.05	0.15	0.2	0.12	0.2	0.35 *	-0.08	-0.07	-0.08	-0.11	-0.08
Txnip	-0.29	-0.71	-0.54	-0.35	-0.58	-0.9 *	0.34	0.17	0.43	0.23	0.19
Inadl	0.09	0.18	0.26	0.3	0.29	0.35 *	0.02	0.07	0.08	0.02	-0.06
Chek1	0.25	0.24	0.26	0.34 *	0.24	0.46 *	0.03	0.1	0.02	0.07	0.03
A1bg	-0.18	-0.48	-0.47	-0.54 *	-0.54 *	-0.49 *	-0.03	-0.01	0	0.08	0
Rcan2	-0.06	0.63	0.82	0.04	0.23	0.88 *	0.02	0.03	-0.16	-0.55	0.09
Abcc3	0.27	0.71	0.78	0.78	1.2	1.24 *	-0.07	-0.35	-0.3	-0.38	-0.16
Msmo1	-0.03	0.16	0.25	0.17	0.24	0.44 *	-0.19	0.01	-0.01	0.18	-0.14
Samsn1	0	-0.08	-0.04	-0.18	-0.23	-0.34 *	-0.26	-0.13	-0.26	-0.05	-0.05
Cbwd1	0.2	0.28	0.19	0.21	0.19	0.37 *	-0.1	0.1	-0.01	0.17	0.14
Hdgfrp2	0.16	0.22	0.28	0.28	0.21	0.3 *	-0.2	-0.05	-0.19	-0.07	-0.21
Rnf39	-0.02	-0.03	-0.11	-0.01	-0.2	-0.43 *	0.19	0.26	0.2	0.06	0.12

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Cyp2c22	0.06	0.31	0.41	0.05	0.44 *	0.65 *	-0.05	-0.28	-0.47	-0.48	-0.09
Cyp2c13	-0.3	1.51	1.35	0.92	1.11	1.72 *	0.03	-0.56	-0.15	-0.51	-0.01
Dnajc21	0.16	0.27	0.23	0.18	0.09	0.38 *	-0.09	-0.17	-0.04	-0.04	0.02
Cxcl10	-0.43	-0.36	-0.32	-0.52 *	-0.38	-0.76 *	-0.46	-0.55	-0.89	-0.66	-0.84
Ppap2c	0.05	0.46	0.57	0.31	0.47	0.85 *	-0.33	-0.01	-0.18	-0.22	0.12
Serpina4	-0.12	-0.24	-0.45	-0.17	-0.32	-0.37 *	0.11	-0.05	-0.05	-0.09	0.01
Mafk	-0.23	-0.25	-0.18	-0.04	-0.14	-0.35 *	-0.14	-0.09	-0.07	-0.23	-0.21
Lzts3	0.13	0.11	0.18	0.32	0.33	0.33 *	0.22	0.04	0.17	-0.01	0.22
Ppif	0.39	0.38	0.21	0.76 *	0.48	0.31	-0.04	-0.09	-0.04	-0.21	0
Gm2a	0.17	0.4	0.69	0.4	0.34	0.62 *	-0.01	-0.14	-0.01	-0.17	0
Atf5	-0.2	-0.03	0.23	-0.1	0.07	0.37 *	-0.08	-0.08	-0.03	-0.02	0.06
Mx2	-0.02	-1.35	-1.19	-0.46	-0.67	-1.53 *	-0.56	-0.9	-0.81	-0.52	-0.42
Cyp2b3	0.24	0.33	0.36	0.26	0.36	0.41 *	0.11	0.22	-0.03	0.13	0.18
Camk2n1	-0.36	-0.5	-0.33	-0.64 *	-0.73 *	-0.6 *	-0.05	0.03	0.03	-0.18	0.08
Rogdi	0.21	0.36	0.33	0.44 *	0.31	0.33	-0.02	0.08	-0.03	-0.14	0.19
Pgp	0.21	0.33	0.41	0.37 *	0.34 *	0.61 *	-0.16	-0.02	-0.13	-0.02	-0.04
Adprm	0.15	0.25	0.23	0.25	0.33	0.46 *	-0.23	-0.12	-0.11	-0.1	-0.14
Tlcd1	0.59	0.61	0.21	1.02	0.89	0.71 *	-0.33	-0.1	-0.34	0	-0.24
Ptplb	-0.13	0.1	0.14	0.05	0.08	0.28 *	-0.05	0.01	-0.01	0.03	0.04
Tpst2	-0.07	0.01	0.2	0.05	0.24	0.43 *	-0.14	-0.21	-0.27	-0.37	-0.11
Hmcn1	0.08	0.32	0.41	-0.08	0.26	0.39 *	0.23	0.08	0.16	0.24	0.27
Slamf9	-0.32	-0.3	-0.24	-0.14	-0.21	-0.58 *	-0.21	-0.26	-0.26	-0.15	-0.09
Eprs	0.04	0.11	0.26	0.21	0.13	0.29 *	0.05	0.04	0.2	0.11	0
Sowahb	0.3	0.24	0.09	0.33	0.35	0.55 *	-0.31	-0.08	-0.18	-0.15	-0.31
Pdhb	0.08	-0.18	-0.26	0.07	0.03	-0.37 *	0.1	0.08	0.16	0.09	-0.03
Tinf2	0	0.2	0.09	0.07	0.05	0.44 *	0.17	0	-0.03	0.19	0.03
Dct	-0.15	-0.17	-0.35	-0.25	-0.12	-0.37 *	0.06	-0.04	-0.34	0.05	-0.15
Isyna1	0.06	0.29	0.43	0.02	0.33	0.61 *	-0.12	-0.11	-0.13	0.04	0
Eci2	0.01	0.2	0.28	0.09	0.23	0.26 *	-0.08	-0.07	-0.12	-0.04	0
Gpr158	-0.28	-0.58	-0.67	-0.49	-0.58	-0.65 *	0.4	0.06	-0.22	-0.09	0.02
Nedd4l	0	0.15	0.4	0.05	0.16	0.39 *	0.02	-0.1	-0.01	-0.04	-0.17
Rprd1a	0.13	0.22	0.2	0.21	0.17	0.28 *	0.01	0.04	-0.04	0.1	0.02
RGD1311805	-0.04	0.1	0.21	0.07	0.12	0.28 *	-0.06	-0.02	0.06	0.03	-0.05
Setd6	0.19	0.28	0.25	0.19	0.28	0.41 *	-0.23	-0.36	-0.3	-0.29	-0.21
Aars	0.04	0.07	0.18	0.07	0.1	0.3 *	-0.11	-0.13	-0.09	-0.11	-0.12
Ypel3	-0.2	-0.41	-0.47	-0.38	-0.44	-0.44 *	0.06	0.12	0.04	0.09	0.15
Irf7	0	-1.04	-0.65	-0.27	-0.48	-1.23 *	-0.5	-0.55	-0.28	-0.16	-0.27
Actr1a	0.11	0.22	0.33	0.21	0.23	0.33 *	-0.01	0.04	-0.07	0.06	0.08
RT1-S3	-0.13	-0.75	-0.57	-0.21	-0.37	-0.71 *	-0.04	-0.08	0.08	0.12	0
Scube3	0.5	0.3	0.69	1.45 *	0.98	1.27 *	0.18	1.23	0.27	0.1	0.2
Supv3l1	0.01	0.16	0.24	0.26	0.11	0.36 *	-0.04	-0.05	0.04	0	-0.04
Man1a1	-0.1	-0.13	-0.13	-0.13	-0.23	-0.25 *	-0.09	0.01	0.04	0.02	0.11
Pcdh18	-0.63	-0.79	-0.79	-1.04 *	-0.95 *	-0.64 *	0.11	0.12	0.31	0.47	0.27
Mab21l3	-0.06	0.65	0.75	0.51	0.72	1.01 *	0.57	0.61	0.38	0.29	0.1
Dph5	0.09	0.29	0.22	0.12	0.26	0.36 *	-0.04	-0.14	-0.38	0.06	-0.27
RGD1309730	0.01	0.37	0.39	0.11	0.3	0.41 *	0.04	0.13	0.02	-0.01	0.07
Secisbp2l	0.08	0.21	0.2	0.14	0.34	0.45 *	-0.25	-0.2	-0.15	-0.3	-0.21
Ttpal	0.04	0.07	0.47	0.07	0.23	0.75 *	-0.2	-0.27	-0.09	-0.13	-0.07
Fam210b	0.24	0.35	0.22	0.35	0.43 *	0.43 *	-0.01	-0.04	0	-0.22	0.06
Psmd5	0.21	0.24	0.12	0.24	0.33	0.38 *	0.01	-0.02	-0.1	0.06	-0.01
Dusp16	0.08	0.1	0.17	0.35	0.11	0.42 *	-0.28	-0.21	-0.45	-0.13	-0.15

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Terf1	0.1	0.29	0.16	0.24	0.16	0.33 *	0.41	0.34	0.47	0.51	0.16
Lactb2	0.36	0.43	0.26	0.49 *	0.46 *	0.48 *	-0.2	-0.03	-0.1	-0.06	-0.13
Mybl1	0.06	0.21	0.17	0.09	0.37	0.94 *	-0.55	-0.58	-0.84	-0.65	-0.75
Mfsd2a	0.23	0.13	0.27	0.86	0.64	0.69 *	-0.78	-0.11	-0.48	-0.18	-0.07
Isg15	-0.15	-1.21	-0.77	-0.48	-0.7	-1.38 *	-0.37	-0.55	-0.4	-0.06	-0.18
Slc30a6	-0.04	0.1	0.34	0.34	0.27	0.52 *	-0.08	-0.02	-0.19	0.07	-0.09
Dhrs7	0.18	0.14	0.37	0.15	0.27	0.35 *	0.15	-0.03	0.07	-0.09	-0.06
Serpina6	-0.05	-0.46	-0.38	-0.11	-0.36	-0.47 *	0.09	0.04	-0.03	0.03	0.05
Mknk2	0.3	0.38	0.26	0.56	0.59 *	0.32	-0.26	-0.13	-0.11	-0.28	-0.15
Top1mt	-0.08	0.22	0.16	0.29	0.29	0.39 *	-0.2	0.01	-0.27	-0.3	-0.18
Hykk	-0.08	-0.14	-0.58	-0.09	-0.46	-1.05 *	0.28	0.88	0.24	0.52	0.56
Fam46a	-0.47	-0.72	-0.56	-0.76	-0.77	-0.94 *	0.03	-0.66	-0.35	-0.22	-0.31
Scap	0.1	0.31	0.31	0.3	0.32	0.52 *	-0.14	-0.13	-0.09	-0.15	-0.21
Crb3	0	0.28	0.31	0.18	0.35	0.31 *	0.03	-0.05	-0.04	-0.14	-0.07
Ankrd44	-0.15	-0.27	-0.29	-0.2	-0.28	-0.46 *	0.04	0.1	0.08	0.11	0.06
Tspan6	-0.11	-0.27	-0.29	-0.27	-0.19	-0.27 *	0.03	0.03	0	0.07	0.05
Mbnl3	-0.42	-0.46	-0.59	-0.41	-0.45	-0.7 *	0.14	0.17	0.17	0.2	-0.11
Tmem97	0.06	0.25	0.16	0.16	0.34	0.4 *	-0.15	-0.19	-0.17	-0.11	-0.05
Tmem98	0.12	0.33	0.6	0.35	0.46	0.55 *	-0.14	-0.12	-0.03	0	0.02
Lig3	0.01	0.18	0.22	0.14	0.24	0.39 *	-0.03	0.08	-0.06	-0.01	0.01
Stxbp4	0.42	0.62	0.44	0.66 *	0.64	0.56 *	0.12	0.24	0.05	0.25	0.13
Snx11	0.13	0.14	0.15	0.28 *	0.22	0.34 *	0.03	0.02	-0.03	0.08	0.03
Abca8a	0.01	-0.25	-0.19	-0.17	-0.25	-0.36 *	0.16	0.21	0.31	0.33	0.22
Ube2o	-0.18	-0.59	-0.36	-0.69 *	-0.57	-0.5	0.02	-0.23	-0.05	-0.07	-0.28
Engase	-0.26	-0.25	-0.23	-0.29	-0.26	-0.34 *	0.1	0.04	0.1	-0.05	0.02
Igf2bp2	-0.2	-0.41	-0.24	-0.19	-0.21	-0.57 *	-0.14	-0.09	-0.21	-0.23	-0.08
St3gal6	0.12	0.24	0.27	0.21	0.25	0.3 *	-0.02	-0.02	0.02	-0.06	0.06
Fam20c	0.33	0.38	0.26	0.53 *	0.44	0.37	-0.12	-0.03	0.01	0.08	-0.15
Zan	-0.33	-0.6	-0.37	-0.44	-0.86 *	-0.54	-0.17	-0.07	0.43	0.21	0.06
Oasl	-0.25	-0.92	-0.52	-0.44	-0.55	-1.13 *	-0.11	-0.46	-0.15	0.05	-0.26
Dcaf15	-0.1	-0.16	-0.28	-0.19	-0.2	-0.3 *	-0.07	-0.06	-0.11	-0.13	-0.13
Diexf	0.06	0.21	0.33	0.21	0.19	0.51 *	0.04	-0.1	-0.09	0.13	0.04
Osbp2	-0.11	-0.11	-0.07	-0.37 *	-0.18	-0.26	0.02	-0.14	-0.12	0.04	0.12
Comtd1	0.17	0.33	0.27	0.27	0.38	0.5 *	-0.04	-0.09	-0.05	-0.1	0
Abhd6	0.24	0.17	0.14	0.47	0.4	0.43 *	-0.16	-0.11	-0.09	-0.02	-0.08
Nek4	0.1	0.05	0.22	0.22	0.22	0.38 *	0.01	-0.05	-0.18	-0.11	0.08
Unc5d	0.06	-0.23	-0.62	-0.22	-0.52	-0.55 *	-0.24	-0.36	-0.05	-0.48	-0.47
Iars	0.02	-0.03	0.13	0.06	0.11	0.3 *	-0.03	-0.04	-0.05	0.14	-0.09
Akr1c19	-0.75	-0.55	-0.41	-0.53	-0.78 *	-0.28	-0.19	0.07	-0.06	0.19	0.03
Rbfa	-0.09	0.32	0.32	0.12	0.3	0.53 *	-0.1	0.11	-0.14	-0.12	-0.13
Lama3	0.82	0.73	0.14	0.97	0.66	0.6 *	-0.04	-0.46	-0.55	-0.29	0.05
Kifc3	0.19	0.63	0.7	0.57	0.84 *	0.89 *	-0.04	-0.17	-0.14	-0.16	-0.14
Def8	0.22	0.23	0.21	0.37 *	0.25	0.42 *	-0.25	-0.17	-0.13	-0.16	-0.09
Pdzd8	-0.21	-0.17	-0.23	-0.2	-0.33 *	-0.31	0.12	0.13	0.24	0.26	0.15
Cnksr3	-0.41	-0.65	-0.84	-0.66 *	-0.82 *	-0.79 *	0.02	-0.08	0.01	-0.31	-0.22
Lrrk1	-0.18	-0.29	-0.33	-0.32	-0.29	-0.43 *	0.19	0.12	0.1	0.03	-0.01
Uvrag	0.03	0.11	0.2	0.15	0.11	0.3 *	0.19	0.13	0.33	0.29	0.21
Egln1	0.14	0.16	0.18	0.15	0.12	0.41 *	-0.03	0.03	0.08	0.15	0.23
Cdr2	0.32	0.39	0.63	0.28	0.5 *	0.73 *	-0.08	0	-0.12	0.16	0.11
Dctn5	0.14	0.22	0.1	0.14	0.21	0.33 *	-0.14	-0.18	-0.11	0.02	-0.17
Slc22a18	0.04	0.22	0.23	0.21	0.15	0.29 *	-0.1	-0.04	-0.06	-0.15	-0.07

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Mrpl49	-0.09	0.01	0	-0.01	0.24	0.51 *	0.04	-0.19	-0.13	0.22	0.15
Nsdhl	0.11	0.31	0.29	0.3	0.37 *	0.38 *	-0.04	-0.05	0.02	0	-0.09
Ranbp6	0.15	0.18	0.16	0.25	0.32	0.43 *	-0.23	-0.03	-0.1	0.01	-0.16
Chuk	0.17	0.22	0.24	0.23	0.28	0.27 *	-0.1	-0.05	0.04	0.05	-0.04
Hps6	0.22	0.14	0.29	0.24	0.22	0.31 *	-0.13	-0.11	-0.01	-0.03	-0.01
Pcgf6	-0.02	0.23	0.31	0.21	0.19	0.39 *	-0.02	0.18	0.17	0.3	0.28
Osmr	-0.7	-1.31	-1.08	-1.09 *	-0.97 *	-1.13 *	-0.65	-0.69	-0.15	0.01	-0.52
Pmvk	-0.25	0.13	0.24	0.09	0.2	0.52 *	-0.09	-0.15	-0.2	-0.3	-0.22
Manba	0.31	0.32	0.55	0.31	0.41	0.49 *	-0.13	-0.12	-0.22	-0.11	-0.11
Rap1gds1	0.19	0.24	0.25	0.08	0.16	0.26 *	-0.04	-0.09	-0.02	0.11	0
Kbtbd4	0.19	0.25	0.26	0.27	0.26	0.29 *	0.08	0.15	0.03	-0.01	-0.01
Rrbp1	-0.03	-0.16	-0.27	-0.14	-0.17	-0.31 *	-0.02	-0.16	-0.12	-0.06	-0.11
Dolk	0.14	0.37	0.32	0.2	0.2	0.37 *	-0.1	-0.13	-0.12	-0.12	-0.12
Crat	-0.97	-0.56	-0.3	-0.91 *	-0.6	-0.09	-0.19	-0.26	-0.43	-0.44	-0.3
Stard8	-0.22	-0.26	-0.25	-0.2	-0.24	-0.38 *	0.03	0.04	0.01	0.13	-0.14
Cnot4	0.08	0.17	0.32	0.21	0.13	0.32 *	-0.16	-0.23	-0.03	0.1	0.02
Dennd2a	0.11	0.31	0.41	0.2	0.3	0.34 *	0.03	0.1	0.12	0	0.17
Usp18	0.22	-0.95	-0.45	-0.14	-0.46	-1.47 *	-0.25	-0.36	-0.22	0.15	-0.21
Xkr8	0.1	0.12	0.12	0.3	0.42	0.52 *	-0.16	-0.17	-0.14	-0.21	-0.25
Yars	0.08	0.15	0.22	0.15	0.25	0.41 *	-0.11	-0.13	-0.13	0.01	-0.07
Tbc1d2	0.11	0.12	0.43	0.38	0.41	0.64 *	-0.31	-0.28	-0.14	-0.63	-0.38
Coro2a	0.04	0.29	0.33	0.31	0.4	0.54 *	-0.09	0	-0.1	-0.11	-0.02
Wdr78	0.28	0.26	0.49	0.39	0.19	0.49 *	0.16	0.03	0.16	0.06	0.3
Tfap2e	-0.1	-0.29	-0.34	-0.17	-0.11	-0.39 *	-0.08	0.06	0.04	-0.07	0.06
Sptb	-0.11	-0.13	0.07	-0.2	-0.05	-0.29 *	0.21	0.06	0.08	0.01	0.14
Otub2	0.19	0.23	0.27	0.42	0.71 *	0.74 *	-0.02	0.15	-0.04	-0.19	-0.07
Tecpr2	0.05	0.19	0.27	-0.07	0	0.4 *	0.01	0	-0.03	0.13	0.13
Tmcc3	-0.04	0.22	0.36	-0.04	0.19	0.54 *	-0.08	-0.04	-0.19	-0.2	-0.06
Irak3	0.33	0.1	0.5	0.68 *	0.33	0.45 *	-0.01	0.34	0.44	0.5	0.34
Tob2	-0.11	-0.19	-0.29	-0.03	-0.18	-0.35 *	-0.13	-0.09	0.04	-0.05	-0.04
Sesn3	0.15	0.23	0.31	0.24	0.18	0.3 *	-0.04	0.02	-0.1	0.01	-0.1
Pih1d2	0.2	0.22	0.11	0.1	0.31 *	0.21	0.16	0.08	-0.02	0.18	0.02
Morc4	-0.17	-0.33	-0.64	-1.01 *	-0.46	-0.58	-0.11	0.11	-0.17	0.08	-0.44
Ccdc51	0.08	0.06	0.04	0.05	0.11	0.35 *	-0.06	-0.21	-0.31	-0.11	-0.13
Abhd5	0.5	0.36	0.58	0.49	0.68	0.83 *	0	-0.07	-0.04	-0.04	0.02
Rpl7l1	0	-0.06	-0.08	0.06	-0.12	0.3 *	-0.12	-0.1	-0.19	-0.14	-0.2
Nek6	-0.37	-0.42	-0.19	-0.65 *	-0.29	-0.12	-0.11	-0.26	-0.09	0.03	-0.26
Igfbp4	-0.29	-0.4	-0.18	-0.18	-0.48 *	-0.5	-0.02	0.01	0.04	0	0
Plekhh3	-0.08	-0.04	-0.27	-0.17	-0.17	-0.34 *	-0.13	-0.05	-0.15	0.08	0.06
Popdc2	-0.31	-0.41	-0.53	-0.5 *	-0.3	-0.26	0.38	-0.02	0.09	-0.25	0.19
Sdsl	0.15	0.32	0.33	0.21	0.25	0.3 *	-0.12	-0.05	-0.17	-0.36	-0.04
Noc4l	0.19	0.29	0.23	0.21	0.27	0.47 *	-0.17	-0.23	-0.28	-0.19	-0.12
Rgs16	0.04	-1.24	-1.75	-0.65	-0.43	-2.46 *	0.73	0.81	0.91	0.18	0.41
Tfdp1	0.22	0.19	0.19	0.24	0.31	0.34 *	0	-0.08	0.05	0.12	-0.01
Mocos	0.22	0.38	1.01	0.57	0.37	0.91 *	-0.54	-0.09	-0.13	-0.19	-0.18
Prrg2	0.04	0.24	0.31	0.38	0.27	0.41 *	0.02	0	0.1	-0.05	0.16
Zfp143	0.02	0.1	0.12	0.03	0.08	0.38 *	-0.06	0	0.04	0.09	0.01
Srpk1	0.09	-0.01	0.1	0.13	0.04	0.43 *	-0.05	-0.04	0.04	0.18	0.13
Srsf3	0.28	-0.03	0.04	0.32	0.34 *	0.28	-0.05	-0.2	-0.09	0.04	-0.37
Mynn	-0.26	-0.27	-0.14	-0.23	-0.38 *	-0.24	0.16	0.29	0.14	0.23	0.13
Agl	0.02	0.2	0.43	0.16	0.16	0.46 *	-0.02	-0.05	-0.13	0.04	-0.07

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
Creb3l1	0.03	0.48	0.36	0.2	0.54	0.8 *	0.11	0.1	-0.02	-0.02	0.06
Mpped2	-0.18	-0.26	-0.12	-0.25	-0.3	-0.47 *	0.01	0	-0.09	0.03	-0.06
Tp53inp2	0.11	0.21	0.19	0.16	0.24	0.32 *	-0.1	-0.06	-0.05	-0.07	0.03
Ncapd2	0.15	0.23	0.29	0.47	0.48	0.65 *	0.19	0.09	0.05	0.17	0.03
Cspp1	0.13	0.18	0.33	0.19	0.16	0.39 *	0.01	0.1	0.15	0.16	0.14
Tox	-0.25	-0.48	-0.73	-0.33	-0.52	-0.64 *	-0.45	-0.16	-0.49	-0.4	-0.23
Tmem55a	0.15	-0.27	-0.29	-0.46	-0.9 *	-0.63	-0.18	-0.2	0.14	-0.11	0.02
Ndufaf4	0.32	0.21	0.15	0.25	0.36	0.44 *	-0.1	-0.04	-0.04	0.23	-0.03
Snapc3	-0.13	-0.35	-0.34	-0.43 *	-0.37	-0.39	0.13	0.1	0.09	0.15	0.15
Gpn2	-0.05	0.08	0.2	0.11	0	0.34 *	-0.08	-0.23	-0.19	-0.05	-0.08
Mthfr	0.15	0.03	0.16	0.19	0.17	0.41 *	-0.22	-0.18	-0.16	-0.07	0.04
Rtcb	0.01	0.07	0.15	0.14	0.08	0.27 *	-0.03	-0.13	-0.1	-0.01	-0.07
Slc41a2	-0.07	-0.03	0.18	0.05	-0.05	0.37 *	-0.03	-0.06	-0.13	-0.11	-0.02
Creld2	-0.29	-0.36	-0.28	-0.5	-0.38 *	-0.26	-0.1	-0.17	-0.26	0.15	-0.1
Nmnat3	-0.28	-0.42	-0.53	-0.63 *	-0.44	-0.53 *	0.17	0.11	0.06	0.22	-0.34
Ube2f	-0.01	0.07	0.17	-0.02	0.11	0.27 *	-0.1	-0.13	-0.19	-0.05	-0.09
Plin4	-0.16	-0.36	-0.3	-0.26	-0.15	-0.45 *	-0.18	-0.06	-0.24	-0.27	-0.11
Pir	0.25	0.29	0.01	0.55	0.68	0.68 *	-0.31	-0.33	-0.48	-0.43	-0.29
Mid2	0.19	0.25	0.3	0.2	0.38 *	0.36 *	-0.03	0.13	0.1	0.06	-0.14
Grap	-0.24	-0.27	-0.17	-0.17	-0.12	-0.42 *	-0.05	0.07	-0.03	0.07	-0.04
Rnaseh2a	0.2	0.27	0.38	0.26	0.15	0.34 *	0.1	0.11	0.1	0.06	0.05
Ssh3	0.15	0.2	0.41	0.41	0.19	0.38 *	0.1	0.33	0.23	0.23	0.22
Pigv	-0.11	-0.15	-0.19	-0.33	-0.22	-0.31 *	0.07	0.23	0.1	0.24	0.21
Slc30a3	0.23	0.42	0.4	0.27	0.32	0.58 *	-0.16	-0.28	-0.16	-0.15	-0.21
Nfe2	0.38	1.12	1.39	1.35	1.7	2.34 *	-0.23	0.3	0.34	0.19	-0.42
Gjb2	-0.28	-0.4	-0.23	-0.44	-0.47 *	-0.43 *	-0.17	-0.11	-0.07	-0.11	-0.1
RT1-CE10	-0.29	-0.89	-0.77	-0.17	-0.47	-0.83 *	-0.09	-0.46	-0.35	-0.15	-0.2
Repin1	-0.17	-0.2	-0.13	-0.18	-0.21	-0.26 *	0.06	0.08	0.04	-0.05	0.05
Xk	0.2	0.41	0.62	0.33	0.31	0.64 *	-0.03	-0.06	-0.18	0.04	-0.18
Ern1	-0.28	-0.37	-0.16	-0.39 *	-0.35	-0.2	-0.02	-0.1	0.09	-0.02	-0.04
Fn3k	0.4	0.74	0.74	0.76 *	0.52	0.64 *	0.18	-0.04	0.11	0.13	0.05
Mapkapk5	0.08	0.16	0.26	0.27	0.26	0.42 *	-0.06	-0.19	-0.01	-0.08	-0.04
Commd6	0.34	0.37	0.29	0.44	0.92 *	0.95 *	0.02	-0.03	-0.05	-0.07	-0.02
Kxd1	0.05	0.22	0.33	0.29	0.38	0.39 *	0.04	-0.13	-0.02	-0.05	-0.08
Idnk	-0.53	-0.61	-0.32	-0.59 *	-0.36	-0.39	-0.2	-0.26	-0.5	-0.25	-0.42
Thnsl1	0.27	0.52	0.64	0.33	0.23	0.76 *	-0.04	-0.23	-0.25	-0.09	0
Myo7b	0.34	1.18	0.6	1.78 *	0.89	1.05	0.36	-0.34	0.44	0.5	0.41
Mrps11	-0.05	0.01	0.08	0.11	0.08	0.33 *	0.13	0	-0.03	0.02	0.12
Eid1	-0.1	-0.23	-0.31	-0.22	-0.2	-0.32 *	0.02	0.02	-0.09	-0.06	-0.08
Plekhd1	0.05	0.47	0.72	0.49	1.25	1.51 *	-0.52	-0.08	-0.16	-0.31	-0.05
RGD1566220	0.03	0.09	0.31	0.15	0.04	0.33 *	-0.2	-0.06	0.15	0.09	0.13
RGD1562114	0.14	0.26	0.33	0.25	0.25	0.37 *	-0.16	-0.29	-0.33	-0.15	-0.11
Lrrc14	0.2	0.25	0.26	0.23	0.2	0.41 *	0.05	-0.04	0.01	-0.02	-0.03
Arhgap42	0.09	0.11	0.22	0.22	0.29	0.39 *	-0.09	-0.04	-0.04	-0.12	-0.01
Bfsp2	-0.02	-0.03	-0.14	-0.05	-0.18	-0.31 *	-0.18	-0.18	-0.24	-0.03	-0.11
Stradb	0.19	0.22	0.19	0.22	0.19	0.27 *	-0.12	0.01	-0.04	-0.02	0.11
Rnf14	0.14	0.15	0.16	0.22	0.22	0.28 *	-0.02	-0.11	-0.01	0.02	0.04
Tmem69	0.16	0.29	0.12	0.3 *	0.32 *	0.3	-0.12	-0.23	-0.2	-0.11	-0.29
LOC678766	0.18	0.29	0.27	0.36	0.41 *	0.52 *	-0.19	-0.36	-0.34	-0.41	-0.36
Fer1l5	-0.03	0.75	1.77	0.5	0.79	1.75 *	0.05	0.23	0.58	0.82	0.82
Sin3b	-0.13	-0.21	-0.27	-0.17	-0.27	-0.25 *	0.02	-0.07	-0.03	-0.06	0

(Table continues)

	Differentially Expressed Genes in Female Rats Uniquely Significant for THS2.2 Group										
GENE ID	3R4F-08 90d	3R4F-15 90d	3R4F-23 90d	THS2.2-15 90d	THS2.2-23 90d	THS2.2-50 90d	3R4F-08 90+42d	3R4F-15 90+42d	3R4F-23 90+42d	THS2.2-15 90+42d	THS2.2-23 90+42d
LOC685078	-0.02	-0.85	-0.79	-0.49	-0.63	-1.03 *	-0.07	-0.22	-0.19	-0.03	-0.43
LOC685619	0.27	0.19	0.16	0.2	0.23	0.45 *	-0.1	0.14	-0.04	0.1	0.08
Cabyr	0	-0.02	-0.2	-0.14	-0.15	-0.28 *	-0.14	-0.08	0.01	-0.3	-0.21
Tprg1l	0.13	0.18	0.24	0.27 *	0.3 *	0.42 *	-0.14	-0.07	-0.04	-0.14	-0.05
LOC691807	0.09	0.13	0.07	0.14	0.2	0.26 *	-0.03	-0.04	-0.05	-0.09	-0.01
LOC691853	0.21	0.28	0.37	0.23	0.21	0.41 *	-0.22	-0.12	-0.16	-0.15	-0.26
Fam177a1	0.21	0.2	0.22	0.25	0.27	0.48 *	-0.12	0.04	0.01	0.07	0.04
Arxes2	-0.13	-0.22	-0.31	-0.25	-0.39 *	-0.2	0.28	0.1	0.05	0.1	0.26
Anapc16	-0.01	0.06	0.31	0.07	0.17	0.33 *	0.08	0.05	0.12	0.13	0.18
Cdkl5	0.18	0.18	0.25	0.26	0.18	0.4 *	-0.09	0.03	-0.06	-0.03	0.2
Snx21	0.14	0.28	0.39	0.15	0.31	0.32 *	0.01	-0.09	-0.13	-0.19	-0.06
LOC100363236	0.25	0.17	0.37	0.16	0.24	0.7 *	-0.04	-0.01	0.06	-0.02	0.08
Cystm1	0.07	0.31	0.37	0.31	0.44	0.61 *	-0.2	-0.21	-0.1	-0.02	0.03
LOC100911353	-0.18	0.2	0.29	-0.22	0.57	1.02 *	-0.41	-0.53	-0.6	-0.59	-0.62
LOC102550246	-0.04	-0.52	-1.04	-0.27	-0.72	-0.89 *	0.56	0.3	0.14	-0.18	-0.17
LOC102551341	-0.21	-0.06	-0.14	-0.1	-0.27	-0.31 *	0.06	0.18	-0.14	0.02	0.01
LOC102553866	-0.25	-0.17	-0.25	-0.14	-0.36 *	-0.21	-0.09	0	-0.13	0.05	-0.04
LOC103694321	-0.19	-0.33	-0.63	-0.42	-0.39	-0.5 *	-0.09	-0.18	-0.13	-0.06	-0.1

14.2 Supplementary Figures

N/A

15 APPENDIX

1. AMENDMENT TO THE OECD (+) ZURICH PROTEOMICS SP (150083)

STUDY PLAN AMENDMENT

AMENDMENT TO THE OECD (+) ZURICH PROTEOMICS SP (150083)

Program Name	SBDAV Zurich
Project Name	SBDAV-Animal Clinic
Work Package Name	
DUC Number	J-1002RDA0143/0001/0001
Author	Ashraf Elamin
Date:	28.04.2014

Review and Approval:

Name	Justification	Date / Signature
Elamin, Ashraf	Approved by Study Director (Supervisor MS-based Proteomics)	30.04.14
Thomas Schneider	Approved by Deputy Study Director (Scientist MS-based Proteomics)	30.04.14
Kogel, Ulrike	Approved by Work Package Leader (OECD +)	30.04.14
Nikolai Ivanov	Approved by Manager Research Technologies	30.04.2014
Hoeng, Julia	Approved by Sys Tox Program Leader	9.5.2014

Table of Contents

1 Purpose....3
2 Experimental Groups....3
3 iTRAQ LC MS/MS Approach....3
4 Timelines....4
5 Related Documents....4
6 Appendix....4

Revision History:

Version	Author	Description of Change	Type of change
1.0	Ashraf Elamin	Original issue	1

1. Major change/new version *2. Minor change* *3. Review without change*

1 Purpose

The main purpose of the amendment to the original signed SP is to perform non-targeted iTRAQ analysis on lung samples of female rat samples. This additional analysis is requested by System Toxicology Program leader and the work package leader so as to generate proteomics datasets that can be compared to the already analyzed lung tissues samples from male rat samples.

2 Experimental Groups.

All experimental groups included in the main study are described in the main study plan, section 4.3: 15006 THS SP.

Only some of the groups are subjects of the analyses included in this particular "omics" study plan (OECD+ part). They are summarized in **Table 1**.

Short Name (main study plan)	Group Code	Target nicotine (µg/L)	Time Point (days)	
			90	90+42
			female	female
SHAM	11	0	6	5
3R4F – low	12	8	-	-
3R4F – medium	13	15	-	-
3R4F – high	14	23	4	5
ZRH – low	15	15	-	-
ZRH – medium	16	23	6	6
ZRH – high	17	50	5	-

Table 1: Experimental groups and dosing

3 iTRAQ LC MS/MS Approach

In the iTRAQ approach, 90 days samples (left lungs from female rats) and 90 + 42 days samples (left lungs from female rats) will be processed and analyzed in one series of experiments. The experimental details are summarized below.

One reference (**Ref mix 1**) mixes will be prepared that represents a pooled samples from all the 90 and 90 + 42 sample sets. For the 90 and the 90 + 42 days samples (left lungs from female rats), six or five or four (depending

on the treatment sample) biological replicates of all treatment groups will be analyzed using iTRAQ 8-plex. Reporter-ion channel 113 will be used for Sham (90 days), 114 Ref mix 1, 115 for Sham (90+42), 116 for 3R4F high (90 days), 117 for 3R4F high (90+42), 118 for ZRH medium (90+42), 119 for Zurich medium (90 days), and 121 for Zurich high (90 days).

Each biological replicate will be processed twice (2 technical replicates) and each technical replicate will be injected twice on the Q-Exactive as described in WKI PMI_RD_WKI_001034:

The sample preparations and the LC MS/MS analysis of the samples will be performed as described in the below SP 15006_150083 THS SP

4 Timelines

The proposed timeline for the sample preparation, labeling with iTRAQ tags, and analysis by LC MS/MS will be **May 23rd 2014**.

5 Related Documents

iTRAQ 8-plex: PMI_RD_WKI_001034

15006_150083 THS SP: Omics analysis Study Plan

15006 THS SP: Study Plan 90 day OECD reg inhalation study 15006

6 Appendix

N/A