STUDY 150155 REPORT SYSTEMS TOXICOLOGY DATA FROM SAMPLES HARVESTED IN THE STUDY 15015: APOE$^{-/-}$ ZRH SWITCHING STUDY TO CHARACTERIZE THE TOXICITY OF THS2.2 IN APOE$^{-/-}$ MICE

Program Name	Systems Toxicology
Project Name	Animal Clinic
PUC Number	
Work Package Name	WP ZRH ApoE switching study
DUC Number	NA
Study Directors	Stephanie Boue
Author(s)	Stephanie Boue, Florian Martin, Bjoern Titz
Study Completion Date	Nov 2014

Table of Contents

Table of Contents 2
List of Figures 7
List of Tables 6
1 Introduction 12
2 Objectives 14
3 Experimental design 15
3.1 Test and reference items 15
3.1.1 Test item: MRTP Zurich FPD 4.2 sensorial medium 15
3.1.2 Reference item: 3R4F conventional cigarettes 15
3.2 Test system 15
3.3 Experimental groups and dosing 16
4 Procedures 17
4.1 Overview of sample generation for transcriptomics analysis 17
4.1.1 Lung 17
4.2 Sample retrieval from Singapore and biobanking 19
4.3 Sample randomization protocol 19
4.4 RNA preparation and whole genome expression arrays 19
4.5 Computational methods and analysis of transcriptomics data 20
4.5.1 Microarray data quality control (QC) and normalization 20
4.5.2 Network-based analysis 20
4.5.3 Biological impact factor (BIF) scoring 21
4.6 Mass spectrometry-based analysis of proteome changes 22
4.6.1 Sample preparation 22
4.6.2 Mass spectrometry for identification and quantification 22
4.6.3 Proteomic data analysis 23
4.6.3.1 Functional network analysis 23
4.6.4 Gene/protein set analysis 24
5 Results and Discussion 25
5.1 Summary omics endpoints 25
5.2 Nasal epithelium transcriptomics 26
5.2.1 QC and number of replicates per group 26
5.2.2 Differentially expressed genes in the RNE 26

5.2.3 BIF analysis in nasal epithelium tissue 30
5.2.4 Biological network perturbation amplitude analysis in nasal epithelium tissue 30
5.2.5 Hyp-level analysis: NPA1 34
5.2.6 NextBio analysis 36
5.2.6.1 3R4F exposure effect 36
5.2.6.2 THS2.2 exposure effects 37
5.2.7 Gene set enrichment analysis 37
5.2.7.1 3R4F effects 37
5.2.7.2 THS2.2 effects 39
5.2.7.3 Cessation vs. switch to THS2.2 effects 40
5.2.8 Gene expression of cytochrome P450 genes 41
5.2.9 Comparison with previous Apoe$^{-/-}$ SMAR findings 41
5.3 Lung transcriptomics 44
5.3.1 QC and number of replicates per group 44
5.3.2 Differentially expressed genes in the lung 44
5.3.3 BIF analysis in the lung 48
5.3.4 Biological network perturbation amplitude analysis in the lung 48
5.3.5 Hyp-level analysis: NPA1 52
5.3.6 NextBio analysis 54
5.3.6.1 3R4F exposure effect 54
5.3.6.2 THS2.2 exposure effects 55
5.3.7 Gene set enrichment analysis 55
5.3.7.1 3R4F effects 55
5.3.7.2 THS2.2 effects 57
5.3.7.3 Cessation vs. switch to THS2.2 effects 58
5.4 Lung proteomics 59
5.4.1 Number of replicates per group and protein identification 59
5.4.2 Differentially expressed proteins in the lung 59
5.4.3 Exposure effect comparison for proteins strongly affected by 3R4F exposure 63
5.4.4 Analysis of activated functional protein networks 66
5.4.5 Gene/protein set analysis 69
5.4.6 Summary of the observed effects on the lung proteome 70
5.5 Thoracic aorta transcriptomics 71

5.5.1 QC and number of replicates per group 71
5.5.2 Differentially expressed genes in the thoracic aorta 72
5.5.3 BIF analysis in the thoracic aorta 75
5.5.4 Biological network perturbation amplitude analysis in the thoracic aorta 75
5.5.5 Hyp-level analysis: NPA1 78
5.5.6 NextBio analysis 79
5.5.6.13R4F exposure effect 79
5.5.7 Gene set enrichment analysis 80
5.5.7.13R4F effects 80
5.5.7.2THS2.2 effects 82
5.5.7.3Cessation vs. switch to THS2.2 effect 83
5.6 Liver transcriptomics 84
5.6.1 QC and number of replicates per group 84
5.6.2 Differentially expressed genes in the liver 84
5.6.3 BIF analysis in the liver 87
5.6.4 Biological network perturbation amplitude analysis in the liver 88
5.6.5 Hyp-level analysis: NPA1 91
5.6.6 NextBio analysis 92
5.6.7 Gene set enrichment analysis 94
5.6.7.13R4F effects 94
5.6.7.2THS2.2 effects 96
5.6.7.3Cessation vs. switch to THS2.2 effect 97
5.7 Liver proteomics 97
5.7.1 Number of replicates per group and protein identification 97
5.7.2 Differentially expressed proteins in the liver 98
5.7.3 Exposure effect comparison for proteins strongly affected by 3R4F exposure 103
5.7.4 Analysis of activated functional protein networks 106
5.7.5 Gene/protein set analysis 110
5.7.6 Summary of the observed effects on the liver proteome 111
5.8 Heart transcriptomics 112
5.8.1 QC and number of replicates per group 112
5.8.2 Differentially expressed genes in the heart 112
5.8.3 BIF analysis in the heart 115

5.8.4 Biological network perturbation amplitude analysis in the heart 116
5.8.5 Hyp-level analysis: NPA1 119
5.8.6 NextBio analysis 121
5.8.6.13R4F exposure effect 121
5.8.6.2THS2.2 exposure effects 122
5.8.7 Gene set enrichment analysis 122
5.8.7.13R4F effects 122
5.8.7.2THS2.2 effects 124
5.8.7.3Cessation vs. switch to THS2.2 effect 125
5.9 Kidney transcriptomics 126
5.9.1 QC and number of replicates per group 126
5.9.2 Differentially expressed genes in the kidney 126
5.9.3 BIF analysis in the kidney 129
5.9.4 Biological network perturbation amplitude analysis in the kidney 130
5.9.5 Hyp-level analysis: NPA1 133
5.9.6 NextBio analysis 134
5.9.7 Gene set enrichment analysis 136
5.9.7.13R4F effects 136
5.9.7.2THS2.2 effects 138
5.9.7.3Cessation vs. switch to THS2.2 effect 139
6 Conclusion 140
7 Abbreviations 141
8 Appendices 143
9 References 143

List of Tables

Table 1: Number of samples analyzed for each tissue, group, and time point 18

Table 2: Number of valid replicates of nasal epithelium transcriptomics profiles 26

Table 3: Number of differentially expressed genes in the nasal epithelium (up-/down-regulated compared with sham at each respective time point) 28

Table 4: Number of valid replicates of lung transcriptomics profiles 44

Table 5: Number of differentially expressed genes in the lung (up-/down-regulated compared with sham at each respective time point) 46

Table 6: Number of valid replicates of lung proteomics profiles 59

Table 7: Number of consistently quantified lung proteins 59

Table 8: Number of valid replicates of thoracic aorta transcriptomics profiles 71

Table 9: Number of differentially expressed genes in the thoracic aorta (up-/down-regulated compared with sham at each respective time point) 73

Table 10: Number of valid replicates of liver transcriptomics profiles 84

Table 11: Number of differentially expressed genes in the liver (up-/down-regulated compared with sham at each respective time point) 86

Table 12: Number of valid replicates of liver proteomics profiles 98

Table 13: Number of consistently quantified liver proteins 98

Table 14: Number of valid replicates of heart transcriptomics profiles 112

Table 15: Number of differentially expressed genes in the heart (up-/down-regulated compared with sham at each respective time point) 114

Table 16: Number of valid replicates of kidney transcriptomics profiles 126

Table 17: Number of differentially expressed genes in the kidney (up-/down-regulated compared with sham at each respective time point) 128

List of Figures

Figure 1: Study design for the Apoe$^{-/-}$ ZRH switching study 15
Figure 2: Liver dissection specification 17
Figure 3: Systems biology endpoints 25
Figure 4: Volcano plots representing the systems response profiles in the RNE: FDR 27
Figure 5: Volcano plots representing the systems response profiles in the RNE: raw *P*-values 28
Figure 6: Overlap between sets of differentially expressed genes in the RNE 29
Figure 7: BIF analysis in the RNE 30
Figure 8: Star plots showing BIF calculated from RNE tissue 31
Figure 9: Heatmap of network perturbation amplitude scores in the RNE 32
Figure 10: Overlap between sets of significantly perturbed networks in the RNE 33
Figure 11: Network perturbation amplitude scores for the xenobiotic metabolism response network 34
Figure 12: Heatmap of NPAI score for all hyps significant in at least one contrast in the RNE 35
Figure 13: NextBio meta-analysis of 3R4F effects on the RNE 36
Figure 14: NextBio meta-analysis of THS2.2 effects on the RNE 37
Figure 15: Gene set enrichment analysis in the 3R4F-exposed RNE 38
Figure 16: Gene set enrichment analysis in the THS2.2-exposed RNE 39
Figure 17: Gene set enrichment analysis in the RNE of animals from the cessation and switch to THS2.2 groups 40
Figure 18: Fold-change of cytochrome P450 gene expression levels in the RNE compared with sham at each time point 41
Figure 19: Xenobiotic metabolism pathway: comparison with a historical study (2) 42
Figure 20: Genes found to be similarly up-regulated in the RNE and lung as in a previous study (Fig. 4B from (2)) 43
Figure 21: Volcano plots representing the systems response profiles in the lung: FDR 45
Figure 22: Volcano plots representing the systems response profiles in the lung: raw *P*-value 46
Figure 23: Comparison of differentially expressed genes between contrasts 47
Figure 24: BIF analysis in the lung 48
Figure 25: Star plots showing BIF calculated from the lungs 49

Figure 26: Heatmap of network perturbation amplitude scores in the lung 50
Figure 27: Overlap between sets of significantly perturbed networks in the lung..... 51
Figure 28: Network perturbation amplitude scores for the xenobiotic metabolism response network 52
Figure 29: Heatmap of NPAI score for all hyps significant in at least one contrast in the lung 53
Figure 30: NextBio meta-analysis of 3R4F effects in the lung 54
Figure 31: NextBio Meta-analysis of THS2.2 effects in the lung 55
Figure 32: Gene set enrichment analysis in the 3R4F-exposed lung 56
Figure 33: Gene set enrichment analysis in the THS2.2-exposed lung 57
Figure 34: Gene set enrichment analysis in the lungs of animals from the cessation and switch to THS2.2 groups 58
Figure 35: Volcano plots representing the proteome response profiles in the lung ... 60
Figure 36: Numbers of differentially expressed proteins in the lung 61
Figure 37: Numbers of differentially expressed proteins over time in the lung 61
Figure 38: Overlap between sets of differentially expressed proteins in the lung 62
Figure 39: Heatmap of differentially expressed proteins in the lung 63
Figure 40: Fold-change expression matrix for proteins strongly affected by 3R4F exposure 64
Figure 41: Boxplots showing the protein expression response of selected proteins from Figure 40 65
Figure 42: Functional interaction network for significantly regulated proteins in the lung after 8 months of 3R4F exposure 66
Figure 43: Comparison of exposure effects on the identified functional network across all exposure conditions (signed log10 adjusted *P*-values) 67
Figure 44: Comparison of exposure effect on the identified functional network across all exposure conditions (log2 fold-changes) 68
Figure 45: Summary of exposure effects on the identified protein clusters 69
Figure 46: Protein set analysis for the lung 70
Figure 47: Volcano plots representing the systems response profiles in the thoracic aorta: FDR 72
Figure 48: Volcano plots representing the systems response profiles in the thoracic aorta: raw *P*-value 73
Figure 49: Overlap between sets of differentially expressed genes in the thoracic aorta 74
Figure 50: BIF analysis in the thoracic aorta 75

Figure 51: Star plots showing BIF calculated from the thoracic aorta 76

Figure 52: Heatmap of network perturbation amplitude scores in the thoracic aorta. 77

Figure 53: Overlap between sets of significantly perturbed networks in the thoracic aorta 78

Figure 54: Heatmap of NPAI score for all hyps significant in at least one contrast in the aorta 79

Figure 55: NextBio meta-analysis of 3R4F effects in the thoracic aorta 80

Figure 56: Gene set enrichment analysis in the 3R4F-exposed thoracic aorta 81

Figure 57: Gene set enrichment analysis in the THS2.2-exposed thoracic aorta 82

Figure 58: Gene set enrichment analysis in the aorta of animals from the cessation and switching to THS2.2 groups 83

Figure 59: Volcano plots representing the systems response profiles in the liver: FDR 85

Figure 60: Volcano plots representing the systems response profiles in the liver: raw *P*-value 86

Figure 61: Overlap between sets of differentially expressed genes in the liver 87

Figure 62: BIF analysis in the liver 88

Figure 63: Star plots showing BIF calculated from the liver 89

Figure 64: Heatmap of network perturbation amplitude scores in the liver 90

Figure 65: Overlap between sets of significantly perturbed networks in the liver 91

Figure 66: Heatmap of NPAI score for all hyps significant in at least one contrast in the liver 92

Figure 67: NextBio meta-analysis of the effects of 3R4F on the liver 93

Figure 68: NextBio meta-analysis of the effects of THS2.2 on the liver 94

Figure 69: Gene set enrichment analysis in the 3R4F-exposed liver 95

Figure 70: Gene set enrichment analysis in the THS2.2-exposed liver 96

Figure 71: Gene set enrichment analysis in the livers of animals from the cessation and switch to THS2.2 groups 97

Figure 72: Volcano plots representing the proteome response profiles in the liver 99

Figure 73: Numbers of differentially expressed proteins for each comparison in the liver 100

Figure 74: Numbers of differentially expressed proteins in the liver during the exposure time 101

Figure 75: Overlap between sets of differentially expressed proteins in the liver 102

Figure 76: Heatmap of differentially expressed proteins in the liver 103

Figure 77: Fold-change expression matrix for proteins strongly affected by 3R4F exposure 104

Figure 78: Boxplots of the protein expression response of selected xenobiotic enzymes from Figure 77 106

Figure 79: Functional interaction network for proteins significantly regulated in the liver after 6 months of 3R4F exposure 107

Figure 80: Comparison of exposure effect on the identified functional network across all exposure conditions (signed log10 adjusted *P*-values) 108

Figure 81: Comparison of exposure effect on the identified functional network across all exposure conditions (log2 fold-changes) 109

Figure 82: Summary of exposure effects on the identified protein clusters 110

Figure 83: Protein set analysis for the liver 111

Figure 84: Volcano plots representing the systems response profiles in the heart: FDR 113

Figure 85: Volcano plots representing the systems response profiles in the heart: raw *P*-value 114

Figure 86: Overlap between sets of differentially expressed genes in the heart 115

Figure 87: BIF analysis in the heart 116

Figure 88: Star plots showing BIF calculated from the heart 117

Figure 89: Heatmap of network perturbation amplitude scores in the heart 118

Figure 90: Overlap between sets of significantly perturbed networks in the heart 119

Figure 91: Heatmap of NPAI score for all hyps significant in at least one contrast in the heart 120

Figure 92: NextBio meta-analysis of 3R4F effects in the heart 121

Figure 93: NextBio meta-analysis of THS2.2 effects in the heart 122

Figure 94: Gene set enrichment analysis in the 3R4F-exposed heart 123

Figure 95: Gene set enrichment analysis in the THS2.2-exposed heart 124

Figure 96: Gene set enrichment analysis in the hearts of animals from the cessation and switch to THS2.2 groups 125

Figure 97: Volcano plots representing the systems response profiles in the kidney: FDR 127

Figure 98: Volcano plots representing the systems response profiles in the kidney: raw *P*-value 128

Figure 99: Overlap between sets of differentially expressed genes in the kidney 129

Figure 100: BIF analysis in the kidney 130

Figure 101: Star plots showing BIF calculated from the kidney 131

Figure 102: Heatmap of network perturbation amplitude scores in the kidney 132

Figure 103: Overlap between sets of significantly perturbed networks in the kidney 133

Figure 104: Heatmap of NPAI score for all hyps significant in at least one contrast in the kidney 134

Figure 105: NextBio meta-analysis of the effects of 3R4F in the kidney 135

Figure 106: NextBio meta-analysis of the effects of THS2.2 on the kidneys 136

Figure 107: Gene set enrichment analysis in the 3R4F-exposed kidneys 137

Figure 108: Gene set enrichment analysis in the THS2.2-exposed kidneys 138

Figure 109: Gene set enrichment analysis in the kidneys of animals from the cessation and switch to THS2.2 groups 139

1 Introduction

It has been hypothesized that the chemical composition and toxicological effects of the aerosol generated by burning or heating tobacco depend to a great extent on temperature (1). The current version of the second generation of tobacco heating systems, THS2.2, which operates at temperatures below 350°C, generates reduced amounts of harmful and potentially harmful constituents (HPHCs) relative to the Kentucky Reference cigarette, 3R4F, under smoking machine conditions.

A pre-clinical assessment program was previously performed on THS2.2 to determine whether adult smokers using THS2.2 are exposed to increased or new hazards compared with those continuing to smoke conventional cigarettes (CCs). The rationale for this pre-clinical program was based on the United States Food and Drug Administration (FDA) Center for Tobacco Product draft guidance (FDA 2012) and included an evaluation of the toxicological properties of materials used to make the product, evaluation of the aerosol through non-targeted chemical screening and measurement of droplet size, measurement of HPHCs, assessment of the genotoxicity and cytotoxicity of the aerosol compared with CCs, and a 90-day inhalation study on rats. *In vitro* tests and the rat inhalation study showed a reduced risk potential for THS2.2 compared with CCs. Results are summarized in the Organisation for Economic Co-operation and Development study report and systems toxicology report (2). Briefly, the rat inhalation study showed that animals exposed to THS2.2 showed much less toxicity signs than those exposed to 3R4F. The biological effects caused by the test item were limited mainly to hematological changes, blood chemistry alterations, and changes in liver weight, which have been shown to be related to nicotine exposure. Lung inflammation and adaptive changes in the histopathology of respiratory tract organs in THS2.2-exposed rats were dramatically lower than those in 3R4F-exposed rats, and were consistent with gene expression and proteomics profiling data in the lung that were suggestive of reduced perturbation of toxicity pathways. In order to go further, i.e. study disease risk in a long-term exposure setup and to benchmark the effects to switching to THS2.2 to smoking cessation, an eight-month mouse switching study was setup.

We previously showed that Apoe$^{-/-}$ mice are an adequate model to study lung-related and cardiovascular-related disease risk in a pre-clinical setting. This earlier study (P13882/P13938) lasted 6 months and tested an early prototype of Yverdon (SMoking ARticle Master Design 2 with Aluminum Disc coating (SMAR MD2AD)) in a 'switching-type' experiment. Following a 3-month-long initial exposure to the reference cigarette 3R4F, exposure to this prototypic modified risk tobacco product (pMRTP) was compared with sham exposure and 3R4F exposure. Each month, mice were necropsied and an array of atherogenesis- and lung-related endpoints was measured. Overall, the switch to SMAR MD2AD resembled the switch to sham (mimicking a cessation protocol), i.e. a number of endpoints representing accrued disease risk either stopped increasing after the switching point, or returned, more or less rapidly, to levels observed in the sham group. This highlighted the potential of heating versus burning to reduce exposure-related risks. Because the mice were not exposed to SMAR for the entire duration of

the study, and because some of the effects of 3 months smoking were not fully reversible with a 3-month cessation period, the present study used a longer study length and set the switching point to 2 months. Therefore, herein, we report an 8-month inhalation switching study to determine the long-term exposure disease risk and to benchmark the effects of switching to THS2.2 compared with smoking cessation. Female Apoe$^{-/-}$ mice were exposed to mainstream aerosol from THS2.2 or 3R4F generated under the Health Canada Intense smoke protocol. This report summarizes the systems toxicology-enabled analyses.

2 Objectives

The aim of this study was to characterize the chronic inhalation toxicity of the aerosol generated from the candidate MRTP THS2.2 and to compare it with that of the Kentucky Reference Cigarette 3R4F.

The endpoints investigated included clinical examinations such as general health status, in-life observations, body weight development, food consumption during and following exposure, and ophthalmologic assessment before and after exposure. Upon necropsy, gross pathology (including organ weights) and histopathology were assessed in respiratory and non-respiratory tract organs, and urinalysis and blood analysis (hematology and clinical chemistry) were performed. Bronchoalveolar lavage fluid was also collected to enumerate free lung cells.

A systems toxicology study was performed to gain a mechanistic understanding in which additional animals were concomitantly exposed in accordance with those in the above-mentioned study, i.e., using the same exposure chambers and regimen. This report focuses on the analyses of transcriptomics and proteomics data generated in this study. For other results, please consult the main study report (3).

The primary objectives of the transcriptomics analyses were to generate gene expression data from respiratory nasal epithelium, lung, liver, thoracic aorta, heart, and kidney from the Apoe-ZRH switching study (Study #15015) to answer the following scientific questions:

- What are the effects of the candidate MRTP (1–8 months of exposure, MRTP group), compared with those of 3R4F and sham on cardiovascular and chronic obstructive pulmonary disease-related endpoints and transcriptomics?
- Are biological changes induced by 3R4F at the transcriptome or proteome levels reversible following cessation or switching to the MRTP, and to what extent?
- Is it possible to correlate inflammation endpoints with lung transcriptional profiles, and what is the agreement between observed biological endpoints and omics data?

An additional objective was to reproduce/verify results obtained in the Apoe$^{-/-}$ SMAR study (P13938) that were published previously (4-7).

3 Experimental design

Five groups of female Apoe$^{-/-}$ mice were exposed to one of the following treatments for up to 8 months (Figure 1): fresh filtered air (sham), mainstream 3R4F smoke, mainstream THS2.2 aerosol (candidate MRTP), or a switching protocol to fresh air (cessation group) or THS2.2 (switching group) following an initial 2-month exposure to 3R4F.

Animals were exposed for 3 h per day, 5 days per week, and were necropsied after 1, 2, 3, 6, or 8 months of exposure.

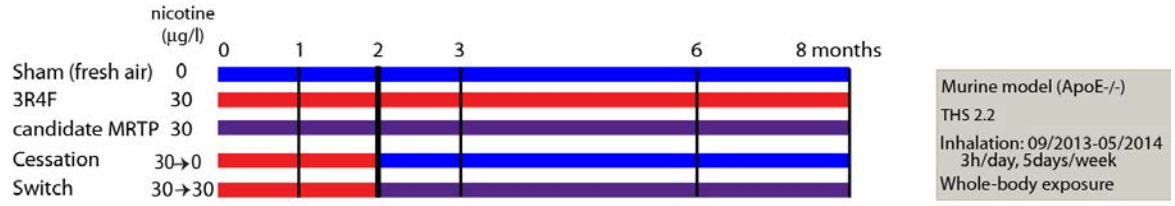


Figure 1: Study design for the Apoe$^{-/-}$ ZRH switching study

3.1 Test and reference items

Test items are described in detail in the Main Study Plan

3.1.1 Test item: MRTP Zurich FPD 4.2 sensorial medium

The candidate MRTP used in this study was the THS2.2 (Zurich FPD 4.2) sensorial medium (flavor system Reynaldo, batches number B-05771, B-06132, and B-08163), consisting 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 nicotine delivery. THS2.2 was produced by Philip Morris International, Neuchâtel, Switzerland.

THS2.2 sticks were conditioned according to International Organization for Standardization (ISO) 3308 (8) prior to use.

3.1.2 Reference item: 3R4F conventional cigarettes

The reference cigarette, 3R4F, was obtained from the University of Kentucky (9) and used as a comparison in the present study. 3R4F cigarettes were conditioned according to ISO 3308 (8) prior to use.

3.2 Test system

Experiments were conducted on female ApoE$^{-/-}$ mice (TC-ApoE-F) bred under specific pathogen-free conditions (Taconic Biosciences, Germantown, NY, USA).

3.3 Experimental groups and dosing

Experimental groups included in the main study are described in the main study plan.

The exposure phase began with an adaptation period lasting for 6 days as follows:
Study days 1–2: target 3R4F total particulate matter (TPM) concentration set to 150 mg/m^3
Study days 3–4: target 3R4F TPM 300 mg/m^3
Study days 5–6: target 3R4F TPM 450 mg/m^3
From study day 7, the target concentration was set to 600 mg/m^3.

Food was obtained from Harlan Laboratories (Blackthorn, UK). A gamma-irradiated (T2914C) rodent pellet diet was used as chow.

Eight animals per exposure group and time point were dedicated to omics analyses.

4 Procedures

4.1 Overview of sample generation for transcriptomics analysis

Animal dissection and sample collection details are described in an amendment to the study plan.

Organs were collected 16–24 h after the last exposure to smoke, and subjected to microarray analysis for the generation of gene expression profiles by an Affymetrix chip (Santa Clara, CA, USA). The following samples were collected: liver (piece #1a, Figure 2), respiratory nasal epithelium (RNE), left lung lobe, thoracic aorta, lower heart, and kidney (months 6 and 8).

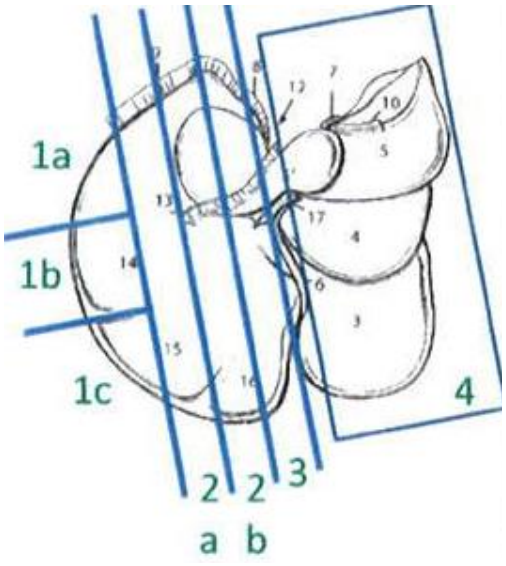


Figure 2: Liver dissection specification

4.1.1 Lung

Lung samples were also analyzed by genomics and proteomics. The left lung lobe was cryosliced by the Tissue Research Laboratory into 20 µm thick slices to ensure availability of homogeneous samples. Slices were placed in rotating order into tubes for genomics and transcriptomics analysis (and a reserve that was used for proteomics if the transcriptomics analysis was successful) until the entire lung lobe was sliced.

The process was performed according to the following standard operating procedure for tissue cryoslicing and laser capture microdissection, 'Preparation of cryosections with cryostat'.

Transcriptomics slices were placed in 2 ml MagNA Lyser Green Beads tubes (Roche Cat. N° 03 358 941 001) containing ceramic beads for homogenization in 600 µl RLT buffer (Qiagen, Hombrechtikon, Switzerland; Cat. N° 79216) and β-mercaptoethanol (Sigma-Aldrich, St. Louis, MO, USA), vortexed for 30 s, and frozen on dry ice until required for RNA extraction.

The number of replicates used for each tissue and time point is summarized in Table 1.

Table 1: Number of samples analyzed for each tissue, group, and time point

Exposure group	Time point (months)	RNE	Lung	Liver	Thoracic aorta	Kidney	Heart
Sham	1	8	8	8	6		8
	2	8	8	8	8		8
	3	8	8	8	8		8
	6	8	8	8	8	8	8
	8	8	8	8	8	8	8
3R4F	1	8	8	8	6		8
	2	8	8	8	8		8
	3	8	8	8	8		8
	6	8	8	8	8	8	8
	8	8	8	8	8	8	8
Candidate MRTP	1	8	8	8	6		8
	2	8	8	8	8		8
	3	8	8	8	8		8
	6	8	8	8	8	8	8
	8	8	8	8	8	8	8
Cessation	3	8	8	8	8		8
	6	8	8	8	8	8	8
	8	8	8	8	8	8	8
Switch	3	8	8	8	8		8
	6	8	8	8	8	8	8
	8	8	8	8	8	8	8
		168	168	168	162	80	168

4.2 Sample retrieval from Singapore and biobanking

Samples and relevant information from the animals' laboratory information management system Provantis were registered in the BSR Biobanking system. Both the dissection and the shipment were tracked. Once the samples arrived in Neuchâtel, the content was checked against the shipment request and confirmed.

The storage location of the samples in Neuchâtel was registered and a pickup manifest was approved by the sample manager before samples were released for processing in the transcriptomics or proteomics labs. The Biobanking system assisted with capturing and tracking the randomization process and printing unique sample labels according to their randomization order to simplify and error-proof laboratory handling. Samples were sorted using liquid nitrogen and dry ice to prevent thawing.

4.3 Sample randomization protocol

For RNE, thoracic aorta, heart, and kidney tissues, the full set of samples was randomized together to ensure a balanced assignment within/across experimental groups to the RNA extraction and Affymetrix hybridization stations.

For the liver, samples were processed in two batches: months 1–3, then months 6 and 8. Aside from the time point batches, which were analyzed separately, all groups were randomized together.

Slicing of lung lobes was performed in two batches: months 1–3, then months 6 and 8. However, RNA extraction and microarray processing were performed all together and fully randomized.

4.4 RNA preparation and whole genome expression arrays

The miRNeasy Mini Kit (Qiagen, Cat. N° 217004) was used to extract and purify total RNA, including micro RNA, except for thoracic aorta for which the miRNeasy micro kit (Qiagen Cat #N° 217084) was used. RNA was dissolved in an appropriate volume of RNAse-free water and quantified by a NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The quality was verified by an Agilent 2100 Bioanalyzer profile (Agilent Technologies, Palo Alto, CA, USA). An RIN number >6 was required for an acceptable quality. For mRNA analysis, 100 ng of total RNA was processed as described in the GeneChip® High Throughput 3′ In Vitro Transcription (HT 3′ IVT) Express User Manual (Affymetrix). The GeneChip® Mouse Genome 430 Array (Affymetrix) was used for hybridization. All RNA samples were labelled using HT 3′ IVT, hybridized to an array, which was then washed and

stained with Fluidic Stations FS450 using Affymetrix GeneChip® Command Console™ (AGCC) Software and protocol FS450_0001. Each array was scanned with the GeneChip® Scanner 3000 7G. Raw images data from the scanner were saved in a DAT file, which were then converted to CEL files containing pixel intensity values collected from the scanner. CEL files were monitored using the AGCC application.

4.5 Computational methods and analysis of transcriptomics data

4.5.1 Microarray data quality control (QC) and normalization

After visual inspection of the chip images to detect artefacts on the chip scan, the data were processed through the standard QC pipeline. Briefly, raw data files were read using the ReadAffy function of the affy package (10) from the Bioconductor suite of microarray analysis tools (11) available for the R statistical environment (12). The quality was controlled by generating and examining RNA degradation plots (AffyRNAdeg function of the affy package), NUSE, and RLE plots (affyPLM package (13)), and calculating the MA(RLE) values. Additionally, an eye-check of the pseudo-images was performed to ensure that no spatial effect was present.

Data processing and scoring methods were implemented in the R statistical environment (R version 3.0.2). Raw RNA expression data were analyzed using the affy and limma packages of the Bioconductor suite of microarray analysis tools available in the R statistical environment. Frozen Robust Multiarray Analysis background correction and quantile normalization were used to generate probe set expression values (14). For each data set, an overall linear model was fitted for all groups of replicates, and specific contrasts of interest (comparisons of "treated" and "control" conditions) were evaluated to generate raw *P*-values for each probe set on the expression array, which in turn were adjusted using the Benjamini–Hochberg (BH) procedure. Probesets in a specific comparison were considered to have a significant change in expression levels if they had an adjusted *P*-value <0.05.

4.5.2 Network-based analysis

Leveraging the "cause-and-effect" network models (15-20) together with Network Perturbation Amplitude algorithms (21,22), the gene expression fold-changes were translated into differential values for each node of the network (denoted by *f*). The node differential values were in turn summarized into a quantitative measure of network perturbation amplitude (NPA). The NPA was computed as a Sobolev norm on the signed directed graph underlying the network (*N*), which can be expressed as a quadratic form $1/\#\text{edges} \cdot f^T Q_N f$. Thus, to summarize, the NPA algorithm considered two main input components: the "cause-and-effect" network

model describing the mechanism and a gene expression data set from a well-designed experiment.

In addition to the confidence intervals of the NPA scores, which accounted for experimental error (e.g., biological variation between samples in an experimental group), companion statistics were derived to inform on the specificity of the NPA score with respect to the biology described in the network. Two permutation tests were implemented (21): the first assessed if the results were specific to the underlying evidence (i.e., gene fold-changes) in the model, leading to a permutation *P*-value (denoted by *O in the figures when <0.05). The second test assessed whether the "cause-and-effect" layer of the network significantly contributed to the amplitude of network perturbation (denoted by K* in the figures when <0.05). The network was considered to be specifically perturbed if both *P*-values were low (typically <0.05), and the perturbation was significant if the confidence interval was also above zero.

4.5.3 Biological impact factor (BIF) scoring

The network models represent functionally distinct biological processes characterizing the systems under consideration. To objectively evaluate the overall biological impact relative to a reference within the experiment (typically a standardized cigarette smoke exposure), the sum of the significant network perturbations for the contrast *i* was normalized with respect to the corresponding sum for the reference. Hence, the relative BIF (RBIF) for the contrast *i* is defined as follows:

$$RBIF(i) = \frac{\sum_{Net} w_i^{Net} f_i^{Net^T} Q_{Net} f_i^{Net}}{\sum_{Net} w_{REF}^{Net} f_{REF}^{Net^T} Q_{Net} f_{REF}^{Net}} = \frac{\sum_{Net} w_i^{Net} NPA_{Net}(i)}{\sum_{Net} w_{REF}^{Net} NPA_{Net}(REF)}$$

where the weights account for three statistics associated with the above outlined NPA algorithm.

The contribution of a given subset, S, of network models (e.g., cell stress sub-networks), for a contrast *i*, is given as follows:

$$Contrib_S(i) = \frac{\sum_S w_i^S NPA_S(i)}{\sum_{Net} w_i^{Net} NPA_{Net}(i)}$$

(because Nets is a disjointed union of subsets of networks, the contributions sum to one).

The relative BIF is therefore decomposed into network components (cell stress, pulmonary inflammation, cell proliferation, apoptosis, necroptosis, senescence, DNA damage, and autophagy) by considering the quantities $Contrib_S(i)\, RBIF(i)$, which are represented as starplots.

Finally, because RBIF is an aggregated quantity, two contrasts can have the same relative biological effect while arising from different network models. To identify those situations, a comparability coefficient was computed as follows:

$$\delta = \frac{\sum_{Net} w_i^{Net} w_{REF}^{Net} f_i^{Net^T} Q_{Net} f_{REF}^{Net}}{\sqrt{\sum_{Net} w_i^{Net} NPA_{Net}(i)} \sqrt{\sum_{Net} w_{REF}^{Net} NPA_{Net}(REF)}}.$$

This coefficient is essentially the cos angle between *i* and REF for the scalar product defined in the NPA algorithm.

4.6 Mass spectrometry-based analysis of proteome changes

4.6.1 Sample preparation

Tissue samples from the right lungs and livers of eight mice (eight biological replicates) were analyzed for each exposure condition and time point. Lung tissue from months 1, 2, 3, 6, and 8 for 3R4F and THS2.2 exposure and months 3, 6, and 8 for cessation and switching (to THS2.2) groups was available for quantitative proteomic analysis. All samples were processed in a random order. Frozen right lung tissue was homogenized using a bead-assisted procedure with Tissue Lyser II (Qiagen, Valencia, CA, USA) in tissue lysis buffer (BioRad, Hercules, CA, USA). Frozen liver samples were pulverized using a CryoPREP™ Impactor (Covaris, Woburn, MA, USA) and lysed with Tissue Lyser II as before in lysis buffer (BioRad). After lysis, the proteins were precipitated with acetone, and the precipitate was resuspended in 0.5 M triethylammonium bicarbonate, 1 M urea, and 0.1% SDS (all Sigma-Aldrich). A total of 50 µg was processed for the tandem mass tag (TMT) 6-plex labeling procedure according to the manufacturer's instructions (Thermo Scientific). Trypsin (Promega, Madison, WI, USA) was added to the samples in a 1:10 trypsin:protein ratio (w/w) followed by overnight digestion at 37°C. Digested samples were labeled with the reporter-ion tags for different treatment groups, and a common reference mix containing 50 µg of all protein extracts from each time point was prepared and labeled with a TMT reporter-ion tag. For each month of exposure, eight replicate TMT 6-plex sets of the five different exposure groups and the reference mix were defined. Within each month, the replicates were randomly assigned to TMT sets and sample-to-reporter mapping was randomized within each set. All labeled samples of one TMT set were pooled and dried in a SpeedVac. Samples were purified of SDS and remaining salts using 0.5 ml bed volume detergent removal columns (Pierce, Rockford, IL, USA) and 1 cc C18 reversed phase SepPak columns (Waters, Milford, MA, USA) according to the manufacturers' instructions. They were then dried in a SpeedVac and resuspended in nanoLC buffer A (5% acetonitrile, 0.2% formic acid (Sigma-Aldrich)).

4.6.2 Mass spectrometry for identification and quantification

Samples were analyzed in a random order using an Easy nanoLC 1000 instrument (Thermo Scientific) connected online to a Q-Exactive™ (Thermo Scientific) mass spectrometer. Peptides were fractionated 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 and a sensitive method as described by Kelstrup et al. (23) on the same column. Both mass spectrometry runs were searched together as merged mass-lists against the mouse reference proteome set (Uniprot, version July 2014, canonical isoforms only) using Proteome Discoverer version 1.4.0.288 software (Thermo Scientific). Mascot (v.2.4.1, Matrixscience, Boston, MA) and SequestHT (implemented in Proteome Discoverer) were used as search tools and resulting protein lists were merged and TMT-reporter ion intensities were determined from Proteome Discoverer software. The Percolator node of the Proteome Discoverer software was used to estimate peptide-level adjusted *P*-values (q-values) and the peptides were filtered for q-values <0.05 (i.e., the false discovery rate (FDR) was controlled at the 5% level). TMT peptide-level quantification data were exported and further processed in the R statistical environment (12). They were filtered for "unique" quantification results as defined by Proteome Discoverer software, e.g. removing redundant quantification results from multiple search engines. A global variance stabilizing normalization was performed with the respective Bioconductor package in R (24,25). Each TMT reporter ion set was normalized to its median, and protein expression values were calculated as the median of these normalized peptide-level quantification values (26). For the detection of differentially expressed proteins, a linear model was fitted for each exposure condition and its respective sham group and *P*-values from a moderated t-statistics were calculated using the empirical Bayes approach (9). The BH FDR method was then used to correct for multiple testing effects. Proteins with an adjusted *P*-value < 0.05 were considered differentially expressed.

4.6.3 Proteomic data analysis

4.6.3.1 Functional network analysis

Functional network analysis was used to identify functional protein clusters affected by the exposure conditions. This analysis was based on the functional association network of the STRING database (medium confidence or higher) (version 9) (27), which constitutes a comprehensive resource for functional protein interactions and integrates diverse sources of evidence including text-mining, co-expression, and pathway databases. The analysis involved four main steps: activated subnetwork identification, functional cluster identification, cluster annotation, and a comparison of exposure effects across all conditions. Depending on the complexity of the response, two different approaches for subnetwork identification were used: simple subgraph induction or a heuristic procedure for the identification of high scoring subnetworks. For subgraph induction, the entire functional interaction network was filtered for those interactions that involved two significantly regulated proteins in a given condition. Only the largest connected component of the induced subgraph was considered for further analysis. For the identification of high-scoring subnetworks, the previously published dnet approach was employed (28). Briefly, the functional network was filtered for proteins with an associated (non-adjusted) *P*-value (of differential expression), the *P*-values were transformed, and the highest scoring connected subnetwork was identified by minimum spanning tree finding and

postprocessing. For thresholding, the target node value was set to 100. Clusters were identified via a spin-glass model and simulated annealing (R igraph package, spins = 25). The Fruchterman–Reingold algorithm was used for the subnetwork layout. Visualization was supported by functions from the dnet and igraph packages. The enrichment of functional categories was evaluated for each cluster using the enrichPathway function from the ReactomePA R package (protein universe: all proteins evaluated for differential expression in the comparison). Based on these enrichment results, a concise functional annotation was associated with the clusters. Additionally, the cluster annotation was supported by the ToppGene Suite (29). Finally, to compare the impact of different exposure conditions on the identified proteins and clusters, the signed –log10 adjusted *P*-values or log2 fold-changes were used to color the protein nodes for each comparison; and the maxmean statistic of the average fold-changes of the cluster members was used to summarize the effects across all conditions (30).

4.6.4 Gene/protein set analysis

We used gene/protein set analysis to identify sets of proteins that demonstrated a significant average up- or down-regulation of its members. For this, we used the gene set analysis algorithm implemented in the piano package of R (31). We used absolute t-scores (versus the sham group) as the protein statistic (non-directional testing), the mean as the protein set statistic, and protein permutation with multiple hypothesis testing correction (BH method) to estimate significance.

5 Results and Discussion

5.1 Summary omics endpoints

A schematic of all omics endpoints analyzed is shown in Figure 3.

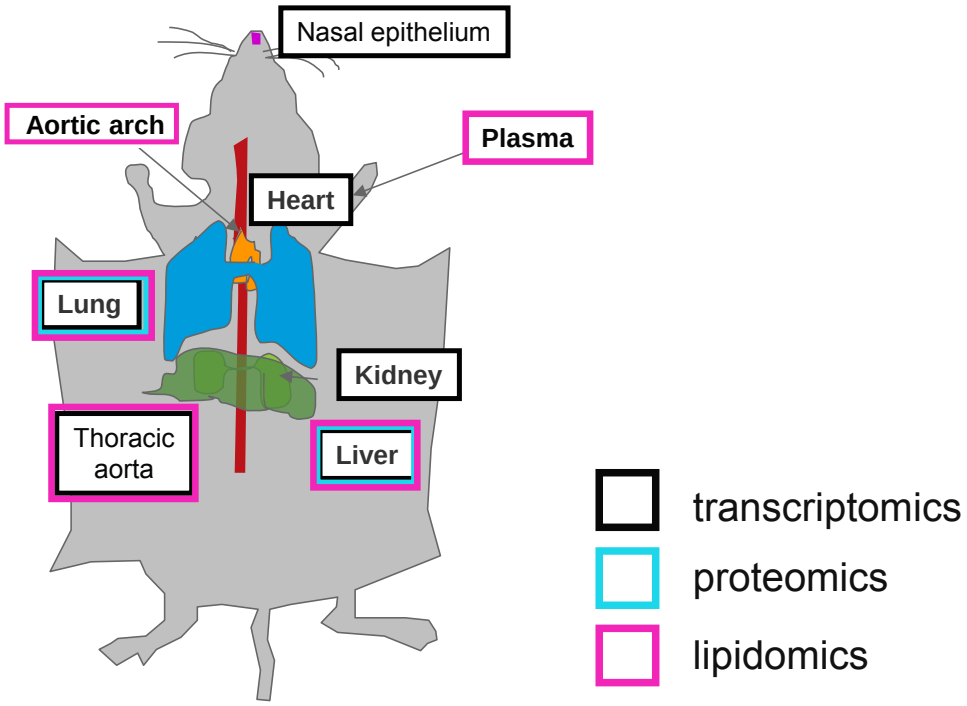


Figure 3: Systems biology endpoints

Results are shown in this report, but a deeper biological interpretation of the significance of the responses will be performed in the future.

5.2 Nasal epithelium transcriptomics

5.2.1 QC and number of replicates per group

RNA from a number of nasal epithelium samples was degraded and could not be hybridized onto microarrays and profiled. The number of replicates available for each group is summarized in Table 2.

Table 2: Number of valid replicates of nasal epithelium transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	8	8	8	7
3R4F	8	8	7	8	7
THS2.2	6	7	6	7	8
Cessation	-	-	6	7	8
Switch	-	-	6	7	8

5.2.2 Differentially expressed genes in the RNE

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value or FDR) of gene expression changes in the RNE for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 4. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 5.

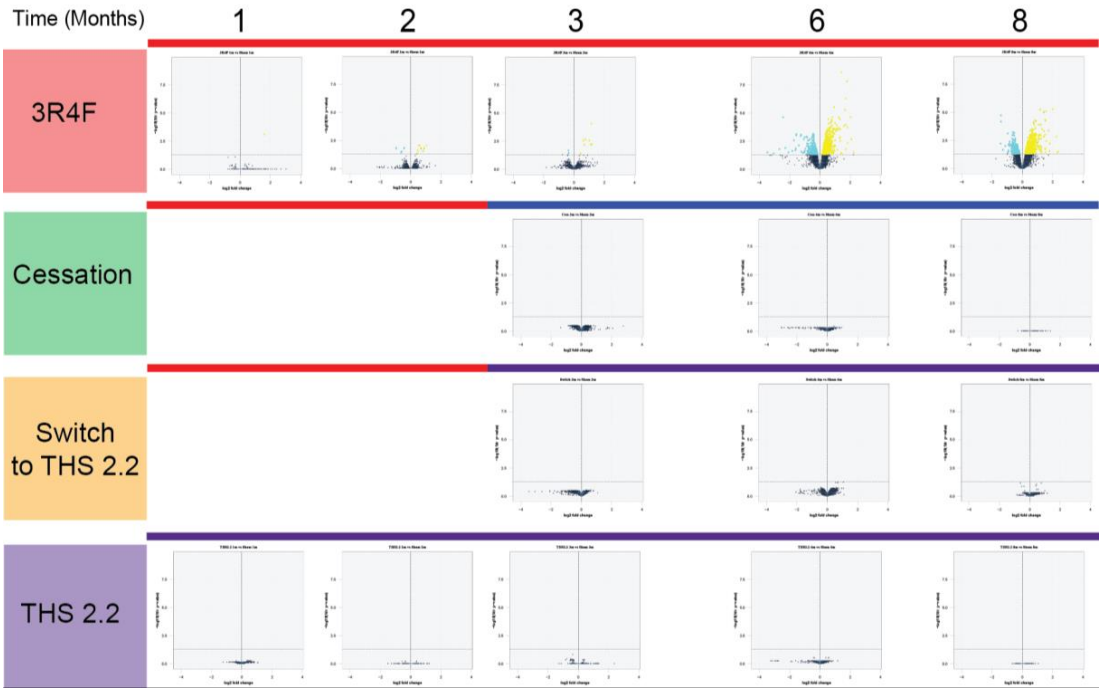


Figure 4: Volcano plots representing the systems response profiles in the RNE: FDR

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that are significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

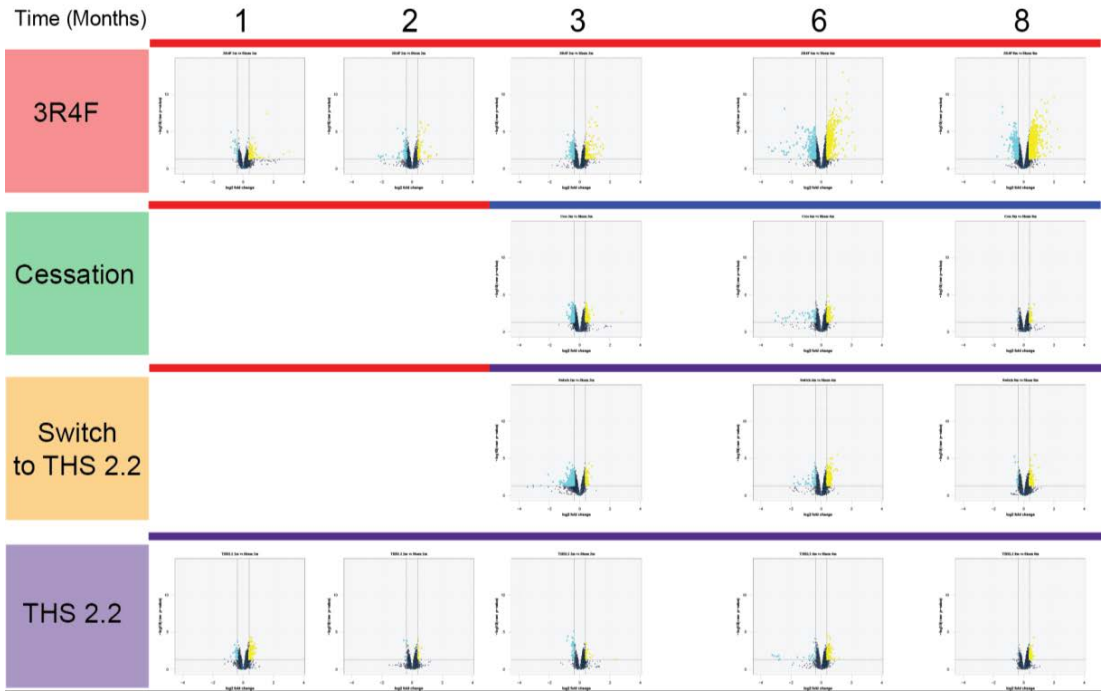


Figure 5: Volcano plots representing the systems response profiles in the RNE: raw *P*-values
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold-change >1.2.

The number of differentially expressed genes compared with sham for each exposure group (considering FDR<0.05) is given in Table 3. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, a pairwise comparison of overlap is shown in Figure 6.

Table 3: Number of differentially expressed genes in the nasal epithelium (up-/down-regulated compared with sham at each respective time point)
REF: reference

Exposure group/time point (months)	1	2	3	6	8
Sham	REF	REF	REF	REF	REF
3R4F	1/0	8/4	8/2	964/366	968/390
THS2.2	0	0	0	0	0
Cessation	-	-	0	0	0
Switch	-	-	0	0	0

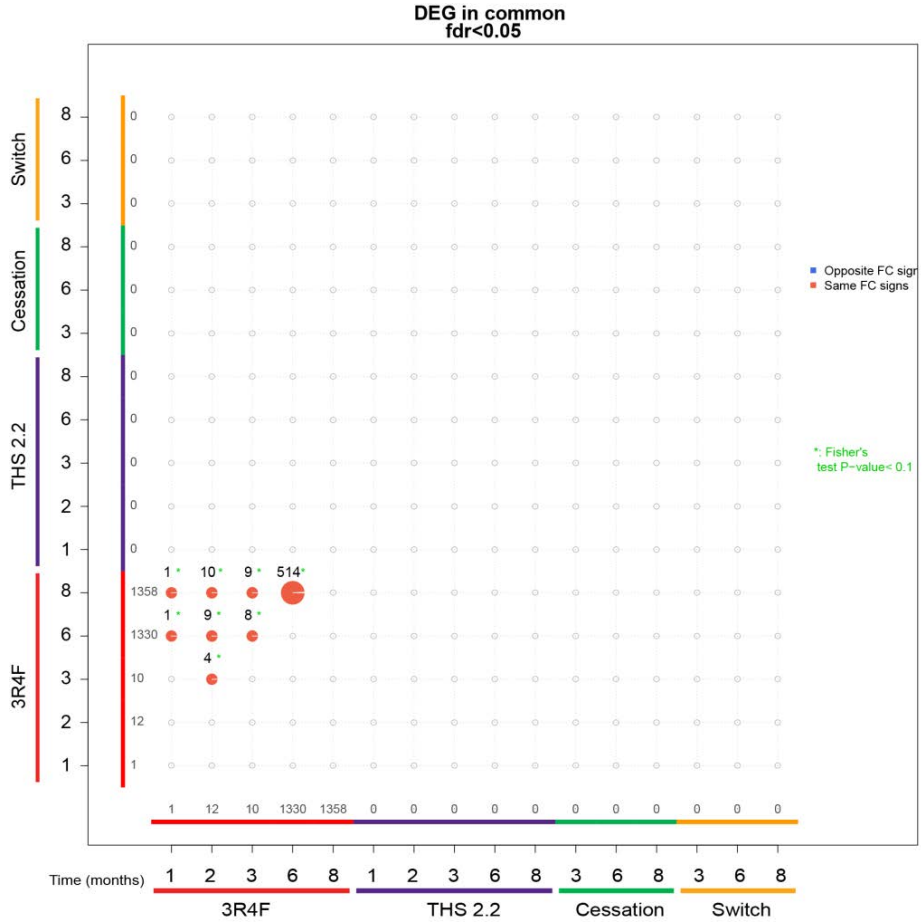


Figure 6: Overlap between sets of differentially expressed genes in the RNE

For each pair of comparisons, the number of shared up- and down-regulated genes is visualized. The overlap of significantly regulated genes that are regulated in the same direction (either both up-regulated or both down-regulated) is indicated in red, while opposite direction changes are shown in blue. A green star signifies that the overlap is significant (Fisher's test $P<0.1$). The number of significantly regulated genes (FDR<0.05) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.2.3 BIF analysis in nasal epithelium tissue

The overall impact of the different exposures was estimated from a BIF (32), expressed as an RBIF (Figure 7).

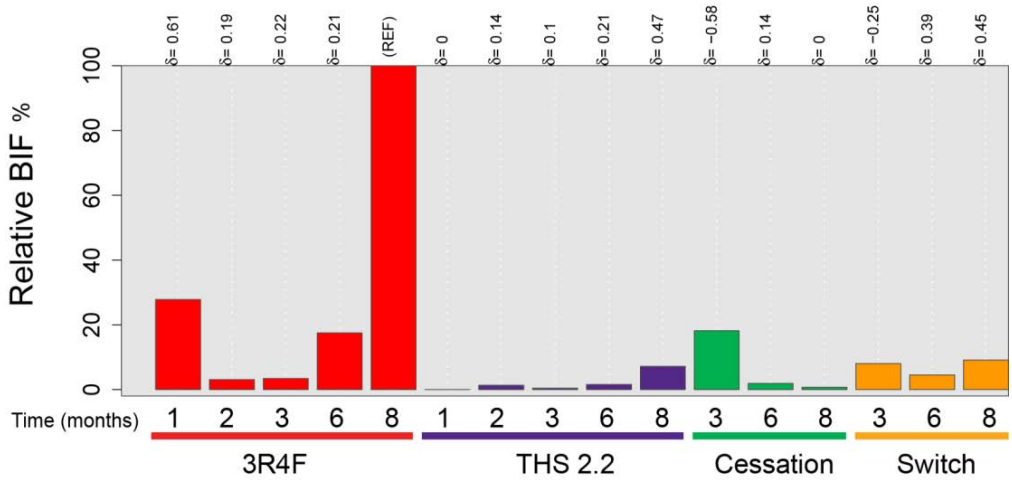


Figure 7: BIF analysis in the RNE

The percentages show the relative biological impact, derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 8 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 8 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.2.4 Biological network perturbation amplitude analysis in nasal epithelium tissue

The star plots in Figure 8 show the biological processes that most contributed to the overall BIF phenotype at each time point.

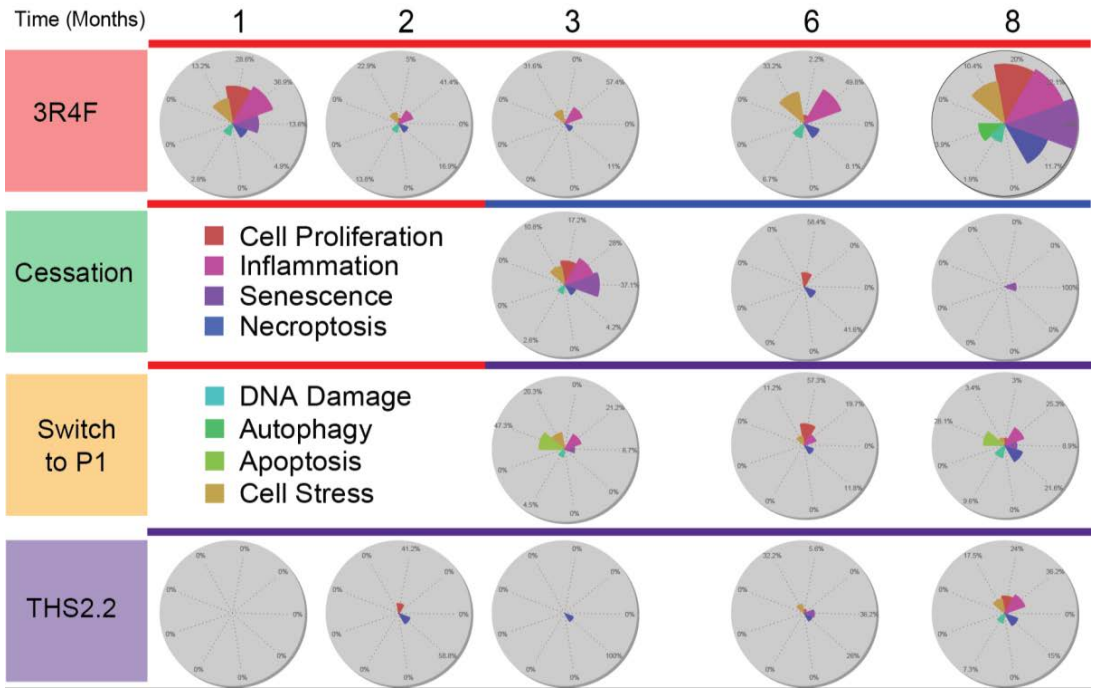


Figure 8: Star plots showing BIF calculated from RNE tissue

The decomposition of the overall RBIF (Figure 7) is shown in its eight mechanistic components (from cell proliferation to inflammation, indicated by segment colors) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

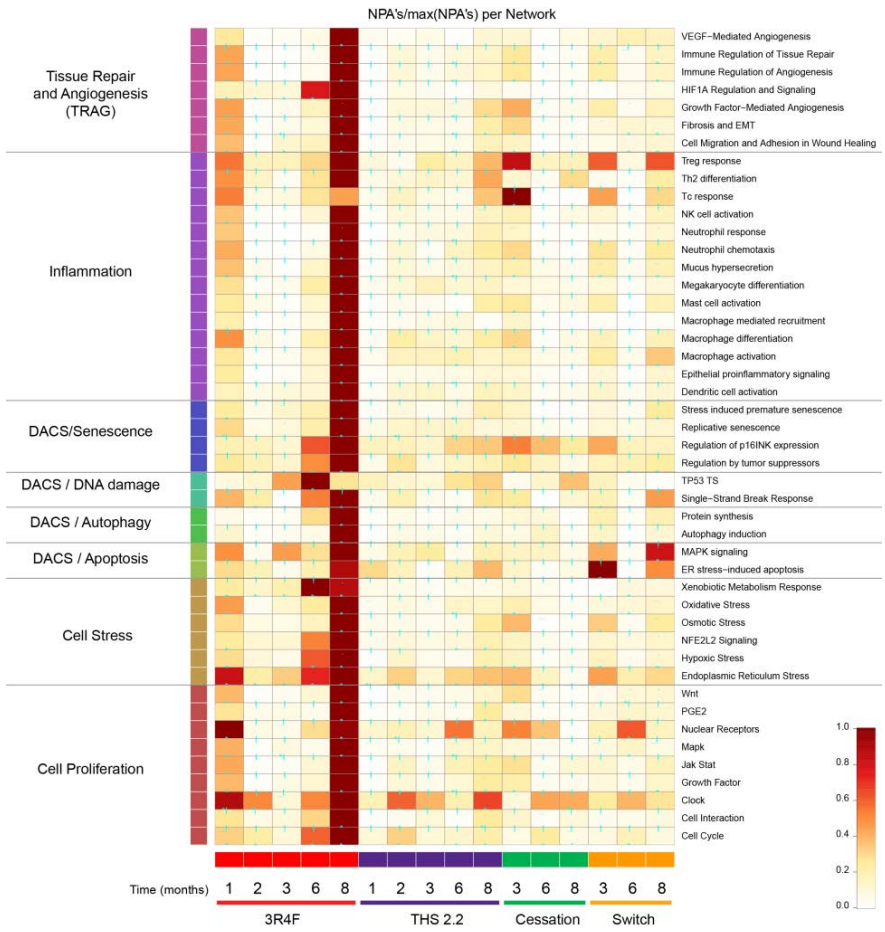


Figure 9: Heatmap of network perturbation amplitude scores in the RNE

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network is considered perturbed if, in addition to the significance of the NPA score with respect to experimental variation, the two companion statistics (O and K) derived to inform on NPA score specificity with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values <0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to the experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrast comparisons, in a similar way to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 10).

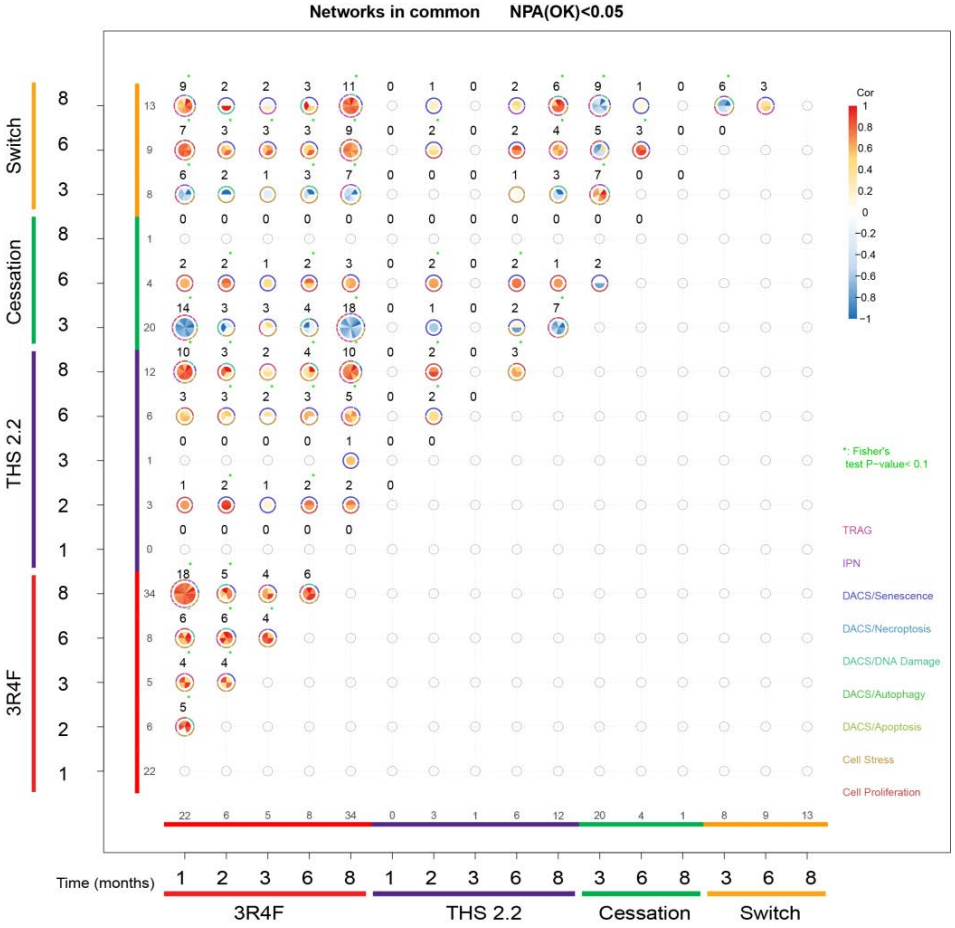


Figure 10: Overlap between sets of significantly perturbed networks in the RNE

For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates a positive correlation, while blue indicates a negative correlation. A green star signifies that the overlap is significant (Fisher's test P<0.1). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point).

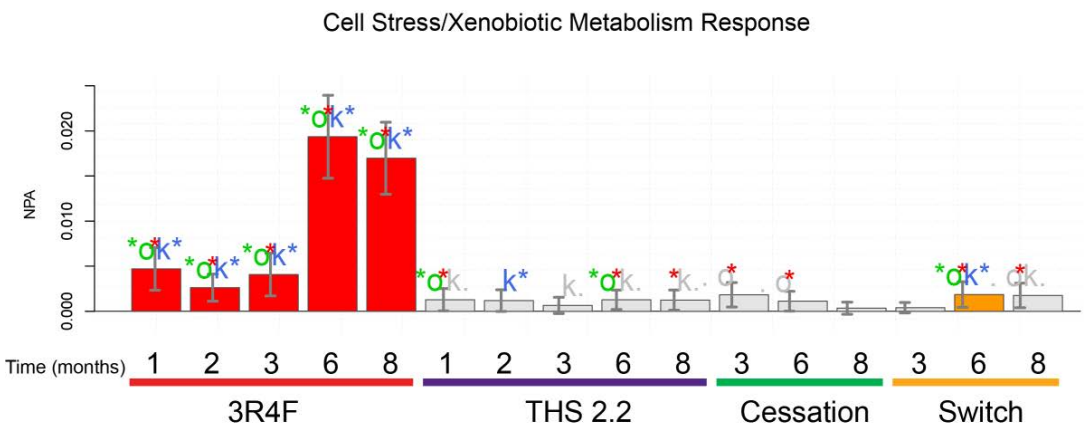


Figure 11: Network perturbation amplitude scores for the xenobiotic metabolism response network

NPA scores for the network model are shown with their confidence interval accounting for experimental variation. Companion statistics derived to inform on the specificity of the NPA score with respect to the biology described in the respective network are shown as *O and K* if their *P*-values were below the significance level.

5.2.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 12).

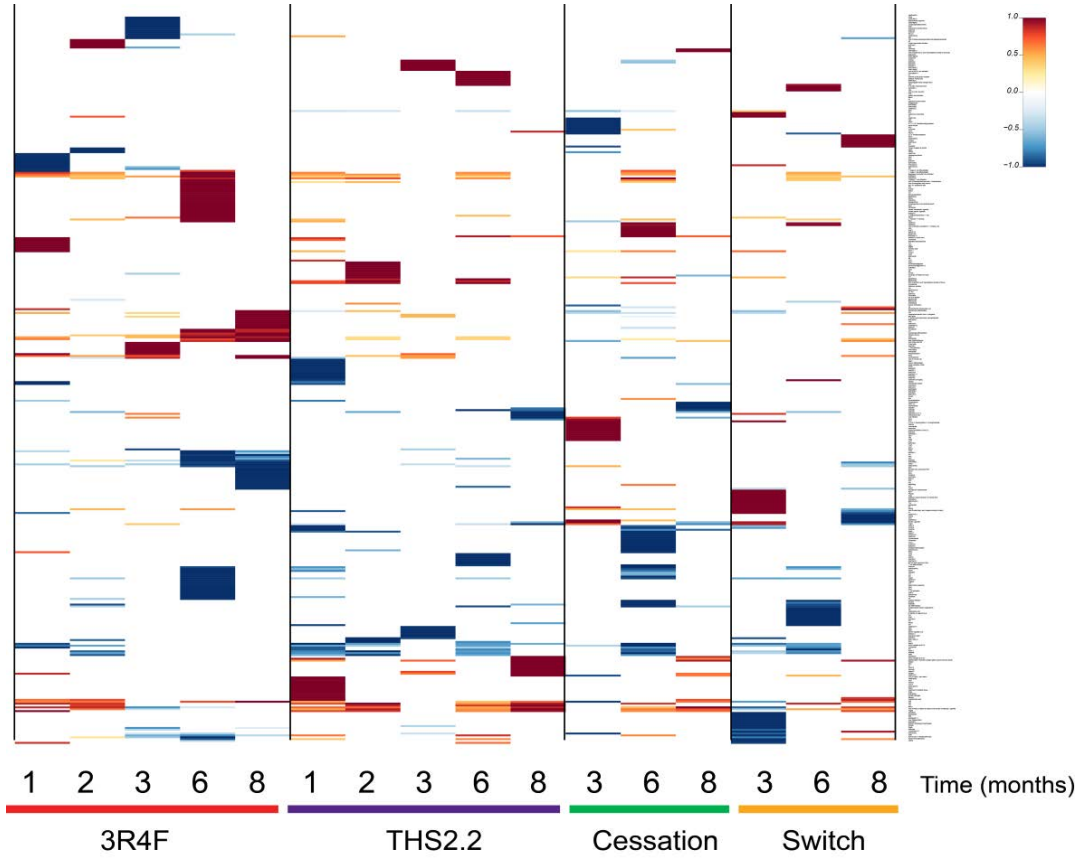


Figure 12: Heatmap of NPAI score for all hyps significant in at least one contrast in the RNE

5.2.6 NextBio analysis

5.2.6.1 3R4F exposure effect

To study the effects of 3R4F on RNE, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed nasal epithelia at five time points.

The top biogroups (Broad MSigDB canonical pathways) were found to be mostly linked to xenobiotic metabolism and lipid transport (Figure 13).

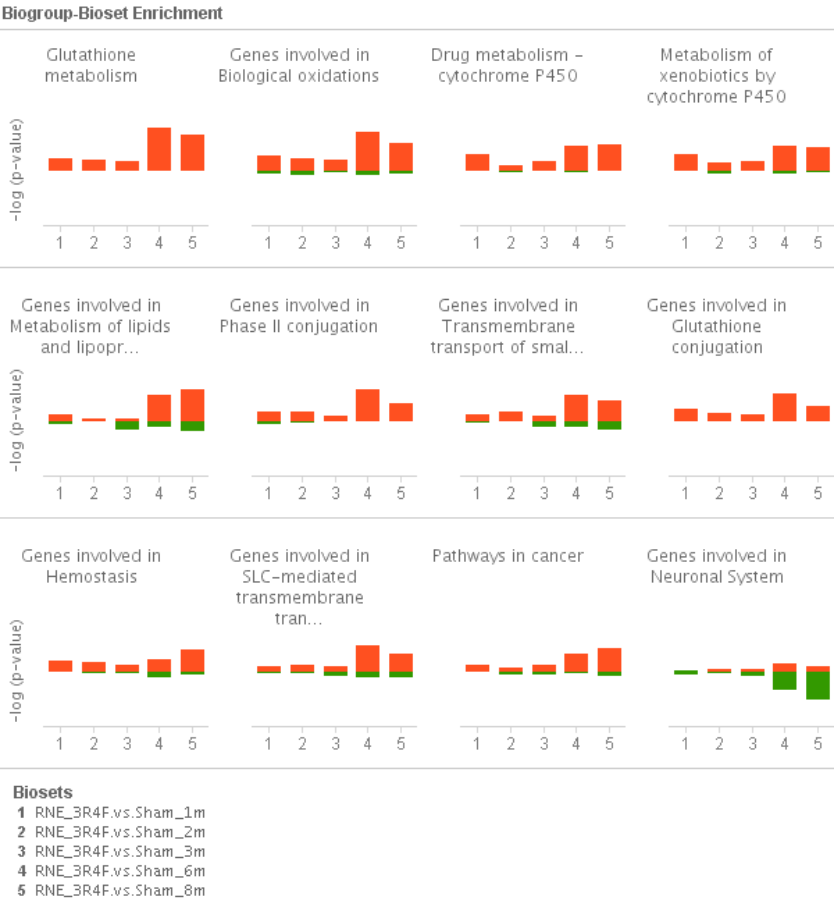


Figure 13: NextBio meta-analysis of 3R4F effects on the RNE

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to 3R4F (compared with sham-exposed control).

5.2.6.2 THS2.2 exposure effects

To study the effects of THS2.2 on RNE, a similar meta-analysis was performed comparing THS2.2-exposed and sham-exposed nasal epithelia at five time points.

The results were less consistent across time points (Figure 14) than those for 3R4F, reflecting the fact that the signal is much lower than in 3R4F-exposed tissues.

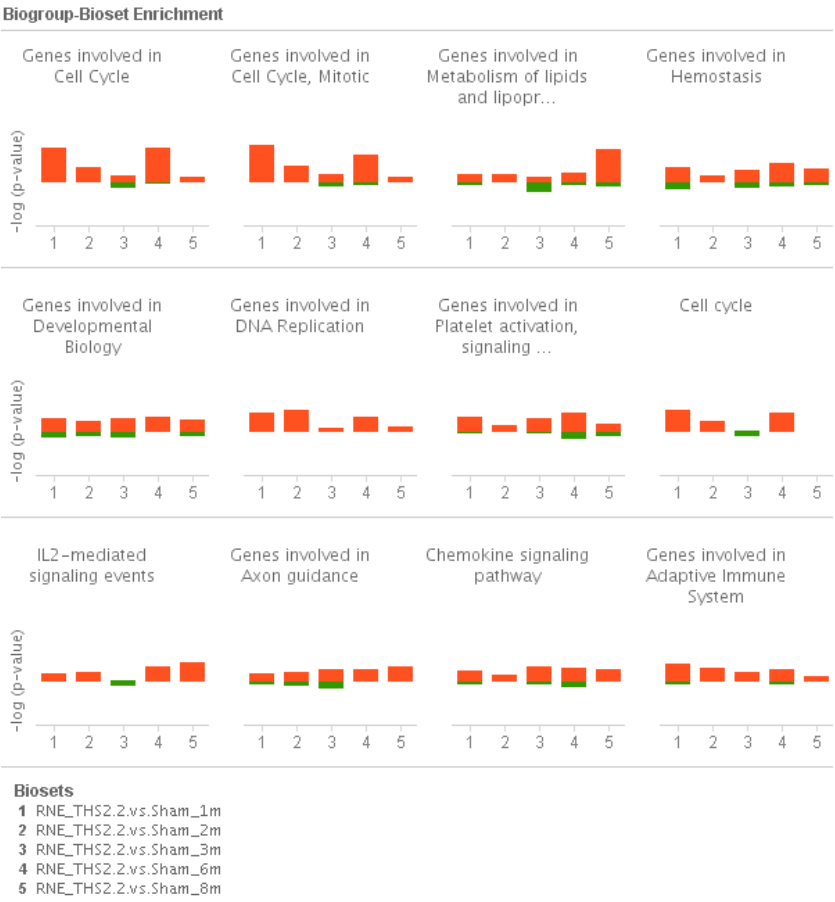


Figure 14: NextBio meta-analysis of THS2.2 effects on the RNE

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to THS2.2 (compared with sham-exposed controls).

5.2.7 Gene set enrichment analysis

5.2.7.1 3R4F effects

To further study trends within each treatment group over time, gene set enrichment analysis was carried out.

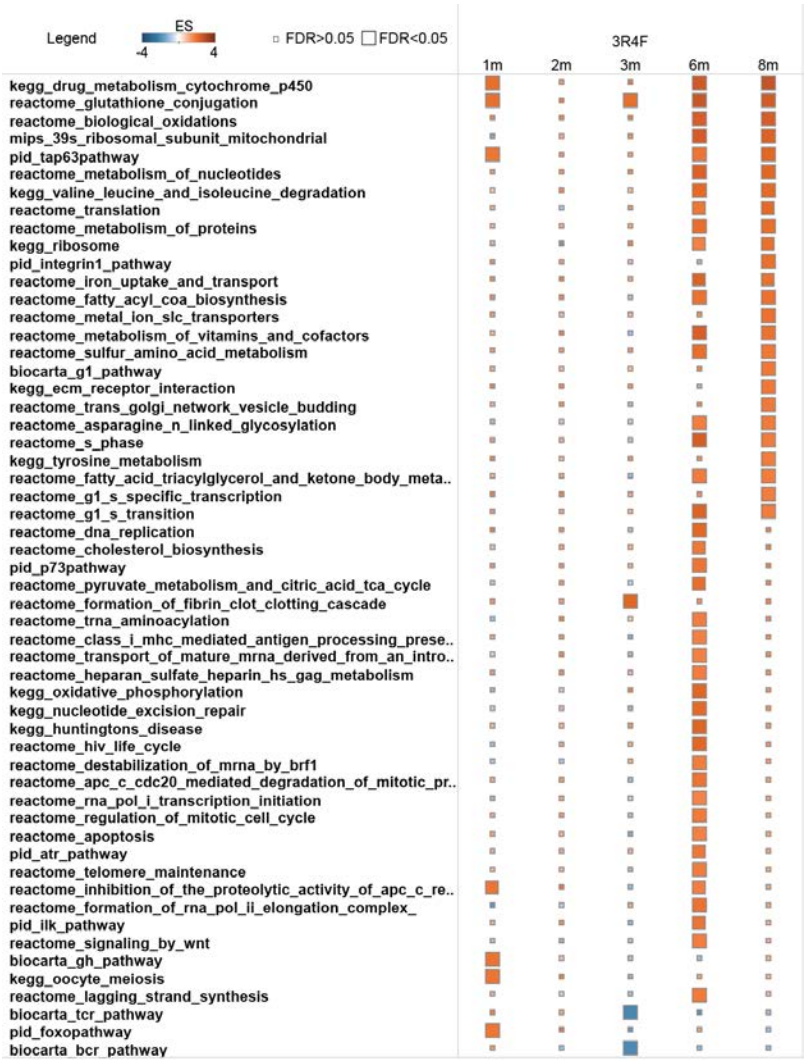


Figure 15: Gene set enrichment analysis in the 3R4F-exposed RNE

5.2.7.2 THS2.2 effects



Figure 16: Gene set enrichment analysis in the THS2.2-exposed RNE

5.2.7.3 Cessation vs. switch to THS2.2 effects

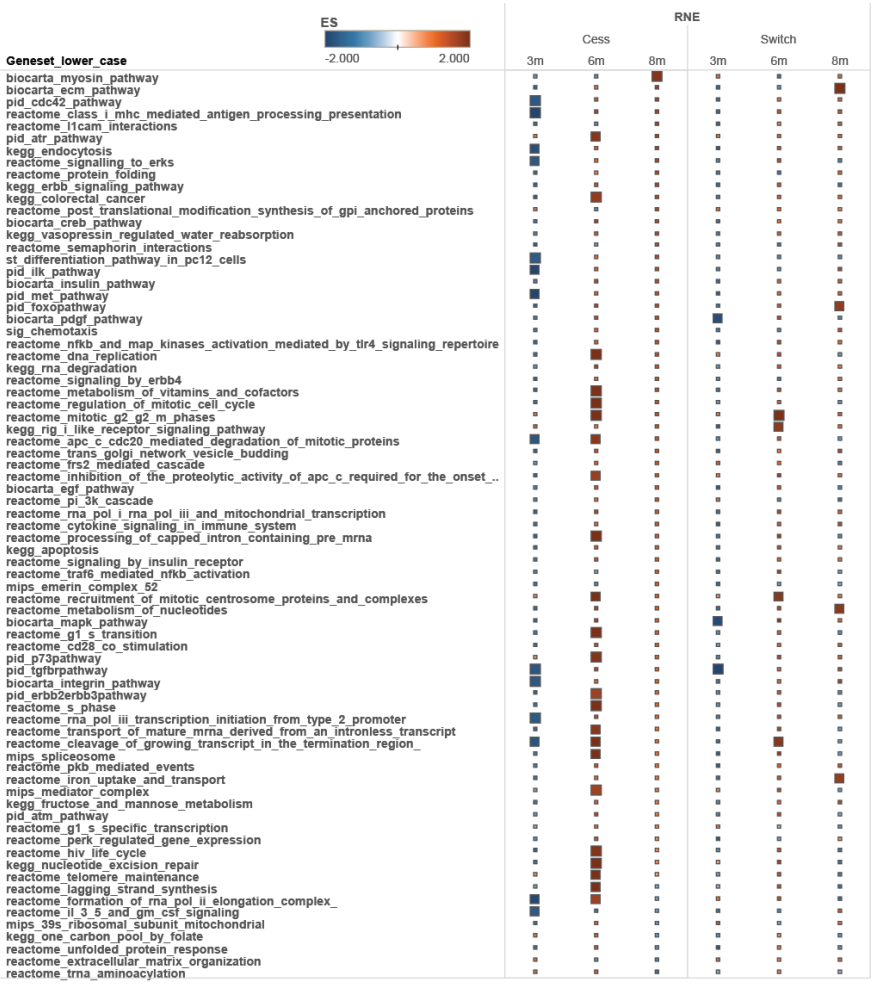


Figure 17: Gene set enrichment analysis in the RNE of animals from the cessation and switch to THS2.2 groups

5.2.8 Gene expression of cytochrome P450 genes

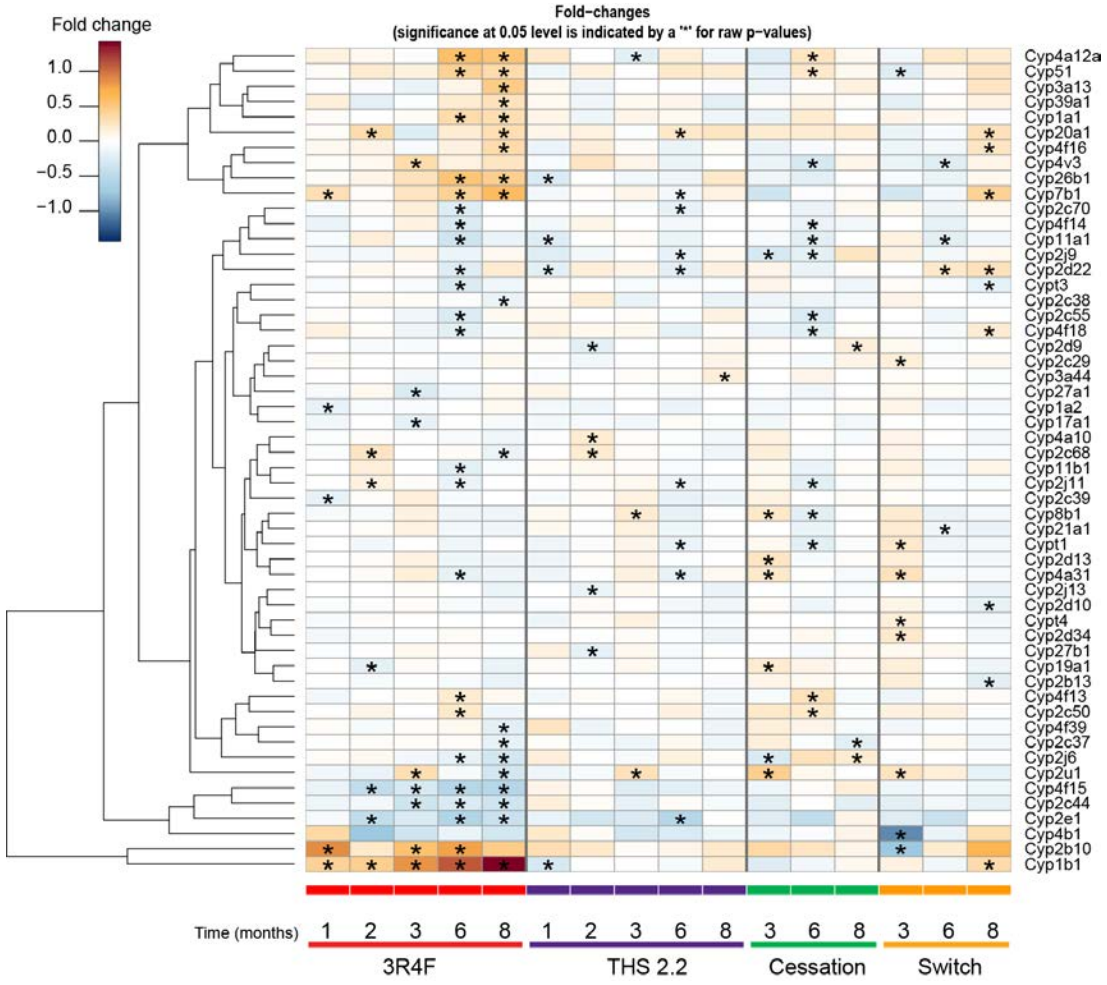


Figure 18: Fold-change of cytochrome P450 gene expression levels in the RNE compared with sham at each time point

5.2.9 Comparison with previous Apoe$^{-/-}$ SMAR findings

Results obtained from a previous Apoe$^{-/-}$ SMAR study have been published in a series of articles (not including SMAR or switch data). RNE transcriptional profiles were previously compared with lung gene expression profiles, and the xenobiotic metabolism pathway was found to be strongly regulated in the RNE (4). The same genes reported in this earlier article were also shown to be up-regulated by cigarette smoke exposure in the present study (Figure 19, Figure 20)

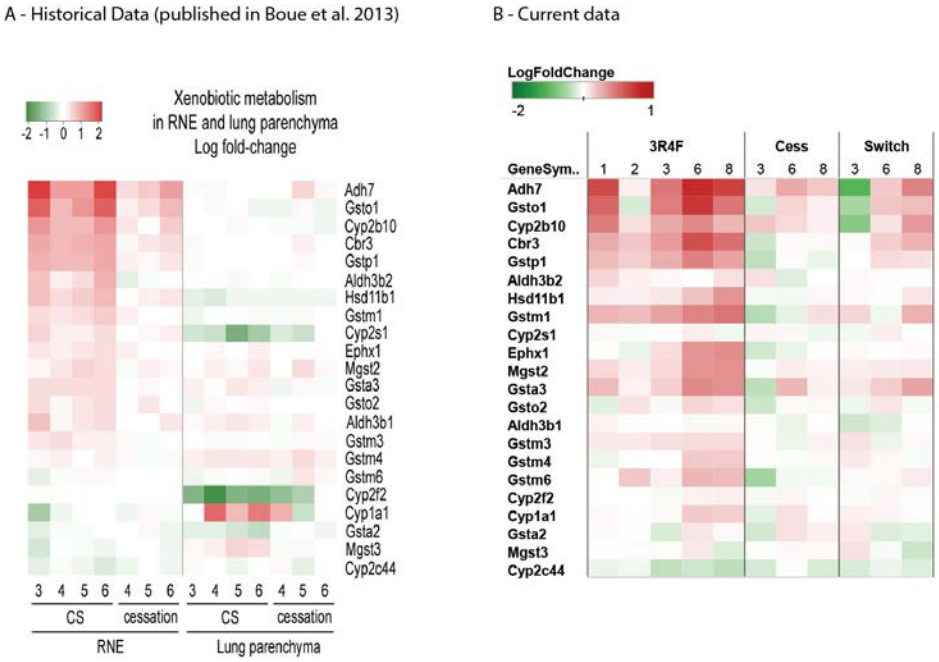


Figure 19: Xenobiotic metabolism pathway: comparison with a historical study (2)

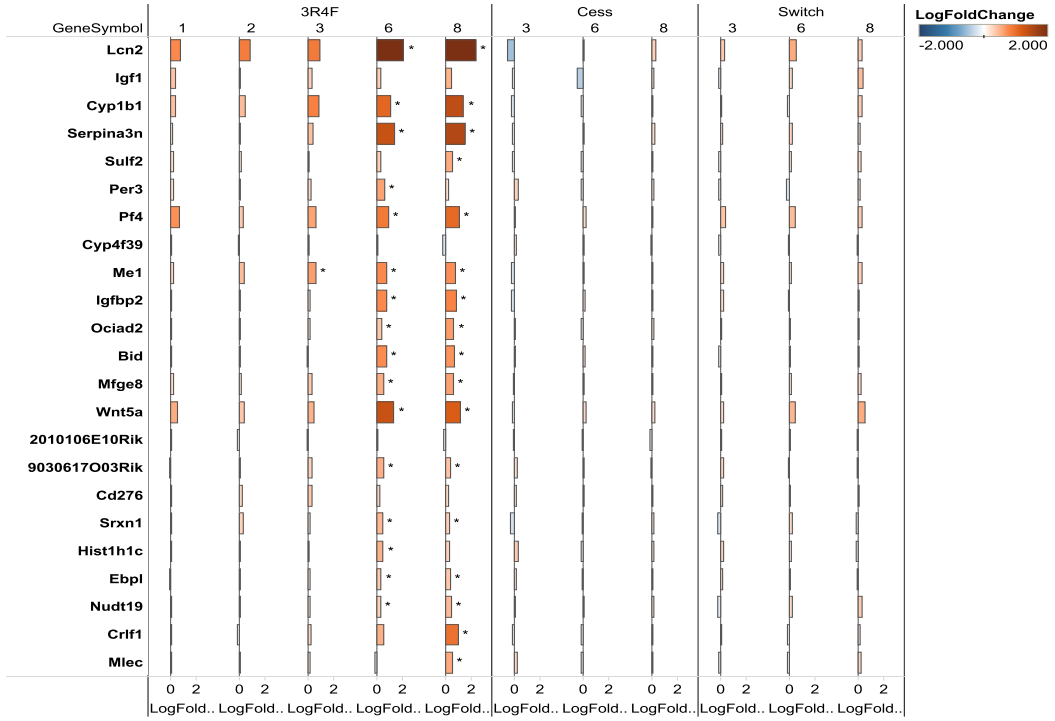


Figure 20: Genes found to be similarly up-regulated in the RNE and lung as in a previous study (Fig. 4B from (2))

5.3 Lung transcriptomics

5.3.1 QC and number of replicates per group

The number of replicates effectively available for each group is summarized in Table 4. Because of possible issues with quality, all tissues from time point 3 were rehybridized.

Table 4: Number of valid replicates of lung transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	7	8	8	7	6
3R4F	8	5	6	7	8
THS2.2	8	7	7	7	8
Cessation	-	-	7	5	8
Switch	-	-	7	7	7

5.3.2 Differentially expressed genes in the lung

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of gene expression changes in the lung for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 21. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 22.

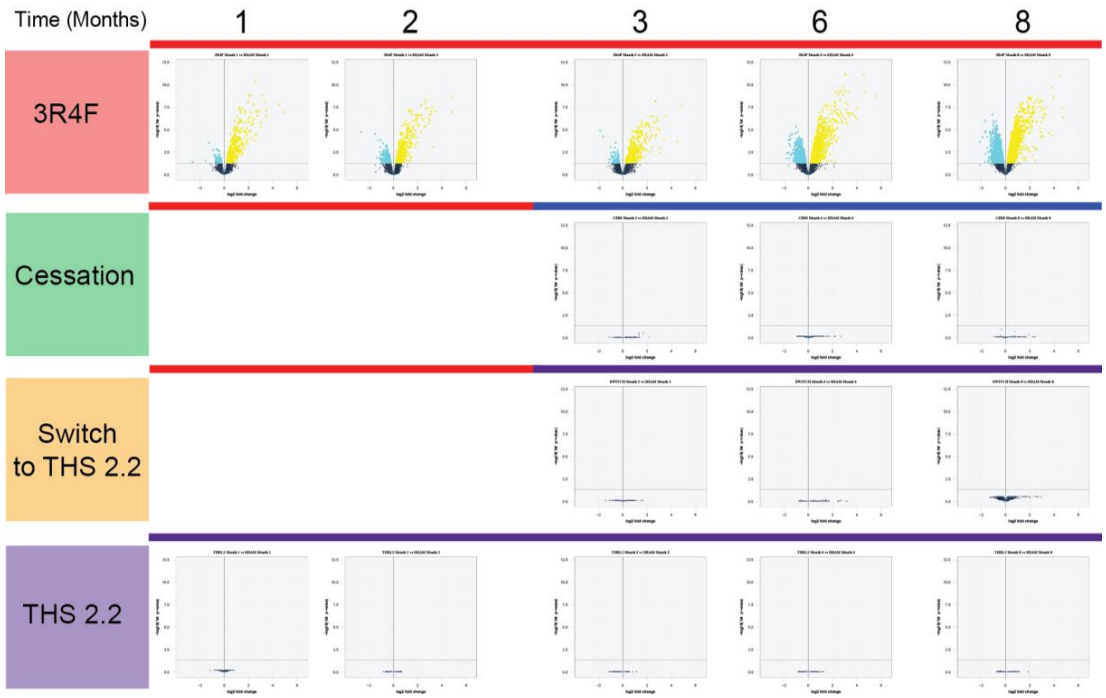


Figure 21: Volcano plots representing the systems response profiles in the lung: FDR

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that are significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

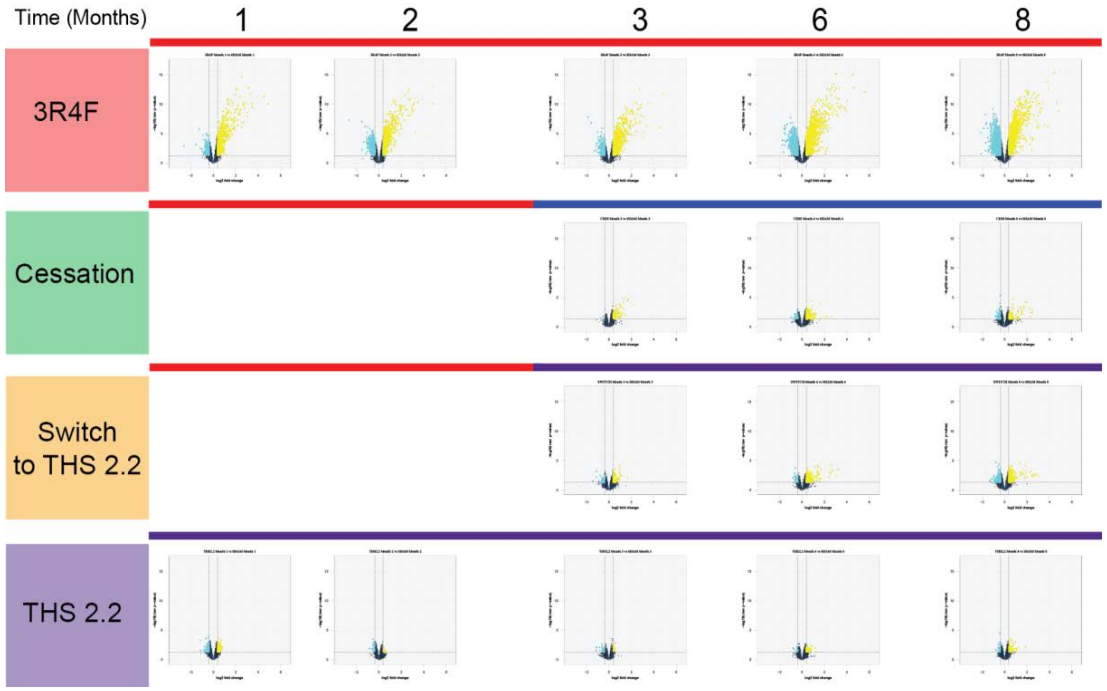


Figure 22: Volcano plots representing the systems response profiles in the lung: raw *P*-value
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold change >1.2.

The number of differentially expressed genes compared with sham for each exposure group is shown in Table 5. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, the overlap between each pair of contrasts is shown in (Figure 23).

Table 5: Number of differentially expressed genes in the lung (up-/down-regulated compared with sham at each respective time point)

Exposure group/time point (months)	1	2	3	6	8
Sham	REF	REF	REF	REF	REF
3R4F	720/123	551/353	503/106	1405/809	2888/2212
THS2.2	0	0	0	0	0
Cessation	-	-	0	0	0
Switch	-	-	0	0	0

REF: reference

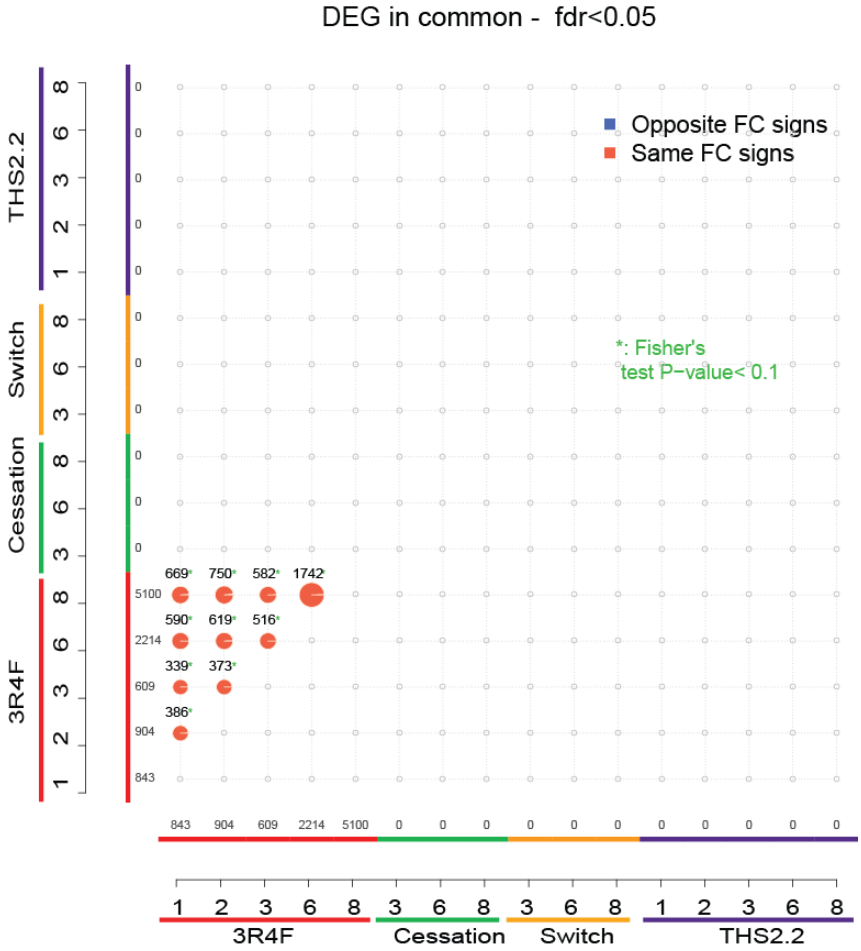


Figure 23: Comparison of differentially expressed genes between contrasts

The pie chart at each intersection represents the overlap between the respective contrasts of differentially expressed genes (FDR<0.05). The red portion of the pie indicates differentially expressed genes regulated in the same direction (either both up- or both down-regulated), while the blue indicates regulation in opposite directions. The significance of the overlap (P<0.1) tested by Fisher's test is indicated by the green star.

5.3.3 BIF analysis in the lung

The overall impact of the different exposures was estimated from a BIF (32), expressed as an RBIF (Figure 24).

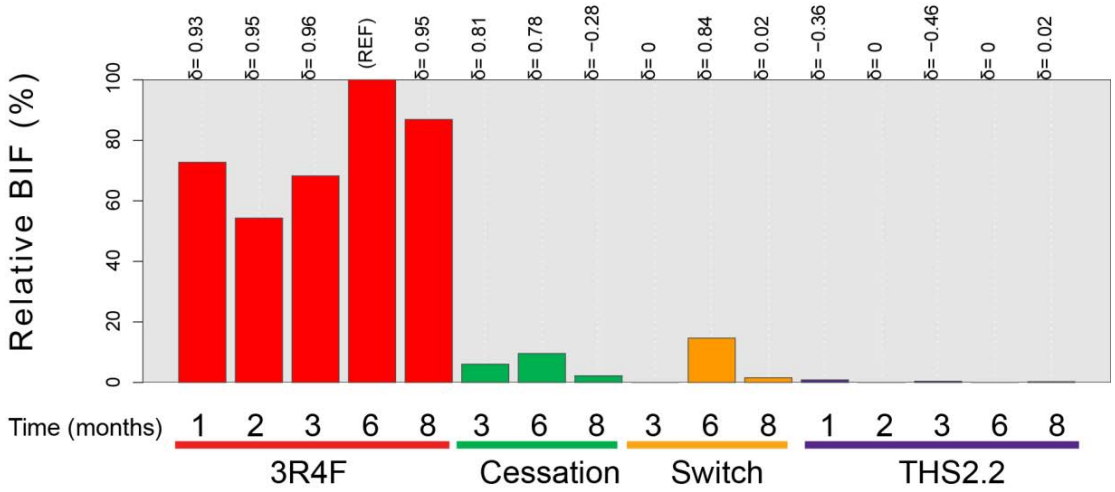


Figure 24: BIF analysis in the lung

The percentages show the relative biological impact, derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 6 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 6 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.3.4 Biological network perturbation amplitude analysis in the lung

The star plots in Figure 25 show the biological processes that most contributed to the overall BIF phenotype at each time point.

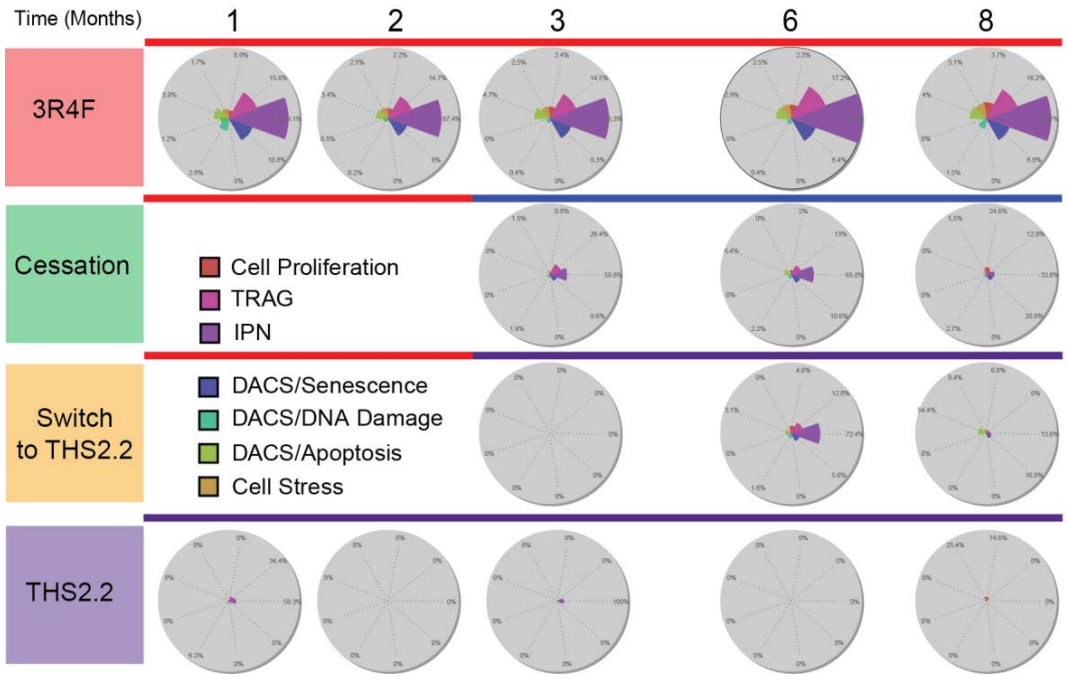


Figure 25: Star plots showing BIF calculated from the lungs

Star plots illustrating the decomposition of the overall RBIF (Figure 24) into its eight mechanistic components (from cell proliferation to inflammation, indicated by segment colors) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

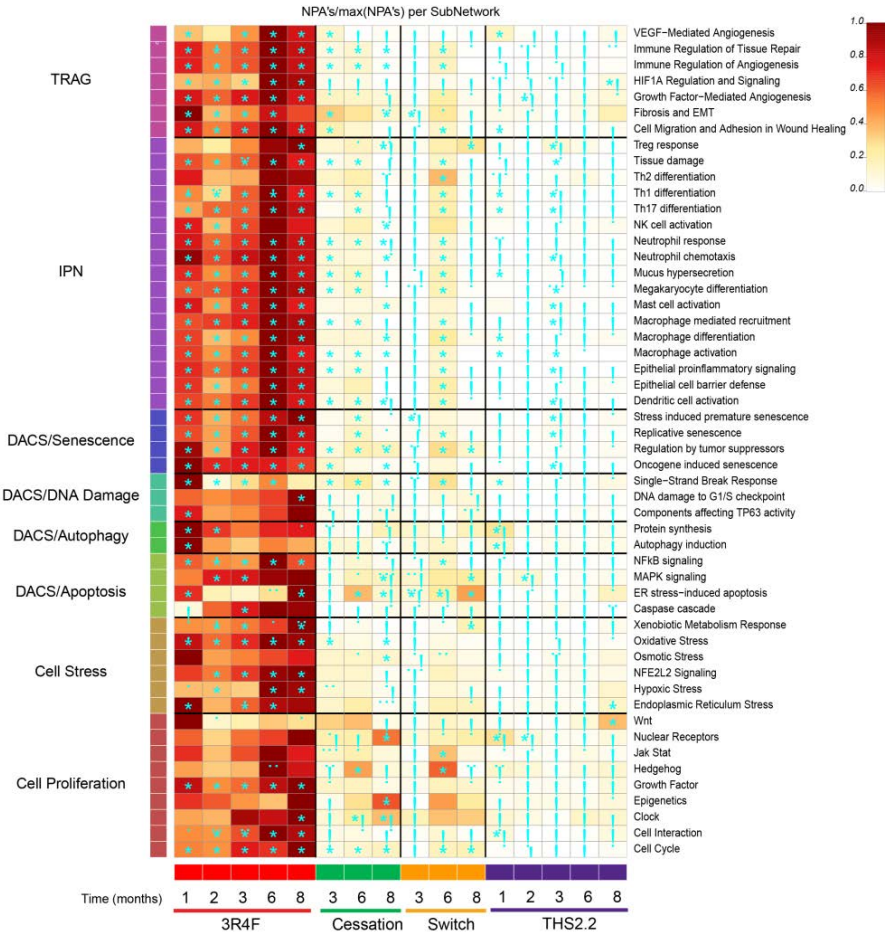


Figure 26: Heatmap of network perturbation amplitude scores in the lung

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network is considered perturbed if, in addition to the significance of the NPA score with respect to experimental variation, the two companion statistics (O and K) derived to inform on specificity of the NPA score with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values <0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrasts comparisons, similar to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 27).

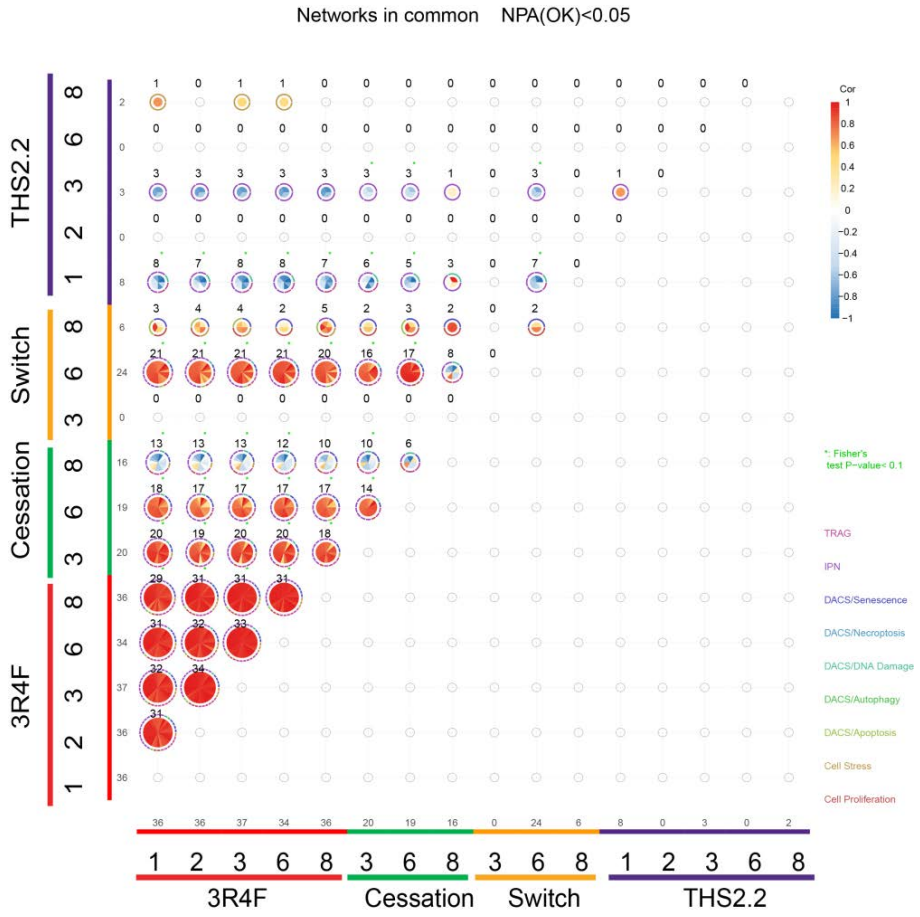


Figure 27: Overlap between sets of significantly perturbed networks in the lung

For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates positive correlation, while blue indicates negative correlation. A green star signifies that the overlap is significant (Fisher's test $P<0.1$). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point).

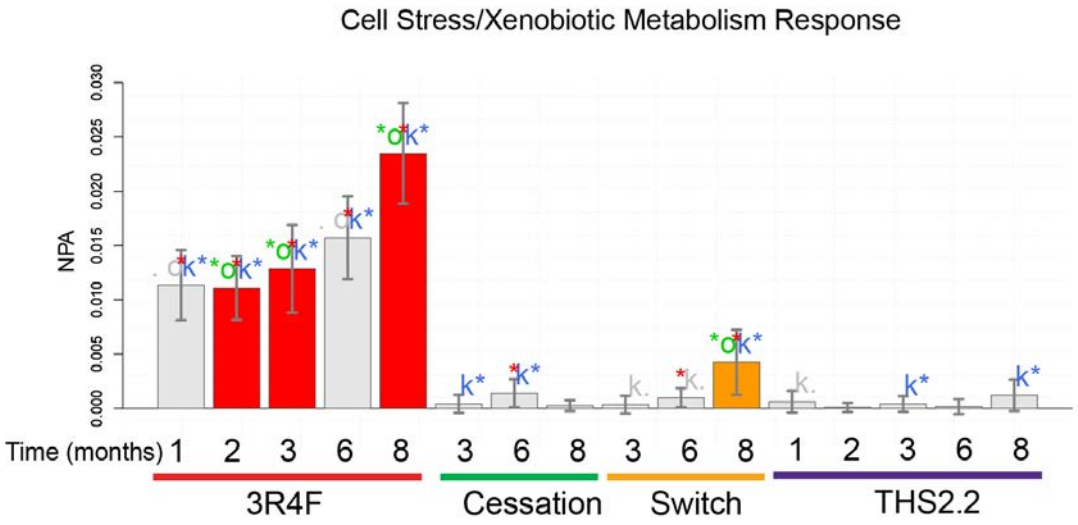


Figure 28: Network perturbation amplitude scores for the xenobiotic metabolism response network

NPA scores for the network model are shown with their confidence interval accounting for experimental variation. Companion statistics derived to describe specificity of the NPA score with respect to the biology described in the respective network are shown as *O and K* if their *P*-values were below the significance level.

5.3.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by the different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 29).

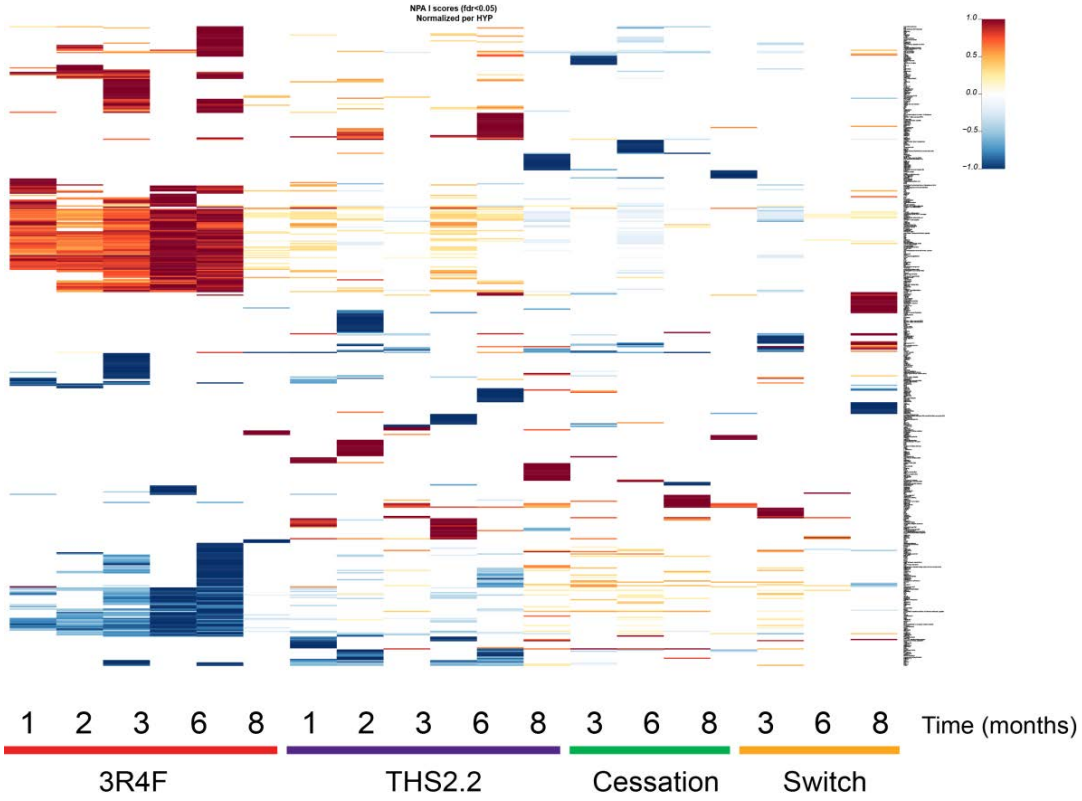


Figure 29: Heatmap of NPAI score for all hyps significant in at least one contrast in the lung

5.3.6 NextBio analysis

5.3.6.1 3R4F exposure effect

To study the effects of 3R4F on the lung, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed lung at five different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly linked to xenobiotic metabolism and lipid transport (Figure 30).

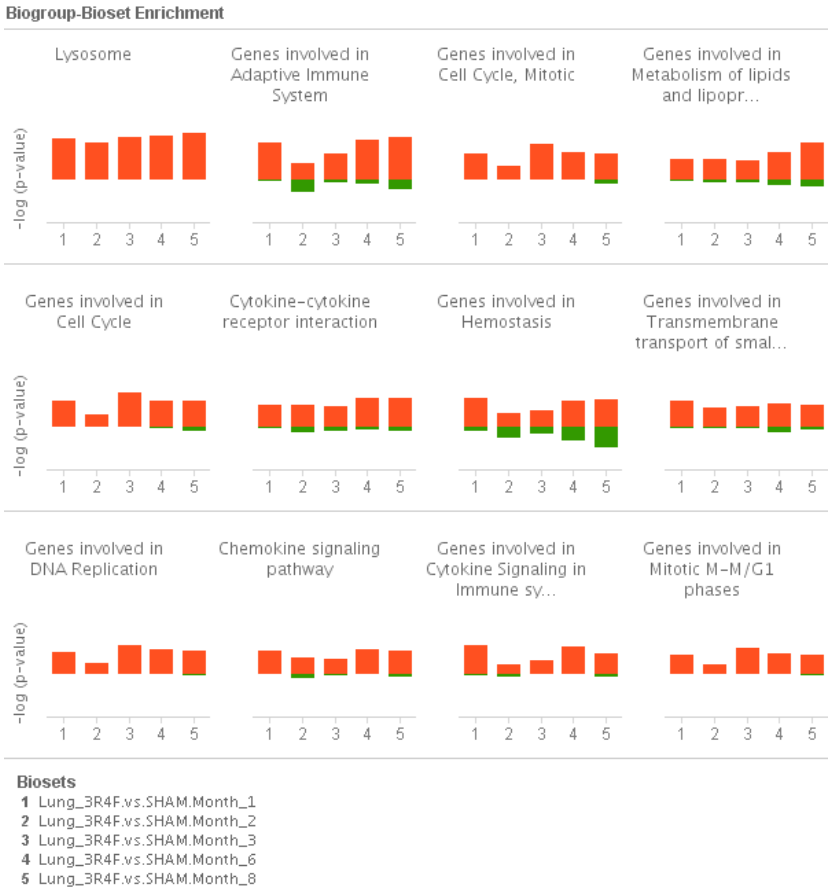


Figure 30: NextBio meta-analysis of 3R4F effects in the lung

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to 3R4F (compared with sham-exposed control).

5.3.6.2 THS2.2 exposure effects

To study the effects of THS2.2 on the lung, a meta-analysis was performed comparing THS2.2-exposed and sham-exposed lungs at five different time points.

Results were less consistent across time points than for 3R4F in the lung (Figure 31), indicating that the signal is much lower than in 3R4F-exposed tissues.

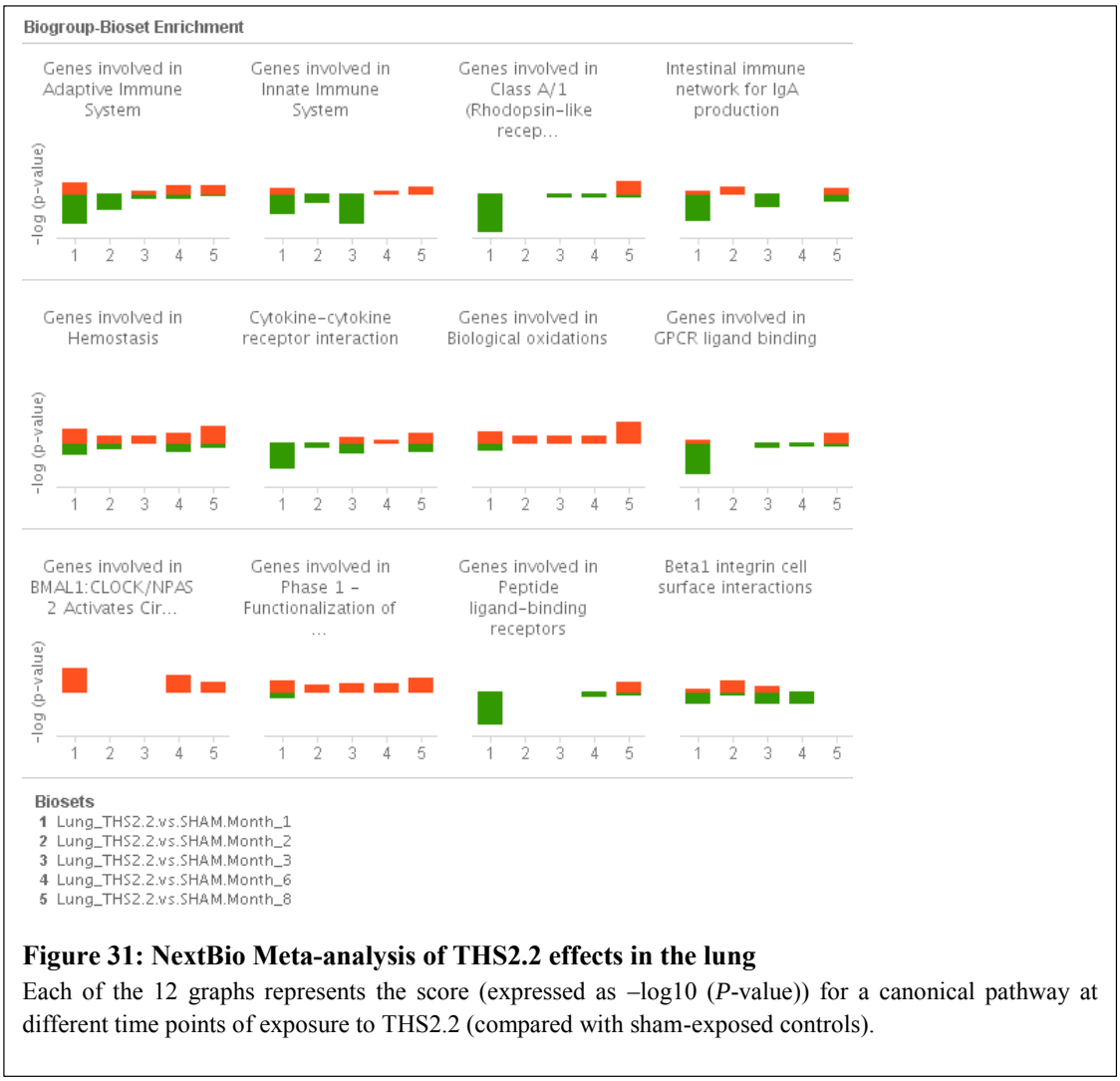


Figure 31: NextBio Meta-analysis of THS2.2 effects in the lung

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to THS2.2 (compared with sham-exposed controls).

5.3.7 Gene set enrichment analysis

5.3.7.1 3R4F effects

To further study the trends within each treatment group over time, gene set enrichment analysis was carried out.



Figure 32: Gene set enrichment analysis in the 3R4F-exposed lung

5.3.7.2 THS2.2 effects

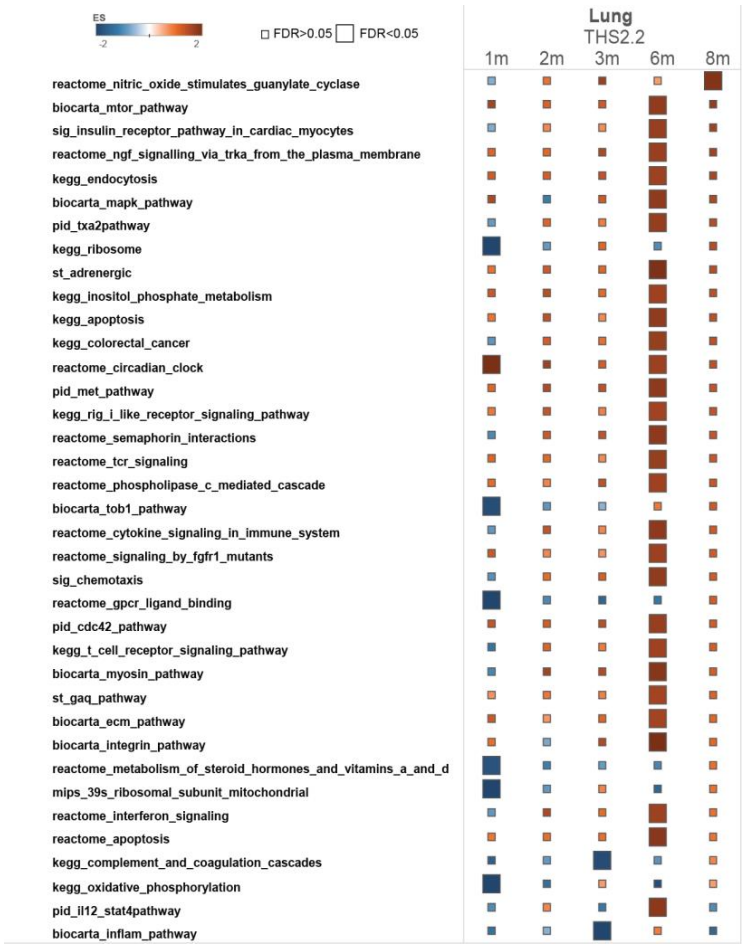


Figure 33: Gene set enrichment analysis in the THS2.2-exposed lung

5.3.7.3 Cessation vs. switch to THS2.2 effects

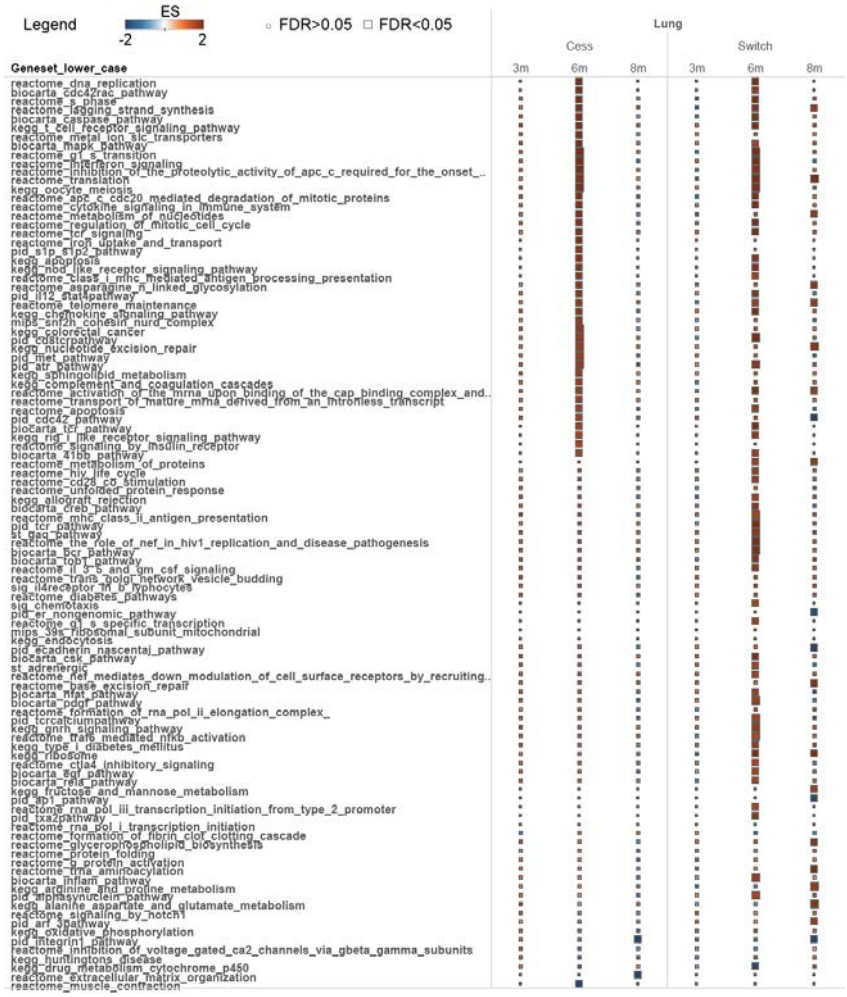


Figure 34: Gene set enrichment analysis in the lungs of animals from the cessation and switch to THS2.2 groups

5.4 Lung proteomics

5.4.1 Number of replicates per group and protein identification

After QC, the following samples were excluded from analysis:

- One sample of the sham group from month 2 (replicate 6): based on principal component analysis (PCA) results

The number of replicates effectively available for each group is summarized in Table 6.

Table 6: Number of valid replicates of lung proteomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	7	8	8	8
3R4F	8	8	8	8	8
THS2.2	8	8	8	8	8
Cessation	0	0	8	8	8
Switch	0	0	8	8	8

In total, 5428 proteins were identified in the experiment. However, for the identification of differentially expressed proteins compared with those in the sham exposure group, only those proteins identified in at least two thirds of one group were considered, as shown in Table 7.

Table 7: Number of consistently quantified lung proteins

Exposure group/time point (months)	1	2	3	6	8
3R4F	2,080	2,324	1,993	1,999	2,273
THS2.2	2,078	2,324	1,993	2,000	2,273
Cessation	-	-	1,993	2,000	2,271
Switch	-	-	1,993	2,000	2,273

5.4.2 Differentially expressed proteins in the lung

For all exposure conditions and time points, the amplitude (expressed as log-fold change, x-axis) and significance (expressed as -log of the adjusted *P*-value) of protein expression changes in lung were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 35.

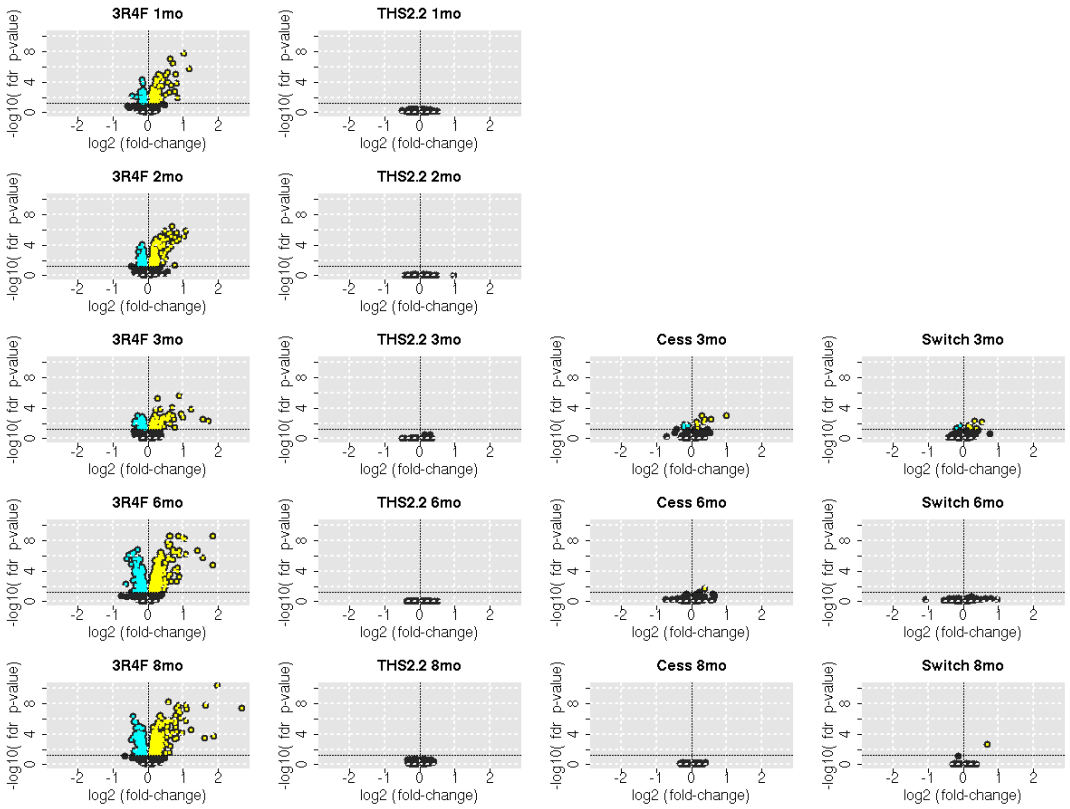


Figure 35: Volcano plots representing the proteome response profiles in the lung

The protein expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight proteins that are significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point (BH-adjusted *P*-value <0.05).

As a summary of the exposure effects, the number of differentially expressed proteins for each comparison with the respective sham group is shown in Figure 36 and Figure 37. In each case, the FDR was controlled at the 5% level using the BH procedure.

Exposure to 3R4F resulted in a clear time-dependent effect on the lung proteome, with the expression of more than 700 proteins significantly altered after an 8-month 3R4F exposure. The number of differentially expressed proteins particularly increased between months 3 and 6. By contrast, exposure to THS2.2 resulted in no significant change to protein regulation during the entire 8-month exposure period. Both cessation and switching to 3R4F resulted in a comparable and rapid loss of the observable 3R4F exposure effects: at the 3-month time point (one month after switching/the start of cessation), differential regulation was still detectable in 13 and eight proteins in the cessation and switching groups, respectively. However, at 8 months, levels had approached those of the sham group, with metallothionein-1 the single remaining up-regulated protein in the switching group.

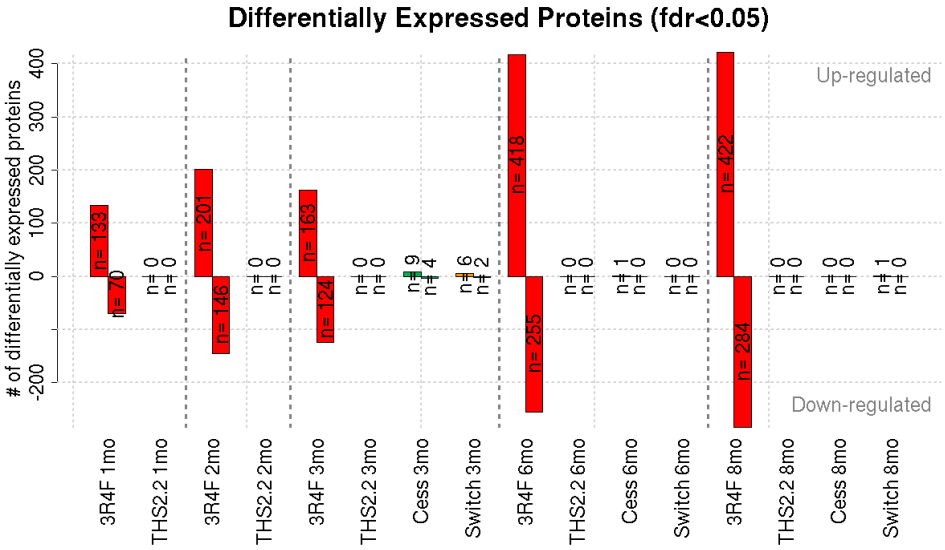


Figure 36: Numbers of differentially expressed proteins in the lung

For each comparison with the sham group, the number of differentially up- and down-regulated proteins is shown. The FDR was controlled at the 5% level using the BH approach.

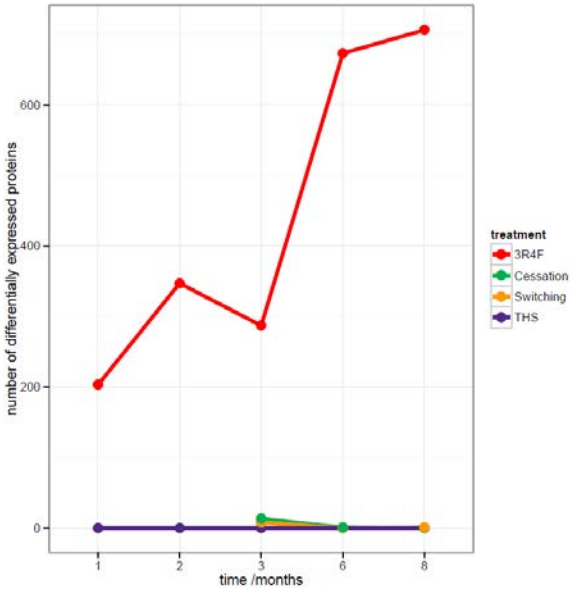


Figure 37: Numbers of differentially expressed proteins over time in the lung

Summary plot of differentially expressed proteins over time for different treatment conditions.

To determine whether the same proteins were differentially regulated in different comparisons, the overlap between the lists of differentially up- or down-regulated proteins was visualized (Figure 38). The generally high overlap between the sets of differentially regulated proteins underlines the consistency of the observed biological responses, e.g. 76% of the proteins up-regulated following 3R4F exposure for 1 month were still up-regulated after 8 months of 3R4F exposure, while 83% of the proteins up-regulated in the switching group after 3 months were also up-regulated after 3 months of 3R4F exposure.

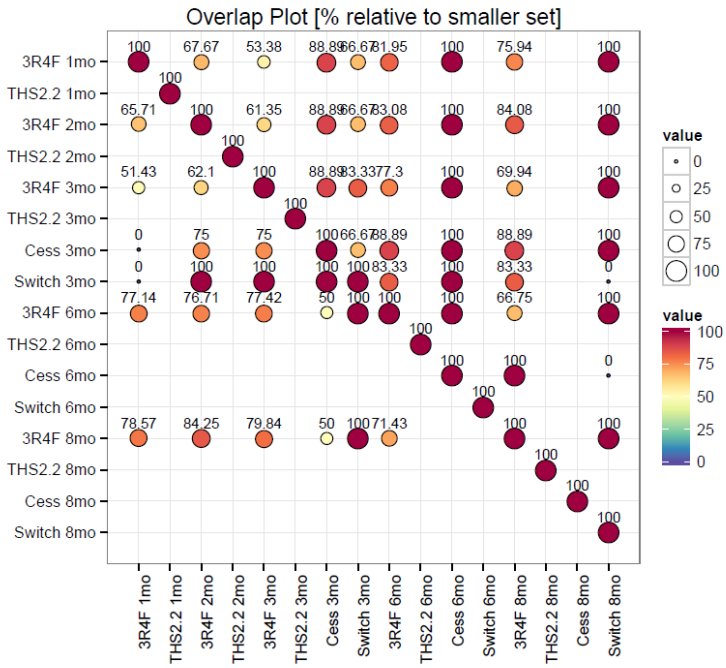


Figure 38: Overlap between sets of differentially expressed proteins in the lung

For each pair of comparisons, the number of shared up- and down-regulated proteins is visualized. The overlap for the up-regulated protein sets is shown in the top right corner, and for down-regulated proteins in the bottom left corner of the matrix. The percent overlap with respect to the smaller set is visualized as the node color and size of each compared pair.

Differentially regulated proteins were also tracked over all conditions in a differential expression heatmap (Figure 39). Again, the time-dependent 3R4F effect is apparent, and only a few of the 3R4F-response proteins remained differentially expressed upon cessation and switching.

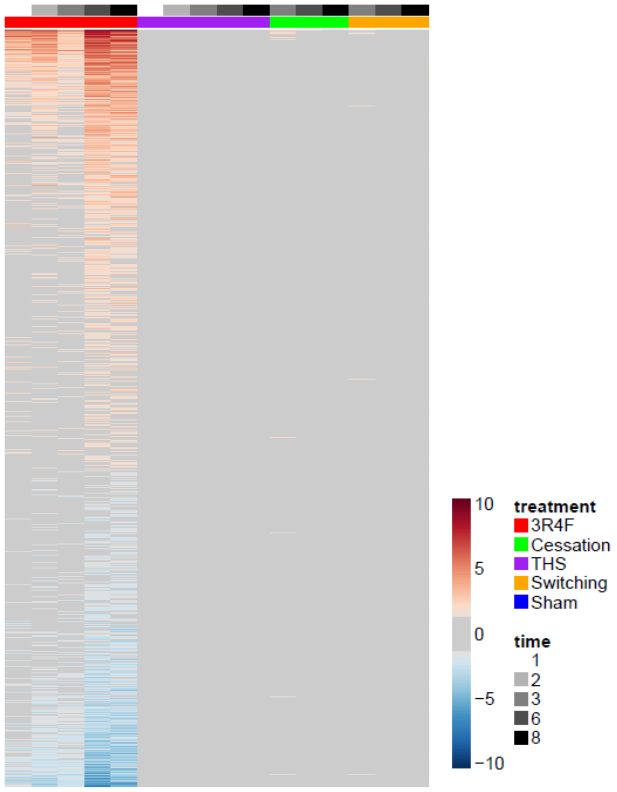


Figure 39: Heatmap of differentially expressed proteins in the lung

Differential protein expression in treatment versus sham groups at the same duration of exposure. Each row represents a differentially expressed protein, each column a sample group comparison (see color key). Only significant changes (FDR<0.05) are colored as the signed –log10 adjusted *P*-value, with red and blue indicating over- and under-expression, respectively, compared with the respective sham group.

5.4.3 Exposure effect comparison for proteins strongly affected by 3R4F exposure

The large number of proteins in the heatmap of Figure 39 made a detailed comparison of their response profiles challenging. Thus, for a more detailed analysis, we focused on the protein subset with the most robust 3R4F response. For this, we selected proteins that were significantly regulated by 3R4F exposure (adjusted *P*-value <0.05) with an up- or down-regulation fold-change >1.3, and compared their response across all conditions (Figure 40). We also compared the exposure effects and variability for selected proteins in this list using boxplots (Figure 41).

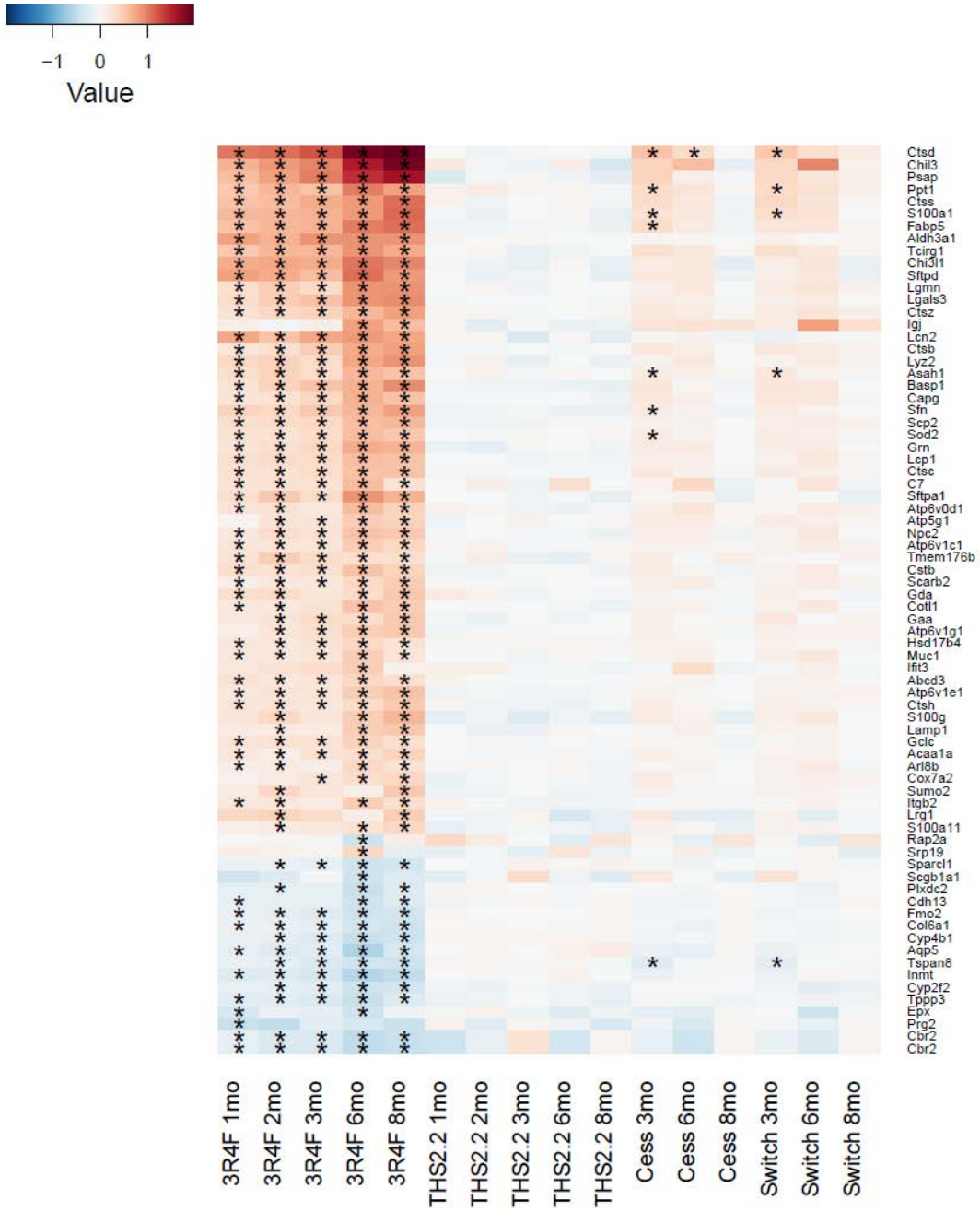


Figure 40: Fold-change expression matrix for proteins strongly affected by 3R4F exposure

Differential protein expression of treatment versus sham at the same exposure duration, visualized as a heatmap. The log2 fold-change is color-coded ("value"). Only proteins significantly regulated by 3R4F exposure with an up- or down-regulated fold-change >1.3 are shown. Differentially regulated proteins with a BH-adjusted *p*-value <0.05 are marked with an asterisk (*).

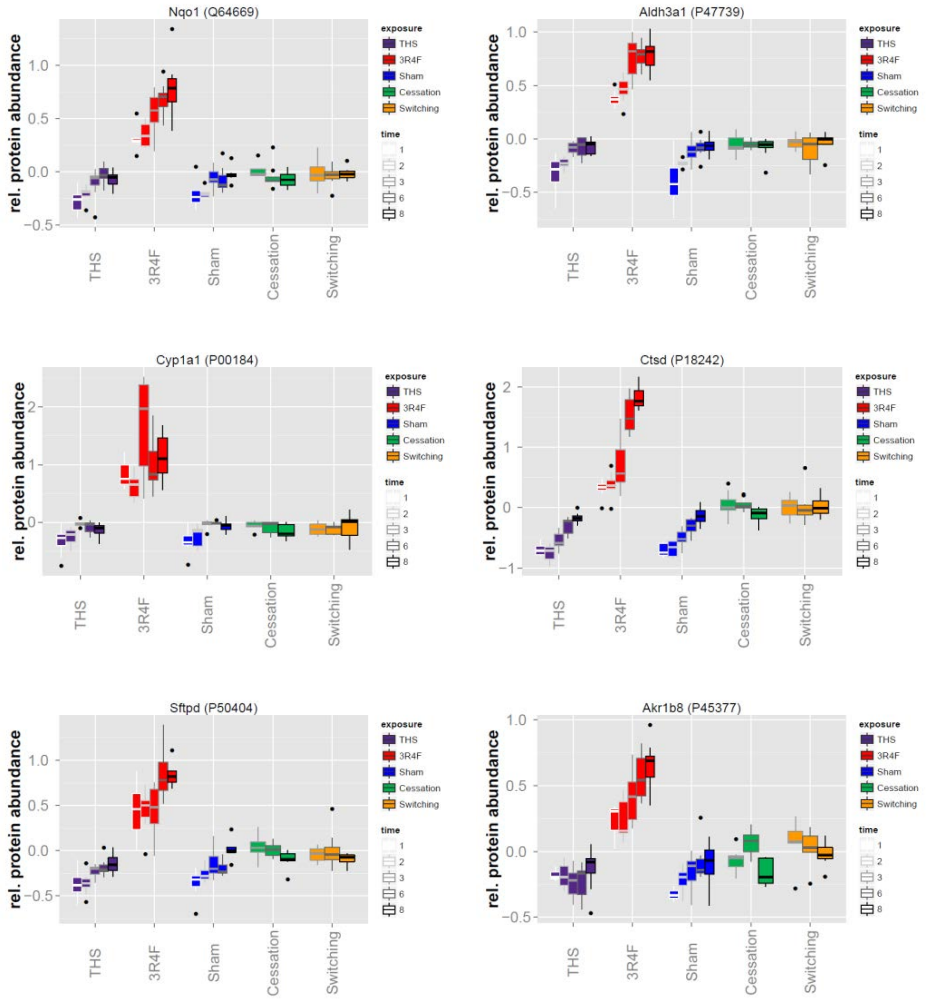


Figure 41: Boxplots showing the protein expression response of selected proteins from Figure 40

Boxplots are grouped by exposure condition and arranged with increasing exposure time from left to right.

This strongly responsive protein subgroup clearly reflected the general response patterns already described for the entire protein set: a time-dependent increase in the 3R4F-exposure effect, a limited to absent response to THS2.2 exposure, and a rapid and further progressing cessation and switching response, which closely approached sham group levels at the 8-month time point.

Comparing the response profiles with (Figure 40) and without (Figure 41) sham normalization highlights the importance of exposure time-matched control groups. For example, cathepsin D (Ctsd) levels appeared to increase with exposure time even in sham-exposed animals.

Proteins with a strong 3R4F response included xenobiotic metabolism enzymes such as Nqo1, Aldh3a1, Cyp1a1, and Akr1b8, the lysosomal Ctsd protein (potentially, macrophage associated (33)), and the lung surfactant protein D (Sftpd) (Figure 41).

5.4.4 Analysis of activated functional protein networks

To understand and compare the functional impact of 3R4F exposure on the lung proteome, we conducted an activated functional subnetwork analysis. For this, the regulated proteins in the group with the largest overall effect (3R4F, 8-month time point) served as the query set. Functional interactions for this protein set were extracted from the STRING database, a comprehensive resource of confidence-scored functional protein interactions (27), and functional clusters in the derived network were identified and annotated (Figure 42). Using this network as a template, regulation of the selected proteins and clusters was compared across all exposure conditions (Figure 43, Figure 44, and Figure 45).

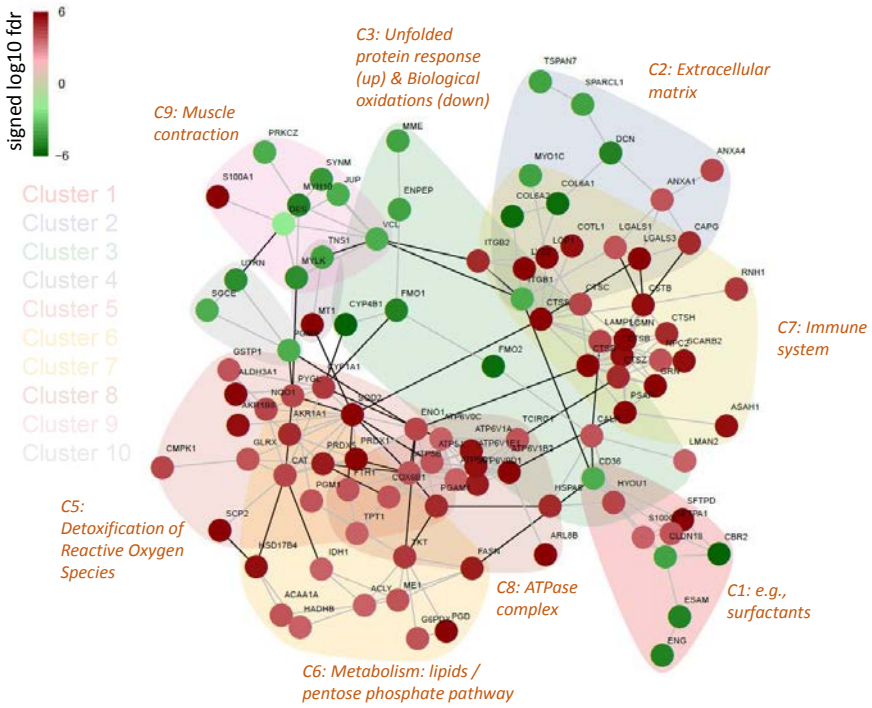


Figure 42: Functional interaction network for significantly regulated proteins in the lung after 8 months of 3R4F exposure

Nodes represent the selected differentially regulated proteins in the 3R4F 8-month exposure group. The node color reflects the signed log10 adjusted *P*-value of differential expression (see color key). The edges/links represent functional interactions from the STRING database. Clusters in this network were identified and marked by colors (see key). A summary annotation is given next to each cluster. The dnet method was used to define this template network (see Methods section).

A total of 10 functional clusters regulated by 3R4F exposure were identified (Figure 42). These included an "immune-system" cluster (C7) that likely reflects induced inflammatory processes (34), a "detoxification of reactive oxygen species (ROS)" cluster (C5) predicted to respond to the oxidative challenge of cigarette smoke and to contribute to direct detoxification processes (35,36) together with the induced metabolic alterations ("metabolism" cluster, pentose-phosphate pathway, C6), and an "unfolded protein response" cluster (C3, lower part) that likely responds to the stress of damaged and misfolded proteins (37). A cluster of "extracellular matrix" proteins (C2) and a "muscle contraction" cluster (C9) were also identified. While the functional relevance of these clusters is less apparent, they could reflect tissue damage processes and may be linked to emphysematous alterations (38,39). However, this requires further investigations.

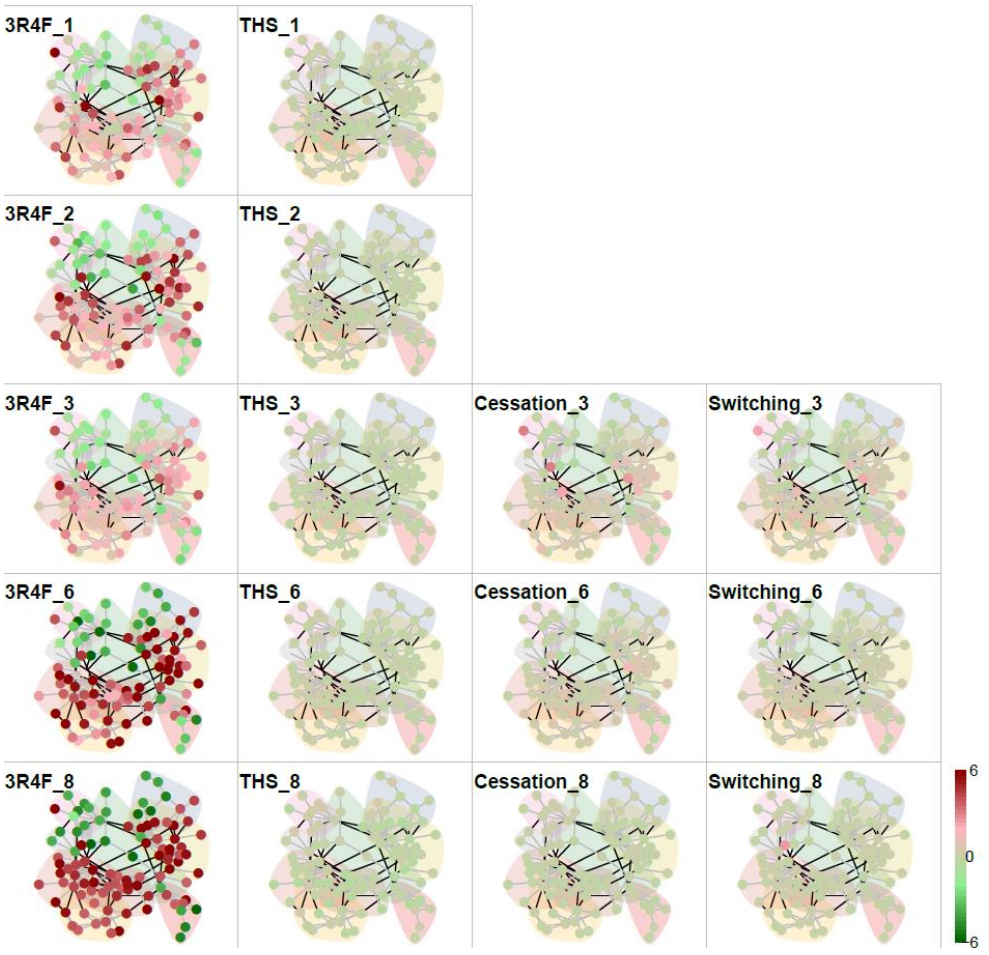


Figure 43: Comparison of exposure effects on the identified functional network across all exposure conditions (signed log10 adjusted *P*-values)

The functional subnetwork from Figure 42 was used as the template and the signed log10 adjusted *P*-value is represented as the protein/node color for each exposure versus sham comparison.

Mapping of the protein alterations onto the functional subnetwork for all exposure conditions underlined the striking difference between 3R4F and the other exposure conditions (Figure 43, Figure 44, and Figure 45). For example, Aldh3a1 and Cyp1a1, two core proteins of the "detoxification of reactive oxygen species" cluster (C5), showed a rapid and strong response after 1 month of 3R4F exposure. Based on the fold-change response (Figure 44), the "immune system" cluster was found to increase in expression over time (C7). This cluster also underwent alterations in the 3-month cessation/switch condition that decreased over time and were absent from the 8-month time point (Figure 45).

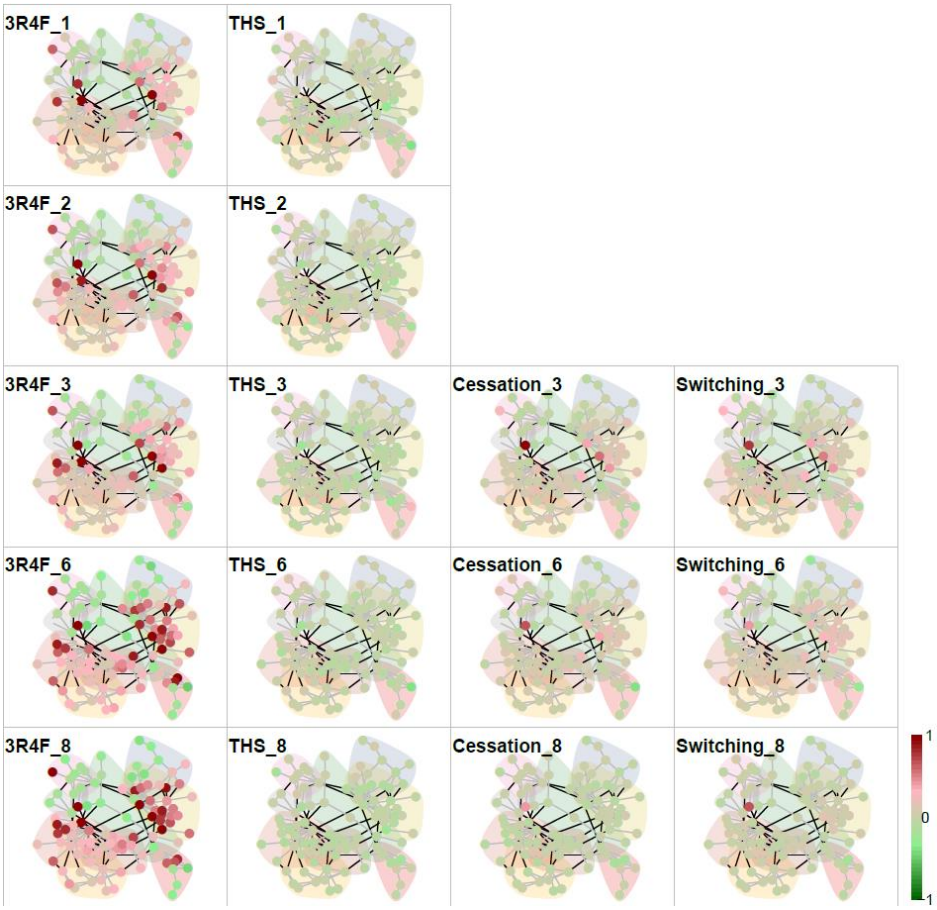


Figure 44: Comparison of exposure effect on the identified functional network across all exposure conditions (log2 fold-changes)

The functional subnetwork from Figure 42 was used as the template and the log2 fold-change value is represented as the protein/node color for each exposure versus sham comparison.

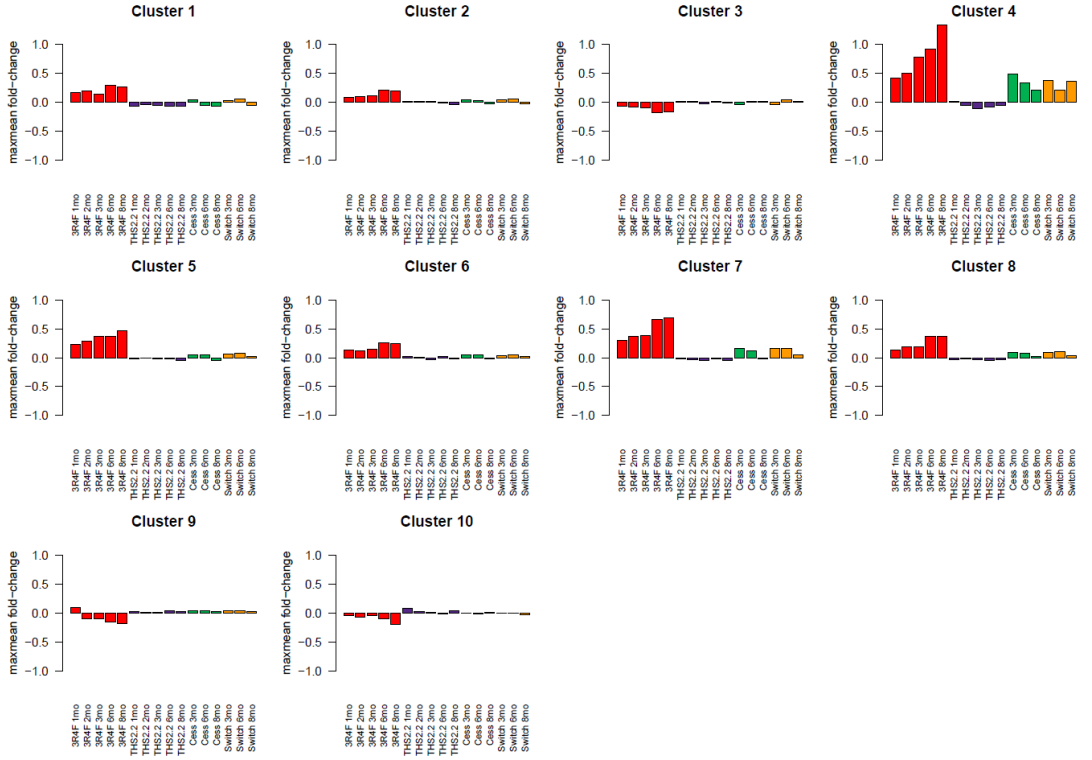


Figure 45: Summary of exposure effects on the identified protein clusters

Maxmean statistics of the fold-changes for each cluster and exposure group. Note that the strong response of cluster 4 is driven by a single up-regulated protein, MT1.

5.4.5 Gene/protein set analysis

We conducted a gene/protein set analysis to complement the functional insights obtained from the subnetwork analysis. We identified protein sets from a large compendium (MSigDB, canonical pathways, (40)) that demonstrated significant average up- or down-regulation of its protein members (Figure 46).

Overall, the protein set summary statistic reflected the patterns observed on the individual protein level: a time-dependent increase in the 3R4F-exposure effect, a limited response to THS2.2 exposure of the main 3R4F-affected protein sets, and a rapid cessation and switching response, which closely approached sham group levels at the 8-month time point.

Protein set analysis commonly results in a less compact representation of the affected biology compared to the functional subnetwork analysis shown above. Nevertheless, identified biological functions affected by 3R4F exposure generally overlapped in our analysis. For example, we identified an effect on the “unfolded protein response” pathway (Cluster 3, REACTOME_UNFOLDED_PROTEIN_RESPONSE), oxidative phosphorylation (Cluster 8/ATPase complex, KEGG_OXIDATIVE_PHOSPHORYLATION), immune system (Cluster 7, REACTOME_MHC_CLASS_II_ANTIGEN_PRESENTATION), and lipid metabolism (Cluster 6, REACTOME_METABOLISM_OF_LIPIDS_AND_LIPOPROTEINS).

A few protein sets were affected by THS2.2 exposure, but their pattern was generally less consistent across conditions and showed a smaller amplitude. Several of these protein sets are linked to translational functions, e.g. KEGG_RIBOSOME and REACTOME_PEPTIDE_CHAIN_ELONGATION, and do not necessarily show a consistent response, e.g. between 3R4F and THS2.2 exposure.



Figure 46: Protein set analysis for the lung

Protein set analysis results based on (absolute) t-scores and the mean as the protein set statistic. Only protein sets that were significantly regulated (BH-adjusted *P*-value <0.05, *) in at least one condition are shown. The color scale reflects the protein set statistic.

5.4.6 Summary of the observed effects on the lung proteome

The exposure effect on the lung proteome was assessed with a quantitative, mass spectrometry-based approach. A good coverage of the lung proteome was achieved, with a total of 5,428 proteins identified and ~2,000–2,300 consistently quantified proteins for each exposure comparison.

Overall, the response patterns were in line with the other endpoints. 3R4F exposure demonstrated an effect on the lung proteome that increased over time with an especially large effect after 6 and 8 months. THS2.2 exposure did not result in any significant differential regulation of the lung proteome. Cessation and switching to THS2.2 resulted in a comparable, rapid reduction of the observable differential regulation of the lung proteome (Figure 35, Figure 36). A reduction in the exposure response for the cessation/switching groups was particularly visible for the selected group of proteins strongly affected by 3R4F exposure (Figure 40).

The functional impact of 3R4F on the lung proteome was assessed by a functional subnetwork approach. This approach highlighted affected biological clusters including an "immune-system", a "detoxification of reactive oxygen species", a "metabolism" (including pentose-phosphate pathway enzymes), a "xenobiotic phase I", and an "ER stress" cluster (Figure 42). Additional identified functional clusters such as "extracellular matrix" and "muscle contraction" warrant further investigation. Overall, all identified functional clusters followed the generally observed response pattern: an exposure time-dependent increase for 3R4F exposure, a limited to absent response to THS2.2 exposure, and a rapid loss of observable exposure effects upon switching and cessation (Figure 43, Figure 44, and Figure 45). For example, 1 month after switching/cessation (i.e., the 3-month time point) a residual exposure effect was observed for the immune system-related cluster. The analysis of affected protein sets further supported these conclusions on the affected biological functions and general exposure response pattern (Figure 46).

In summary, analysis of the lung proteome identified several well-supported biological effects of cigarette smoke exposure including a xenobiotic-, immune-, and metabolic- response. By contrast, THS2.2 exposure demonstrated a limited to absent response of the lung proteome. Switching to THS2.2 and cessation of smoking resulted in a comparable and rapid loss of the cigarette smoke exposure effects.

5.5 Thoracic aorta transcriptomics

5.5.1 QC and number of replicates per group

The number of replicates effectively available for each group is summarized in Table 8.

Table 8: Number of valid replicates of thoracic aorta transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	8	7	8	7
3R4F	8	8	7	8	7

THS2.2	8	8	7	7	7
Cessation	-	-	8	8	8
Switch	-	-	6	7	5

5.5.2 Differentially expressed genes in the thoracic aorta

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of gene expression changes in the thoracic aorta for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 47. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 48.

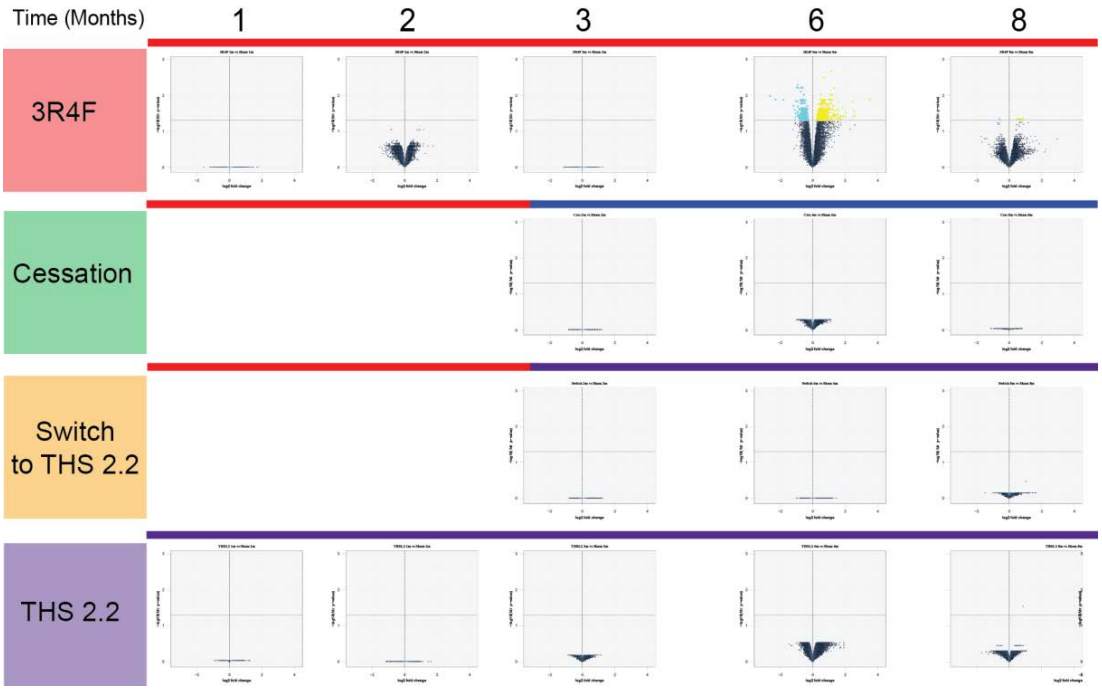


Figure 47: Volcano plots representing the systems response profiles in the thoracic aorta: FDR
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that are significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

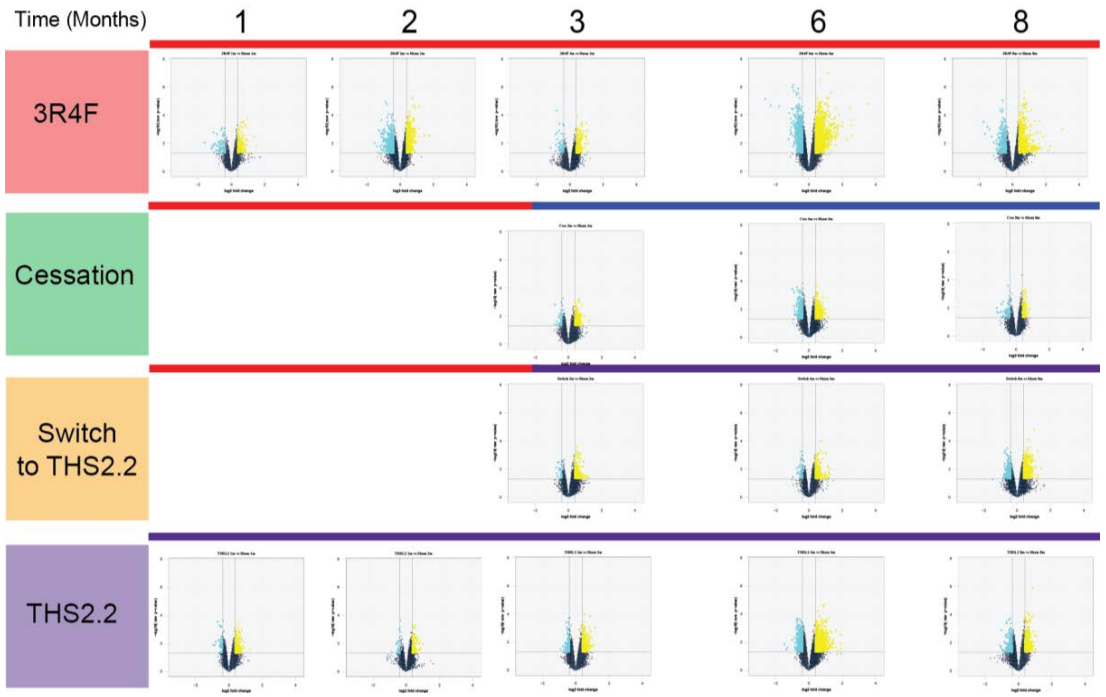


Figure 48: Volcano plots representing the systems response profiles in the thoracic aorta: raw *P*-value

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold-change >1.2.

The number of differentially expressed genes compared with sham for each exposure group is shown in Table 9. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, a pairwise comparison of the overlap was carried out, as shown in Figure 49.

Table 9: Number of differentially expressed genes in the thoracic aorta (up-/down-regulated compared with sham at each respective time point)

Exposure group/time point (months)	1	2	3	6	8
Sham	REF	REF	REF	REF	REF
3R4F	0	0	0	325/226	9/1
THS2.2	0	0	0	0	1/0
Cessation	-	-	0	0	0
Switch	-	-	0	0	0

REF: reference

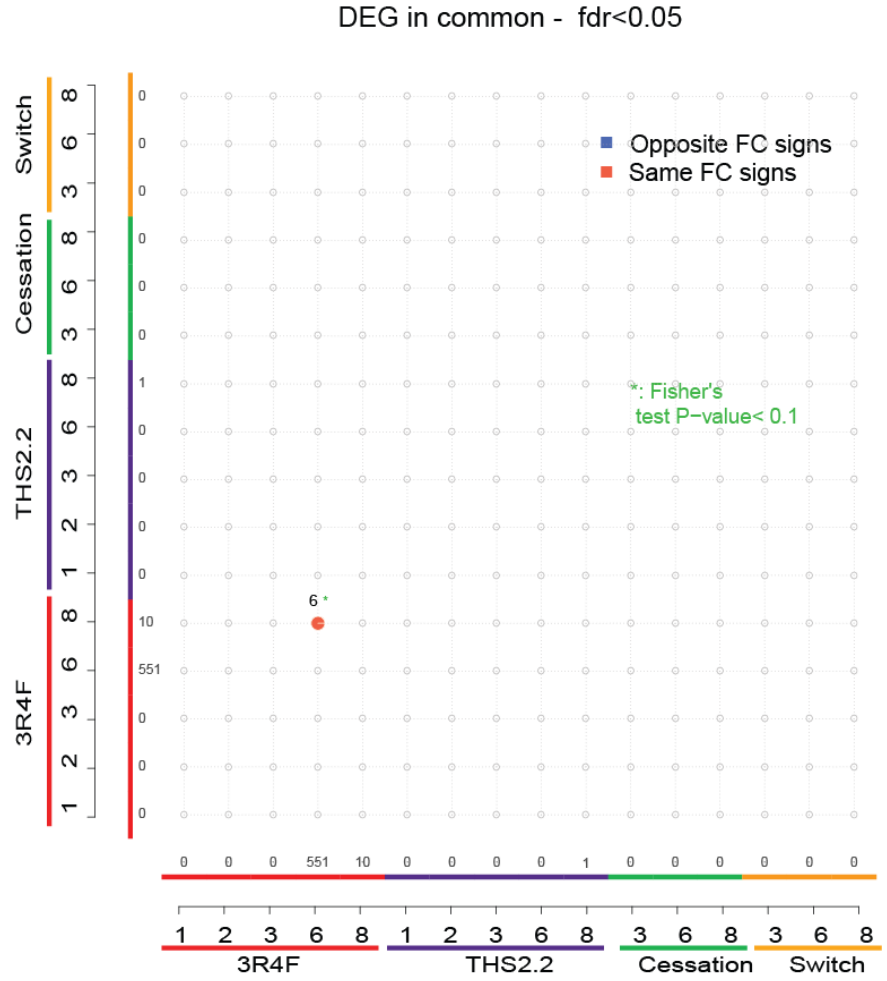


Figure 49: Overlap between sets of differentially expressed genes in the thoracic aorta

For each pair of comparisons, the number of shared up- and down-regulated genes is visualized. The overlap of significantly regulated genes regulated in the same direction (either both up- or both down-regulated) is indicated in red, while opposite direction changes are shown in blue. A green star signifies that the overlap is significant (Fisher's test *P*<0.1). The number of significantly regulated genes (FDR<0.05) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.5.3 BIF analysis in the thoracic aorta

The overall impact of the different exposures was estimated from a BIF (32), expressed as an RBIF (Figure 50). It is important to note that not all networks were built for a vascular context. Their primary application was meant for pulmonary tissues, so the results need to be interpreted with caution.

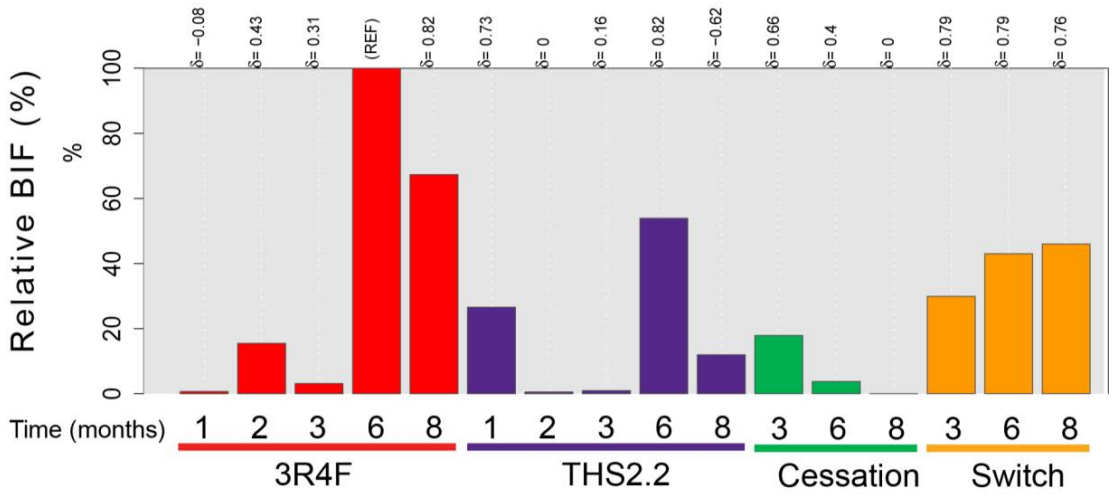


Figure 50: BIF analysis in the thoracic aorta

The percentages show the relative biological impact derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 8 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 8 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.5.4 Biological network perturbation amplitude analysis in the thoracic aorta

The star plots in Figure 51 show the biological processes that most contributed to the overall BIF phenotype at each time point.

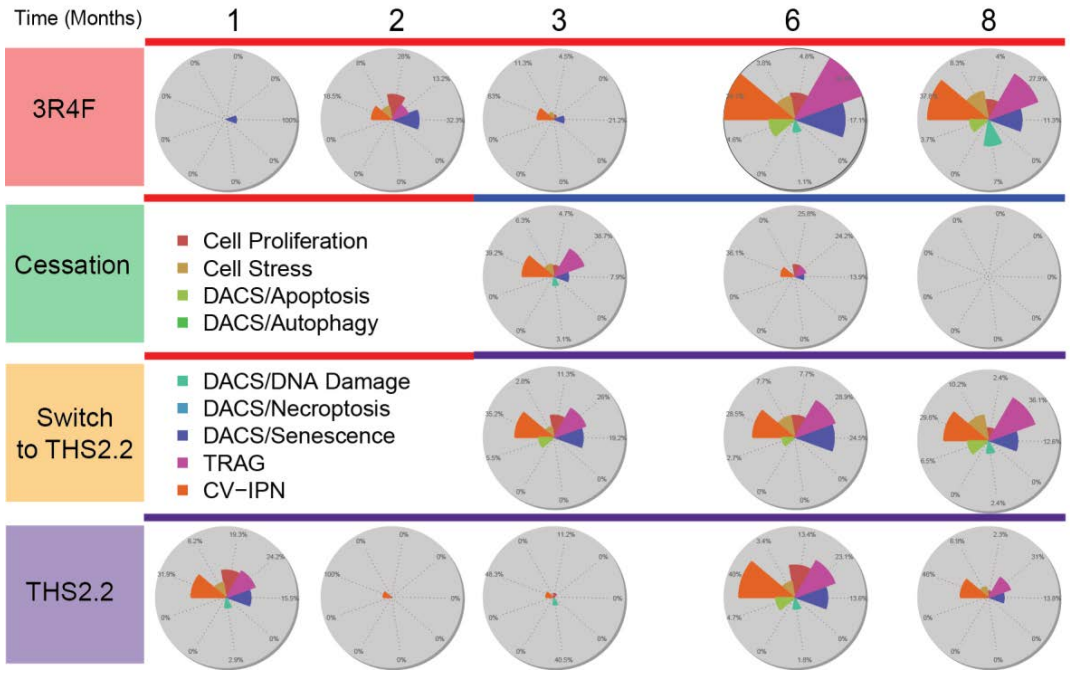


Figure 51: Star plots showing BIF calculated from the thoracic aorta

Star plots illustrating the decomposition of the overall RBIF (Figure 50) into its eight mechanistic components (from cell proliferation to inflammation, indicated by the colors of the segments) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

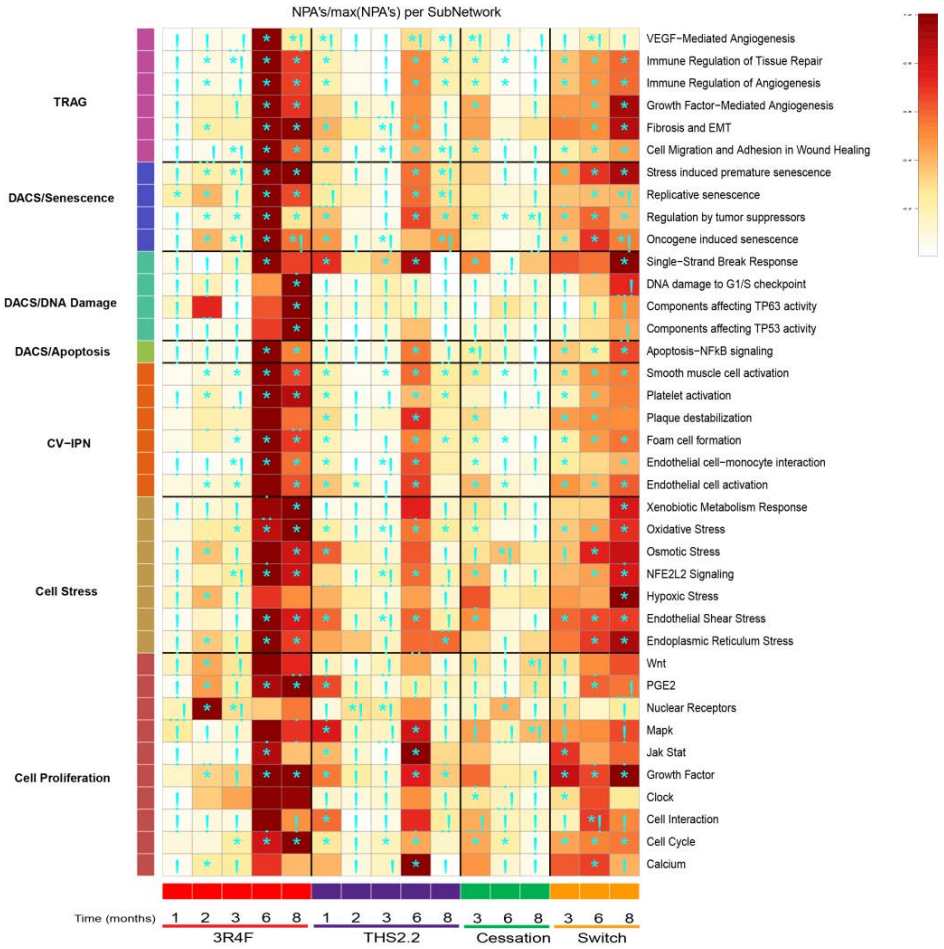


Figure 52: Heatmap of network perturbation amplitude scores in the thoracic aorta

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network was considered perturbed if, in addition to the significance of the NPA score with respect to the experimental variation, the two companion statistics (O and K) derived to inform on NPA score specificity with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values <0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to the experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrast comparisons, similar to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 53).

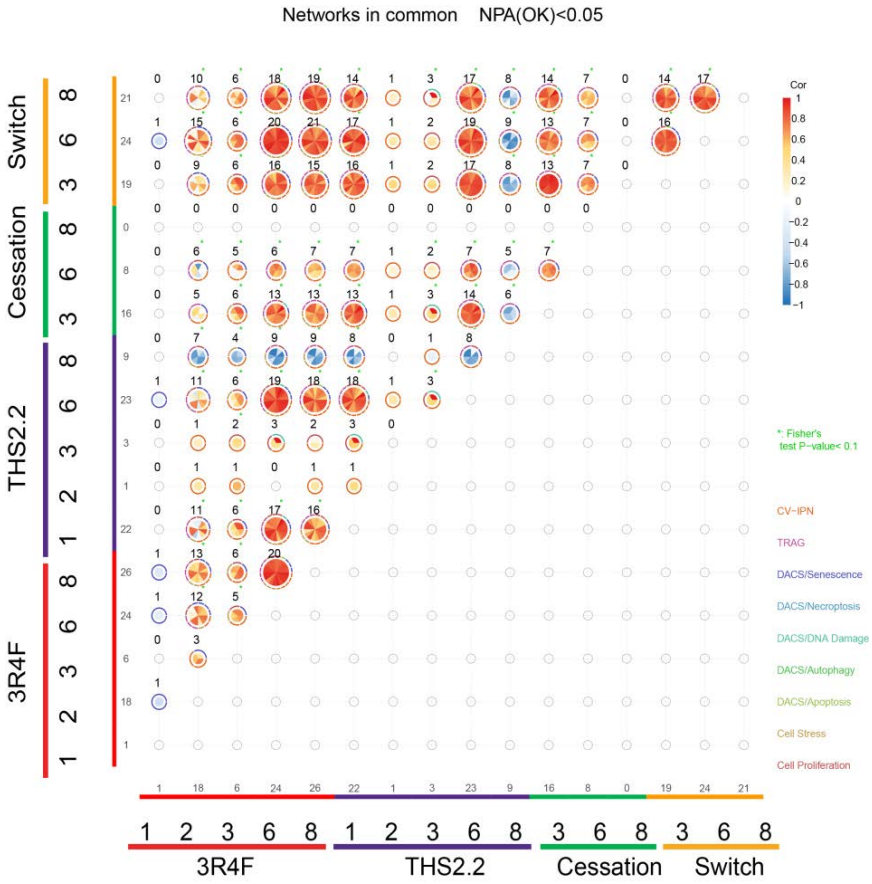


Figure 53: Overlap between sets of significantly perturbed networks in the thoracic aorta
For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates positive correlation, while blue indicates negative correlation. A green star signifies that the overlap is significant (Fisher's test P<0.1). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point)

5.5.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by the different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 54).

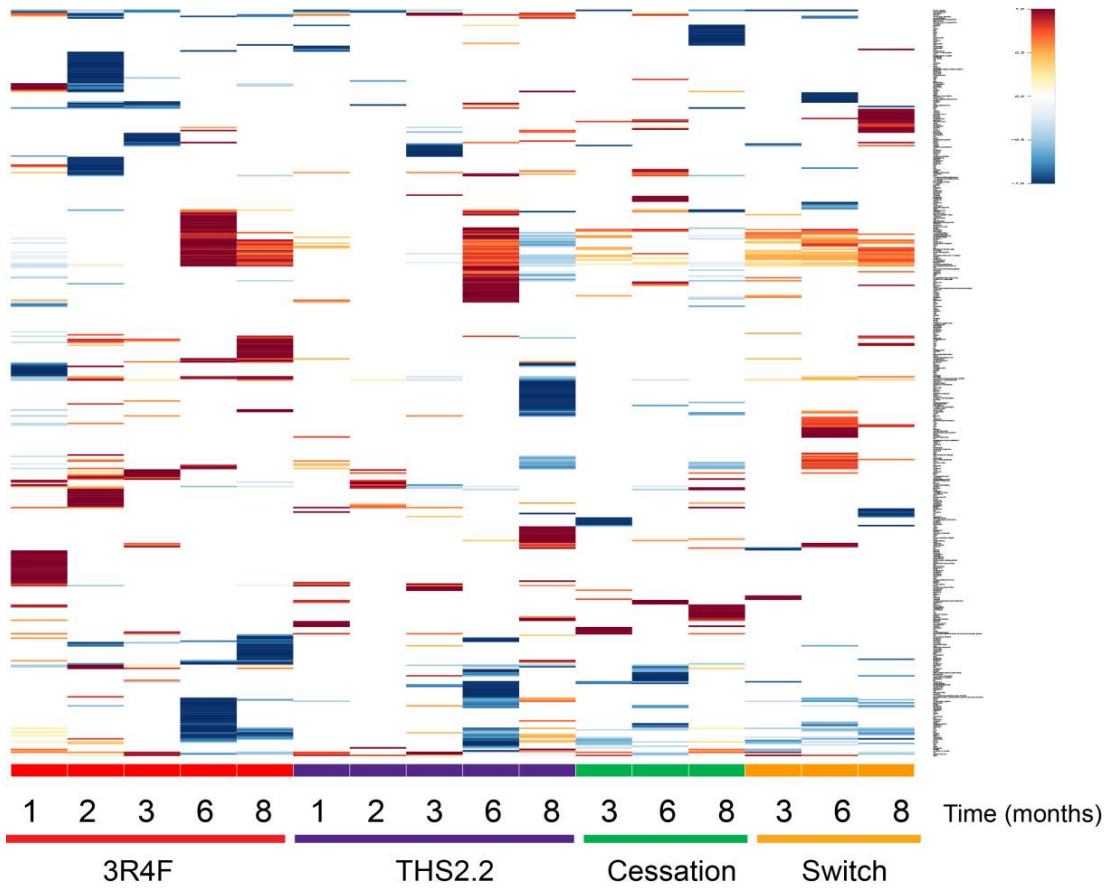


Figure 54: Heatmap of NPAI score for all hyps significant in at least one contrast in the aorta

5.5.6 NextBio analysis

5.5.6.1 3R4F exposure effect

To study the effects of 3R4F on the aorta, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed thoracic aortas at five different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly associated with xenobiotic metabolism and lipid transport (Figure 55).



Figure 55: NextBio meta-analysis of 3R4F effects in the thoracic aorta

Each of the 10 graphs represents the score (expressed as –log10 (*p*-value)) of a canonical pathway for the different time points of exposure to 3R4F (compared with sham-exposed control). Each bar is composed of two parts: the red part is based on the strength of the overlap between the biogroup and the up-regulated genes, while the green part is based on the strength of the overlap between the biogroup and the down-regulated genes.

5.5.7 Gene set enrichment analysis

5.5.7.1 3R4F effects

To further study the trends within each treatment group over time, gene set enrichment analysis was carried out.

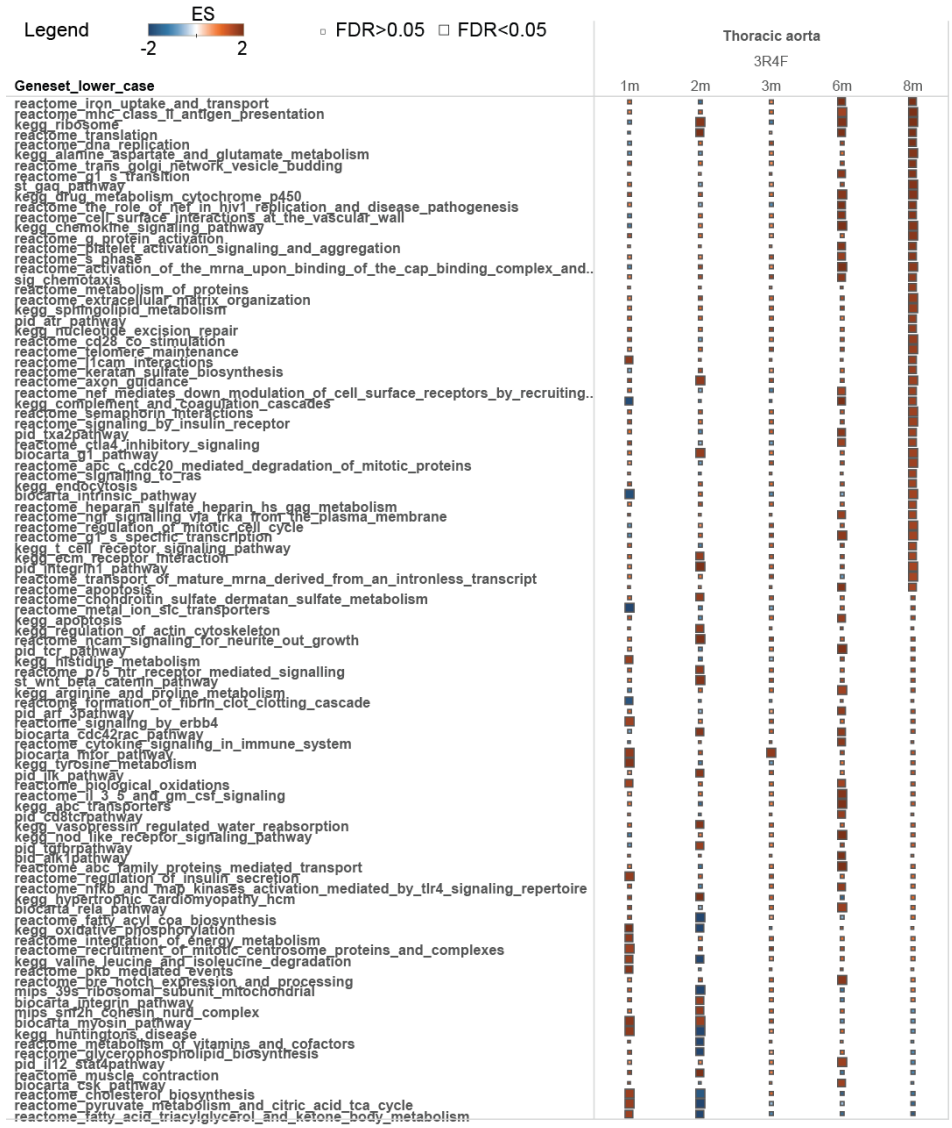


Figure 56: Gene set enrichment analysis in the 3R4F-exposed thoracic aorta

5.5.7.2 THS2.2 effects

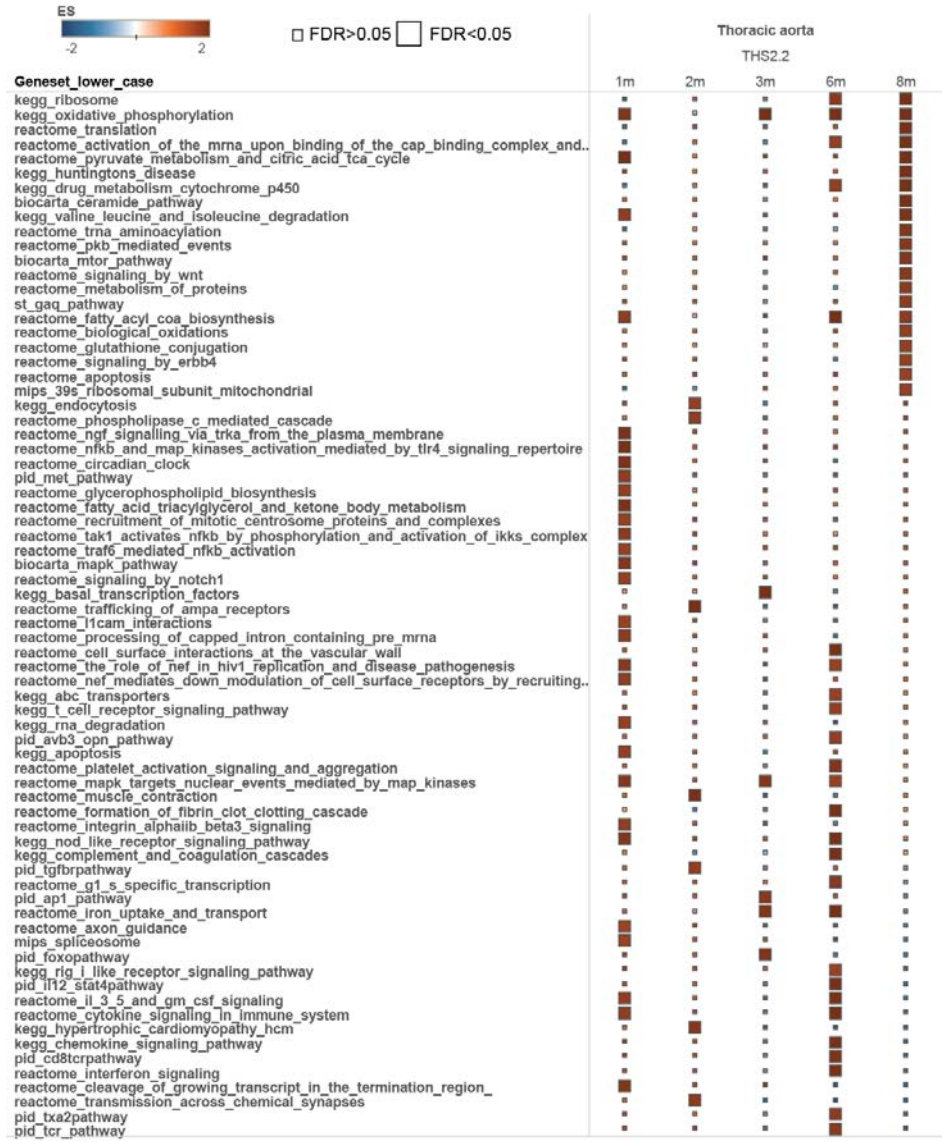


Figure 57: Gene set enrichment analysis in the THS2.2-exposed thoracic aorta

5.5.7.3 Cessation vs. switch to THS2.2 effect

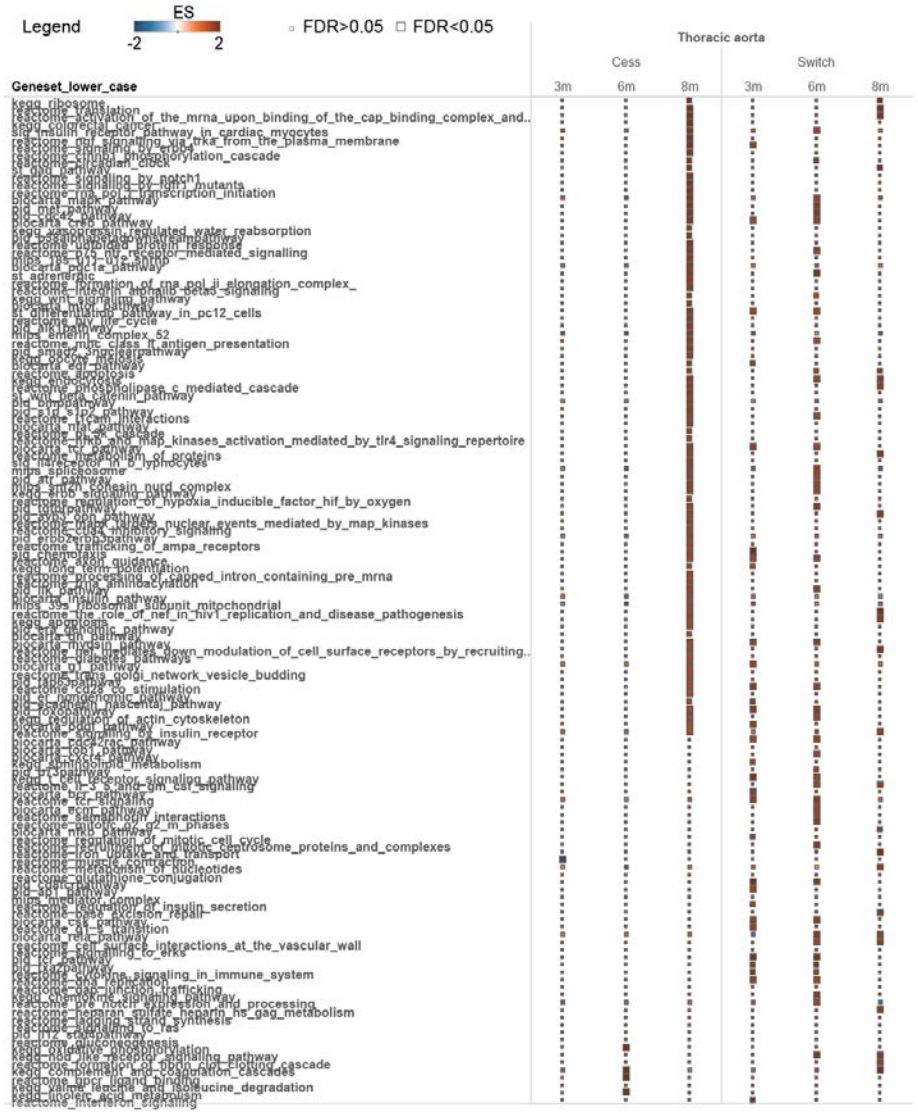


Figure 58: Gene set enrichment analysis in the aorta of animals from the cessation and switching to THS2.2 groups

5.6 Liver transcriptomics

5.6.1 QC and number of replicates per group

The number of replicates effectively available for each group is summarized in Table 10.

Table 10: Number of valid replicates of liver transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	8	8	8	8
3R4F	8	8	8	8	8
THS2.2	8	8	8	8	8
Cessation	-	-	8	8	8
Switch	-	-	8	6	8

5.6.2 Differentially expressed genes in the liver

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of gene expression changes in the thoracic aorta for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 59. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 60.

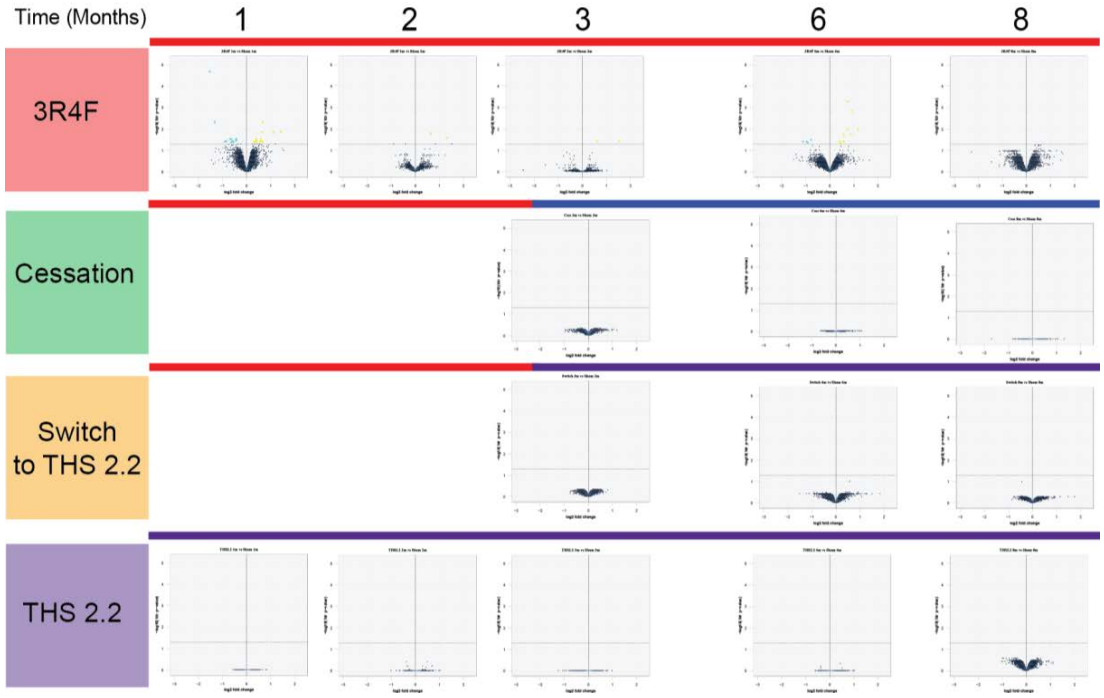


Figure 59: Volcano plots representing the systems response profiles in the liver: FDR

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that are significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

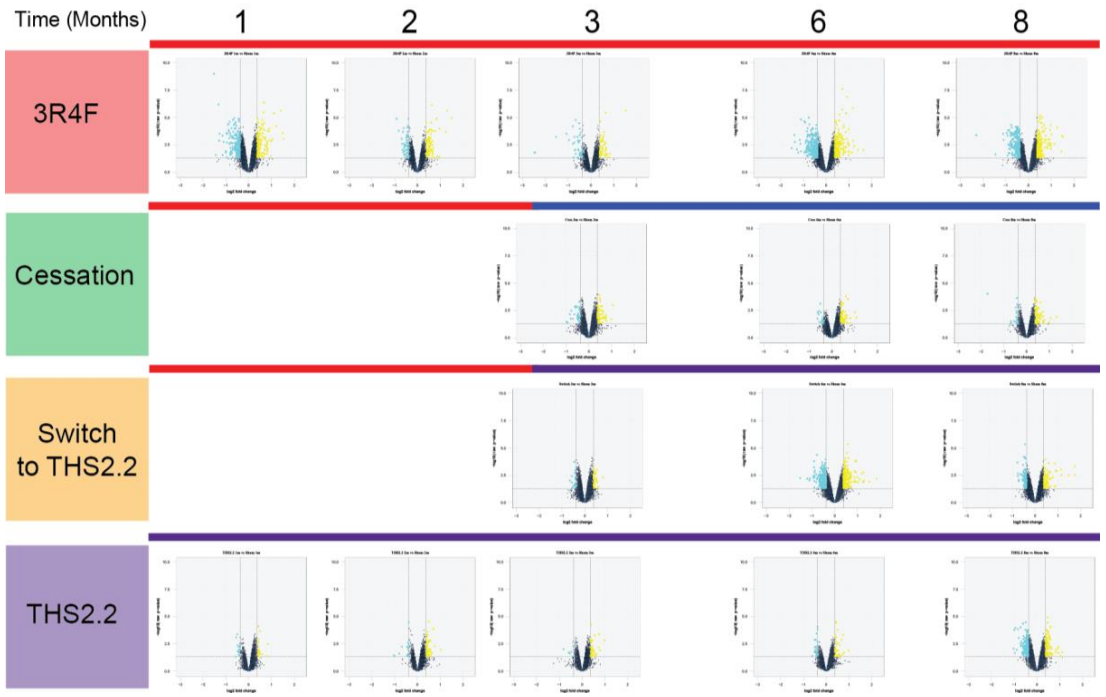


Figure 60: Volcano plots representing the systems response profiles in the liver: raw *P*-value
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold-change >1.2.

The number of differentially expressed genes compared with sham for each exposure group is shown in Table 11. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, the overlap between each pair of contrasts is shown in Figure 61.

Table 11: Number of differentially expressed genes in the liver (up-/down-regulated compared with sham at each respective time point)

Exposure group/time point (months)	1	2	3	6	8
Sham	REF	REF	REF	REF	REF
3R4F	14/14	2/0	2/0	10/4	0
THS2.2	0	0	0	0	0
Cessation	-	-	0	0	0
Switch	-	-	0	0	0

REF: reference

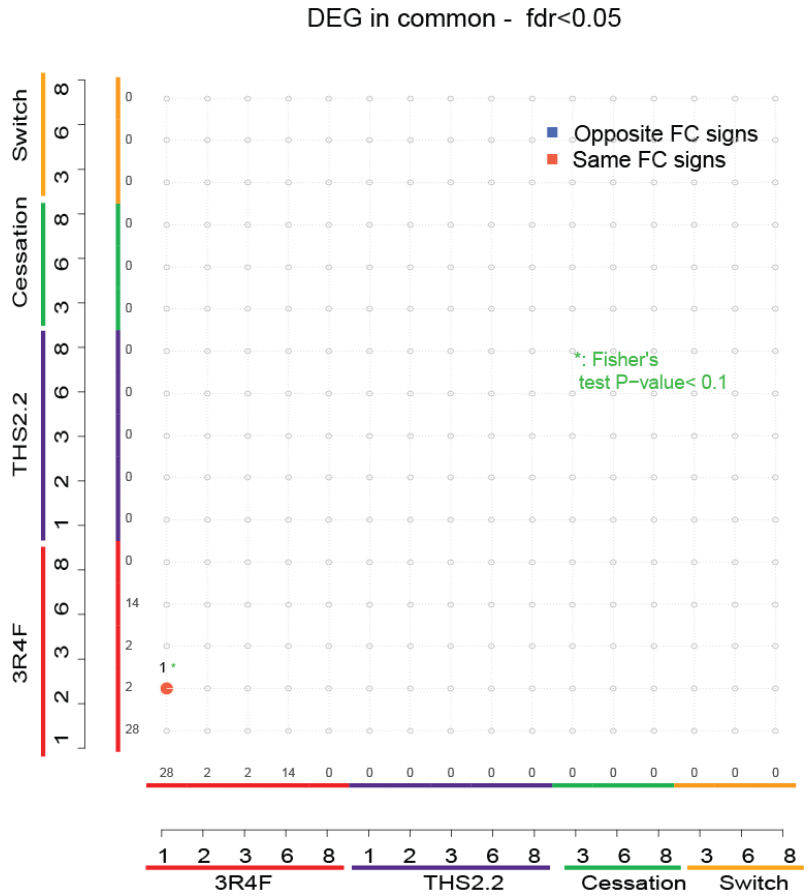


Figure 61: Overlap between sets of differentially expressed genes in the liver

For each pair of comparisons, the number of shared up- and down-regulated genes is visualized. The overlap of significantly regulated genes regulated in the same direction (either both up- or both down-regulated) is indicated in red, while opposite direction changes are shown in blue. A green star signifies that the overlap is significant (Fisher's test *P*<0.1). The number of significantly regulated genes (FDR<0.05) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.6.3 BIF analysis in the liver

The overall impact of the different exposures was estimated from a BIF (32), expressed as a RBIF (Figure 62).

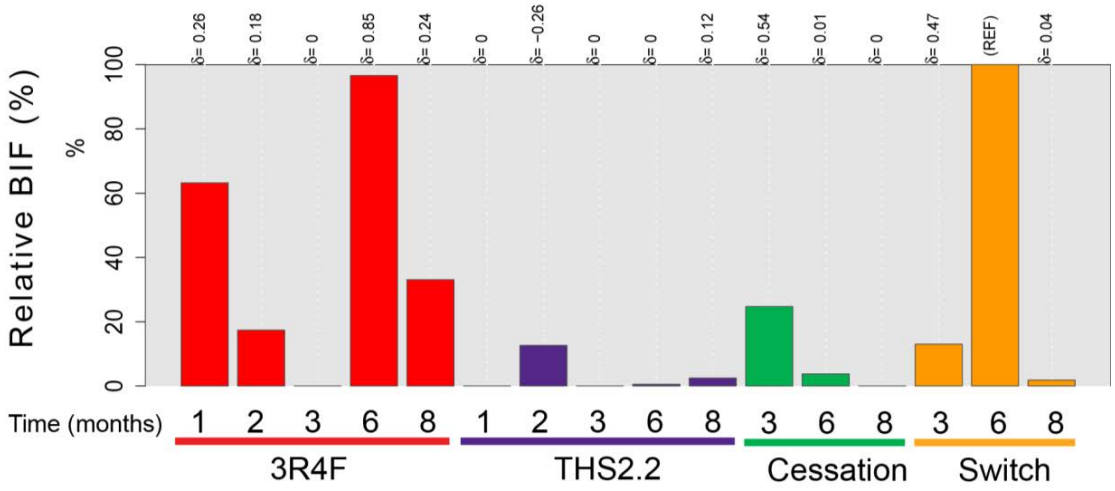


Figure 62: BIF analysis in the liver

The percentages show the relative biological impact, derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 6 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 6 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.6.4 Biological network perturbation amplitude analysis in the liver

The star plots in Figure 63 show the biological processes that most contributed to the overall BIF phenotype at each time point. It should be noted that the networks used were not designed in the liver context and may not be 100% relevant.

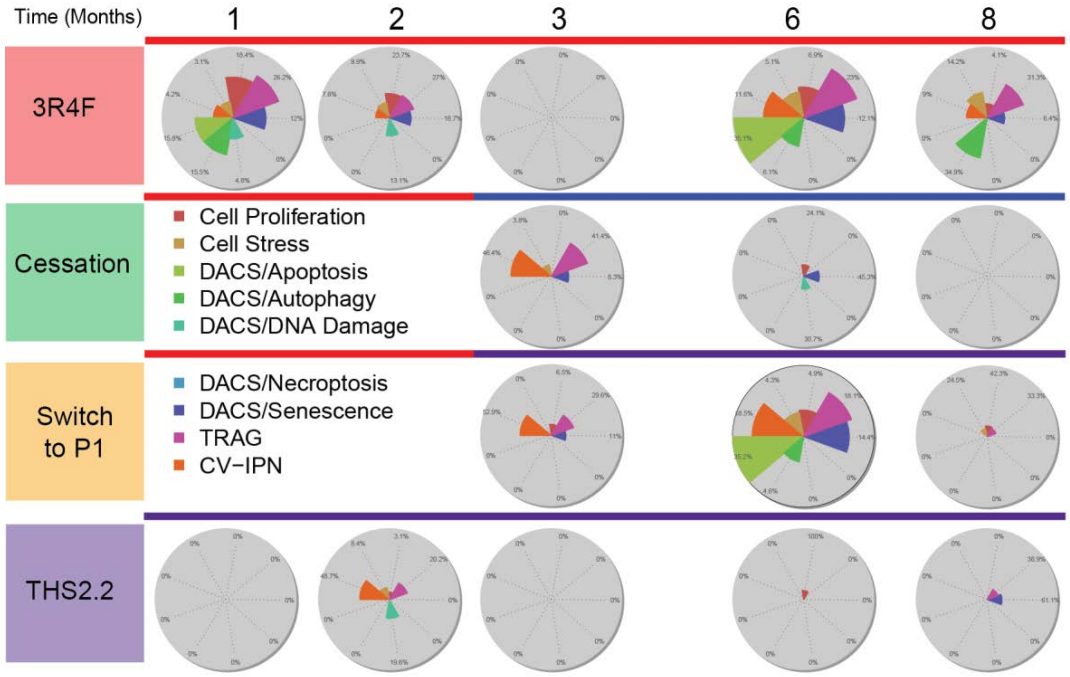


Figure 63: Star plots showing BIF calculated from the liver

Star plots illustrating the decomposition of the overall RBIF (Figure 62) into its eight mechanistic components (from cell proliferation to inflammation, indicated by the colors of the segments) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

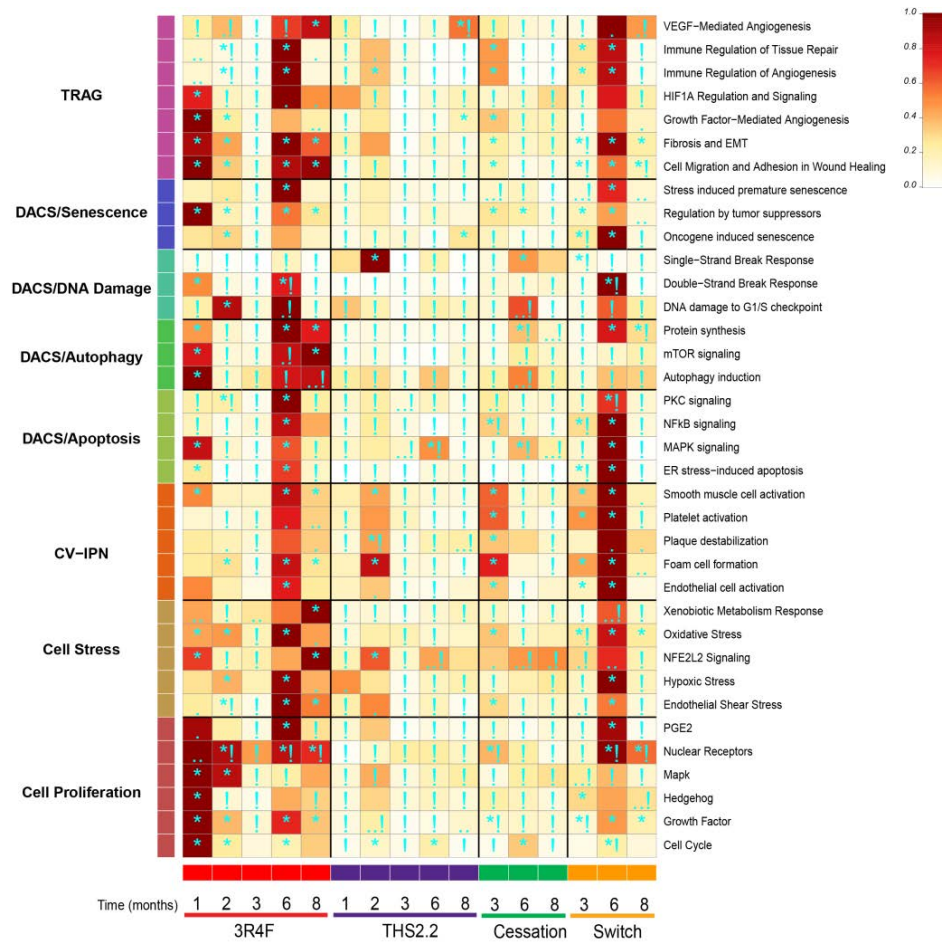


Figure 64: Heatmap of network perturbation amplitude scores in the liver

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network is considered perturbed if, in addition to the significance of the NPA score with respect to the experimental variation, the two companion statistics (O and K) derived to inform on NPA score specificity with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values <0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to the experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrasts comparisons, similar to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 65).

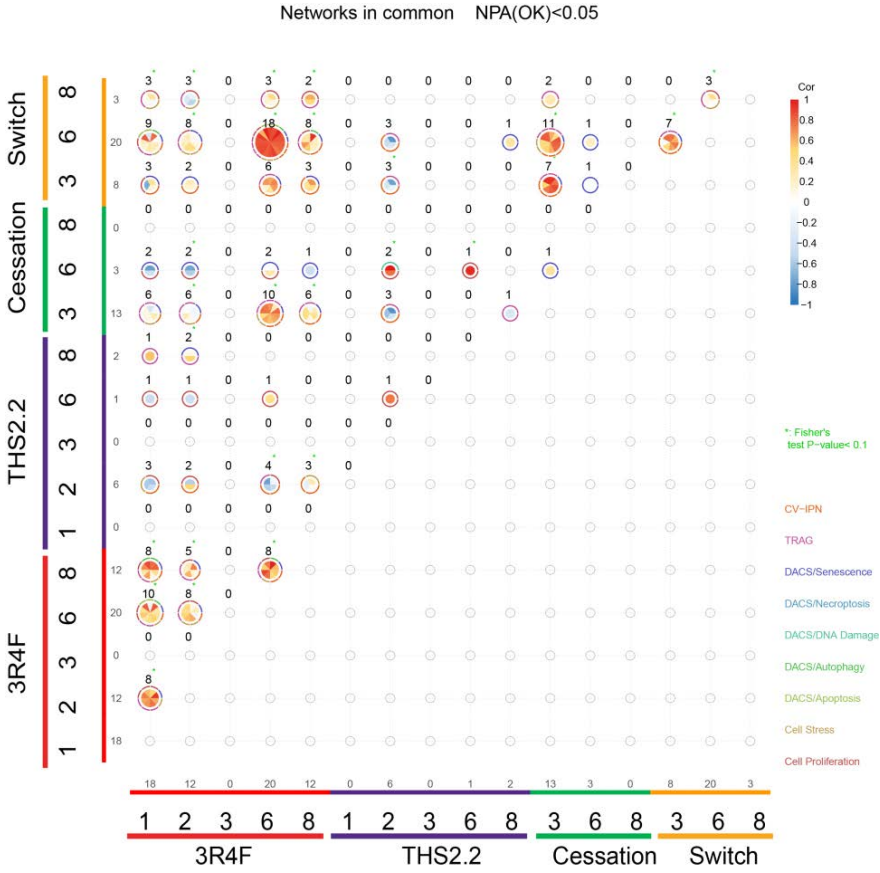


Figure 65: Overlap between sets of significantly perturbed networks in the liver

For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates positive correlation, while blue indicates negative correlation. A green star signifies that the overlap is significant (Fisher's test P<0.1). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.6.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by the different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 66).

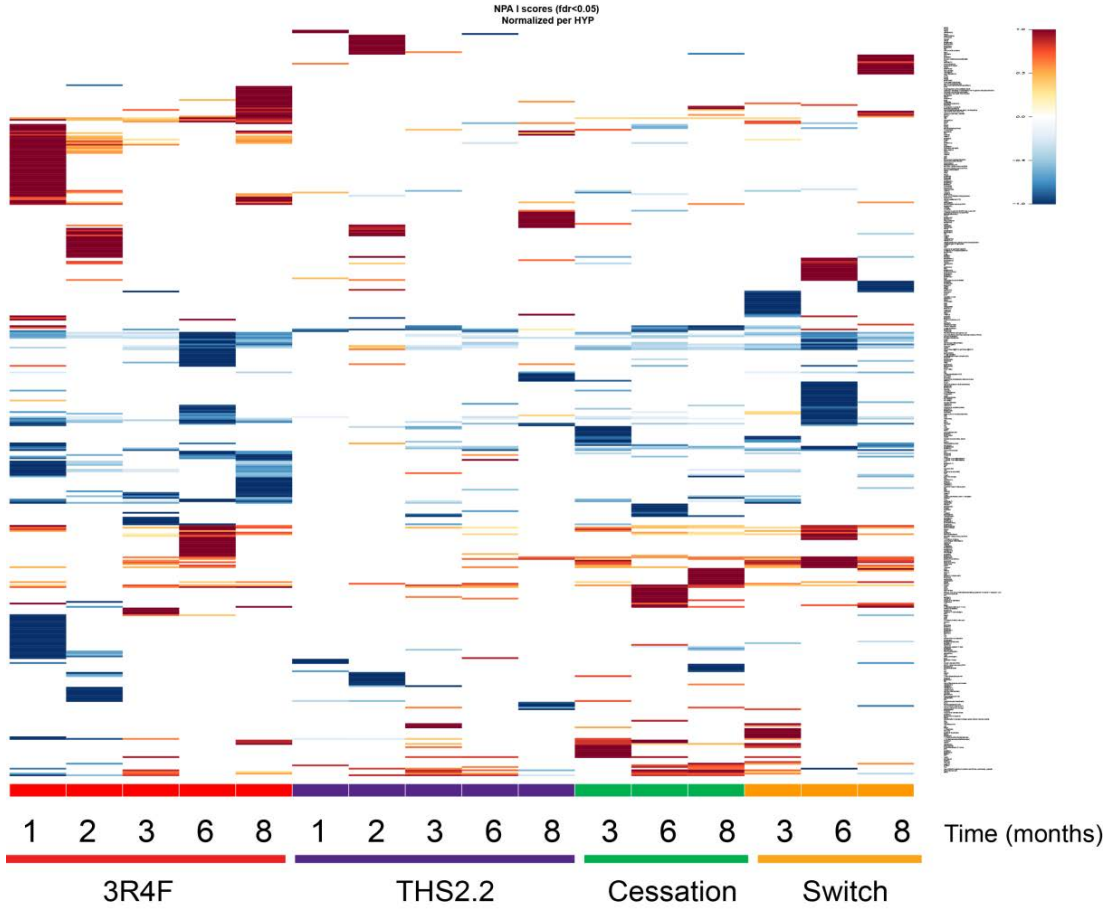


Figure 66: Heatmap of NPAI score for all hyps significant in at least one contrast in the liver

5.6.6 NextBio analysis

To study the effects of 3R4F on the liver, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed livers at five different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly linked to xenobiotic and/or lipid metabolism and the cell cycle (Figure 67).



Figure 67: NextBio meta-analysis of the effects of 3R4F on the liver

Each of the 10 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to 3R4F (compared with sham-exposed control). Each bar is composed of two parts: the red part is based on the strength of the overlap between the biogroup and up-regulated genes, while the green part is based on the strength of the overlap between the biogroup and down-regulated genes.

To study the effects of THS2.2 on the liver, a meta-analysis was performed comparing THS2.2-exposed and sham-exposed livers at five different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly linked to xenobiotic metabolism and the cell cycle (Figure 68).



Figure 68: NextBio meta-analysis of the effects of THS2.2 on the liver

Each of the 10 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to THS2.2 (compared with sham-exposed control). Each bar is composed of two parts: the red part is based on the strength of the overlap between the biogroup and up-regulated genes, while the green part is based on the strength of the overlap between the biogroup and down-regulated genes.

5.6.7 Gene set enrichment analysis

5.6.7.1 3R4F effects

To further study trends within each treatment group over time, gene set enrichment analysis was carried out.

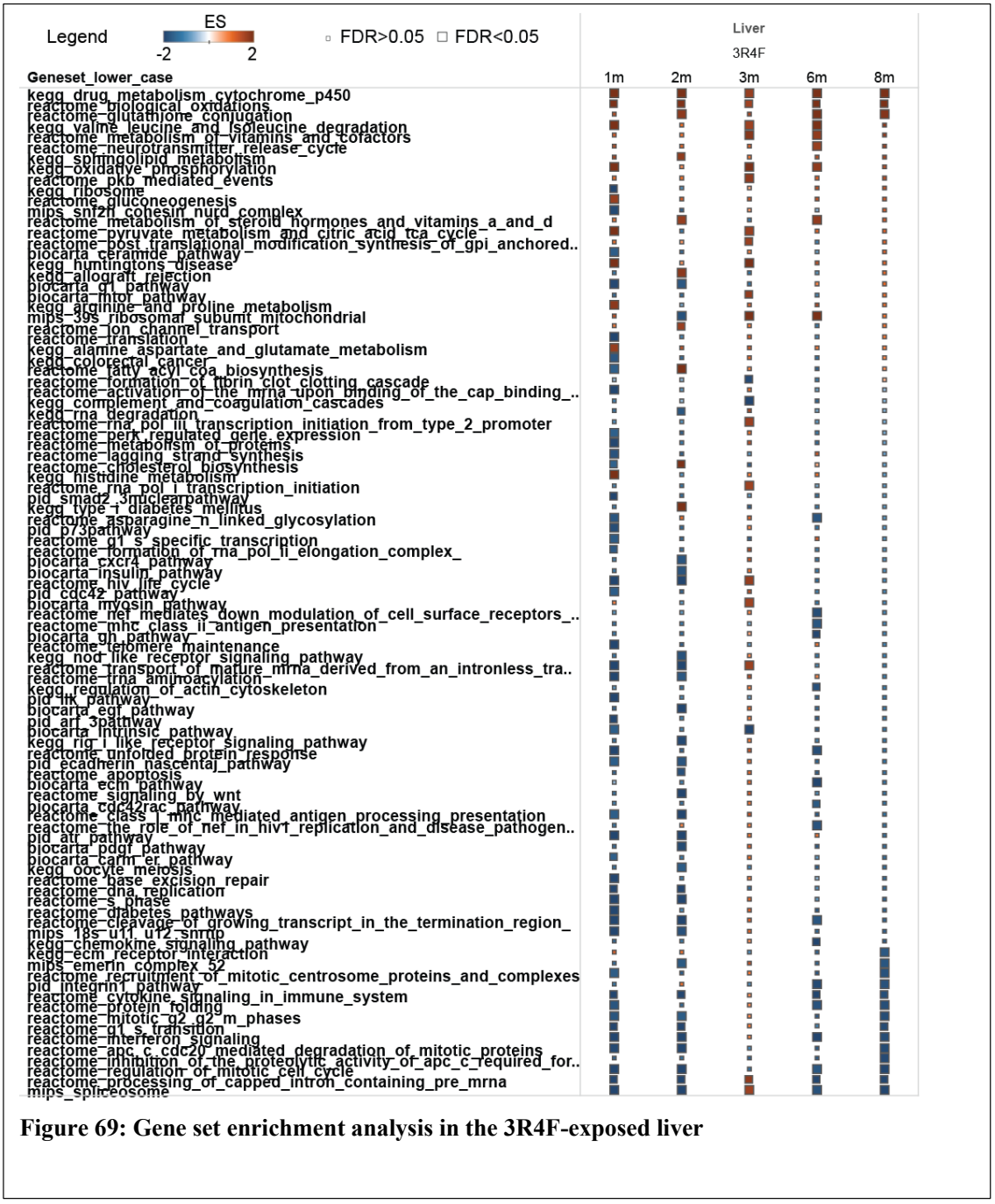


Figure 69: Gene set enrichment analysis in the 3R4F-exposed liver

5.6.7.2 THS2.2 effects



Figure 70: Gene set enrichment analysis in the THS2.2-exposed liver

5.6.7.3 Cessation vs. switch to THS2.2 effect

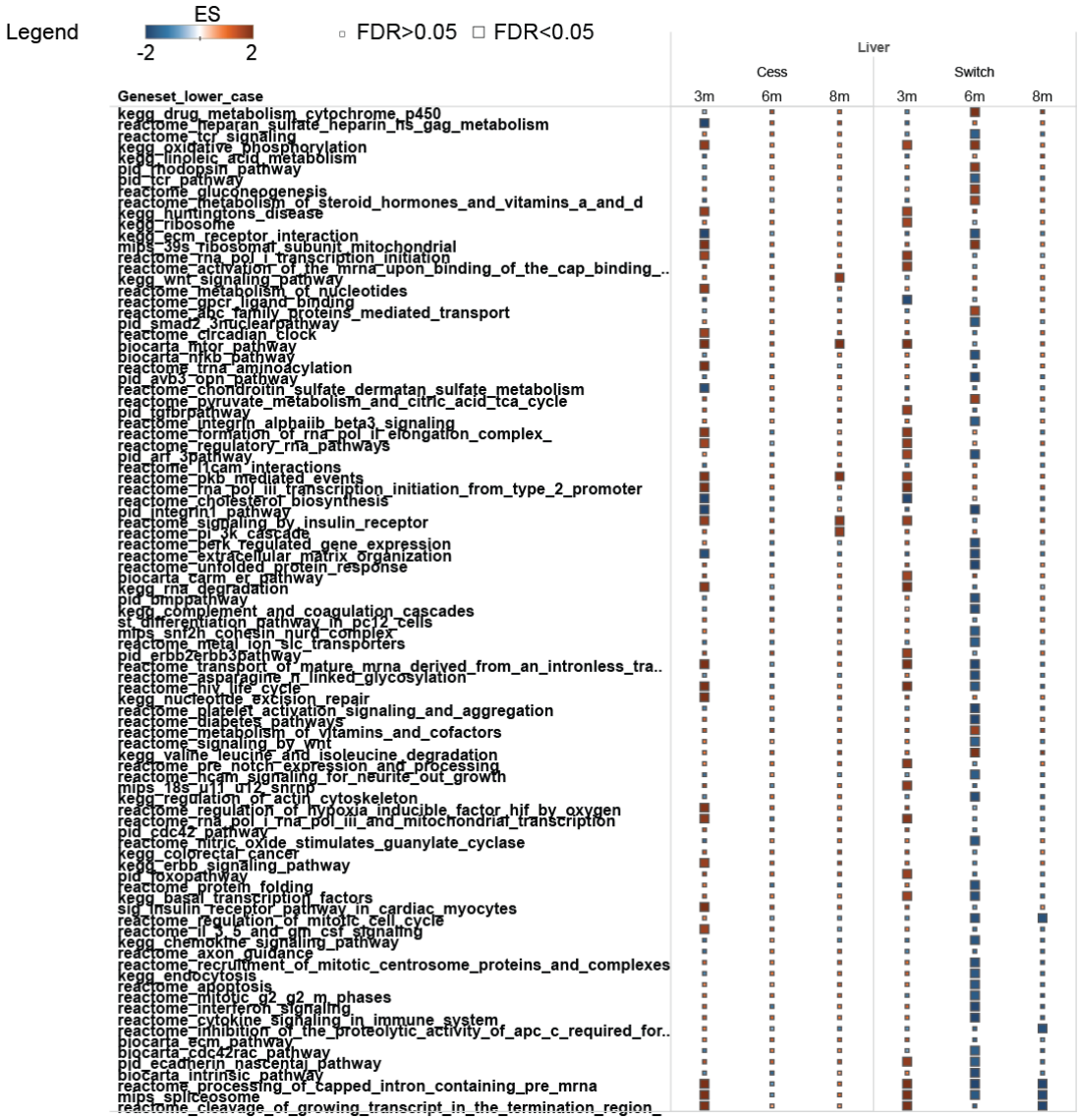


Figure 71: Gene set enrichment analysis in the livers of animals from the cessation and switch to THS2.2 groups

5.7 Liver proteomics

5.7.1 Number of replicates per group and protein identification

After quality control, the following samples were excluded from analysis:

- One full TMT set from month 8 (replicate 7): based on sample QC of both technical replicate mass spectrometry runs
- One switching sample from month 8 (replicate 8): based on PCA analysis results

The number of replicates effectively available for each group is summarized in Table 12.

Table 12: Number of valid replicates of liver proteomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	8	8	8	7
3R4F	8	8	8	8	7
THS2.2	8	8	8	8	7
Cessation	-	-	8	8	7
Switch	-	-	8	8	6

In total, 5213 proteins were identified in the experiment. However, for differential expression in comparison with the sham exposure group, only proteins identified in at least two thirds of one compared group were considered. The number of proteins considered in each sham comparison is given in Table 13.

Table 13: Number of consistently quantified liver proteins

Exposure group/time point (months)	1	2	3	6	8
3R4F	2,327	2,194	2,206	2,174	2,381
THS2.2	2,330	2,194	2,205	2,175	2,380
Cessation	-	-	2,206	2,175	2,381
Switch	-	-	2,206	2,174	2,496

5.7.2 Differentially expressed proteins in the liver

For all exposure conditions and time points, the amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of protein expression changes in the liver were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 72.

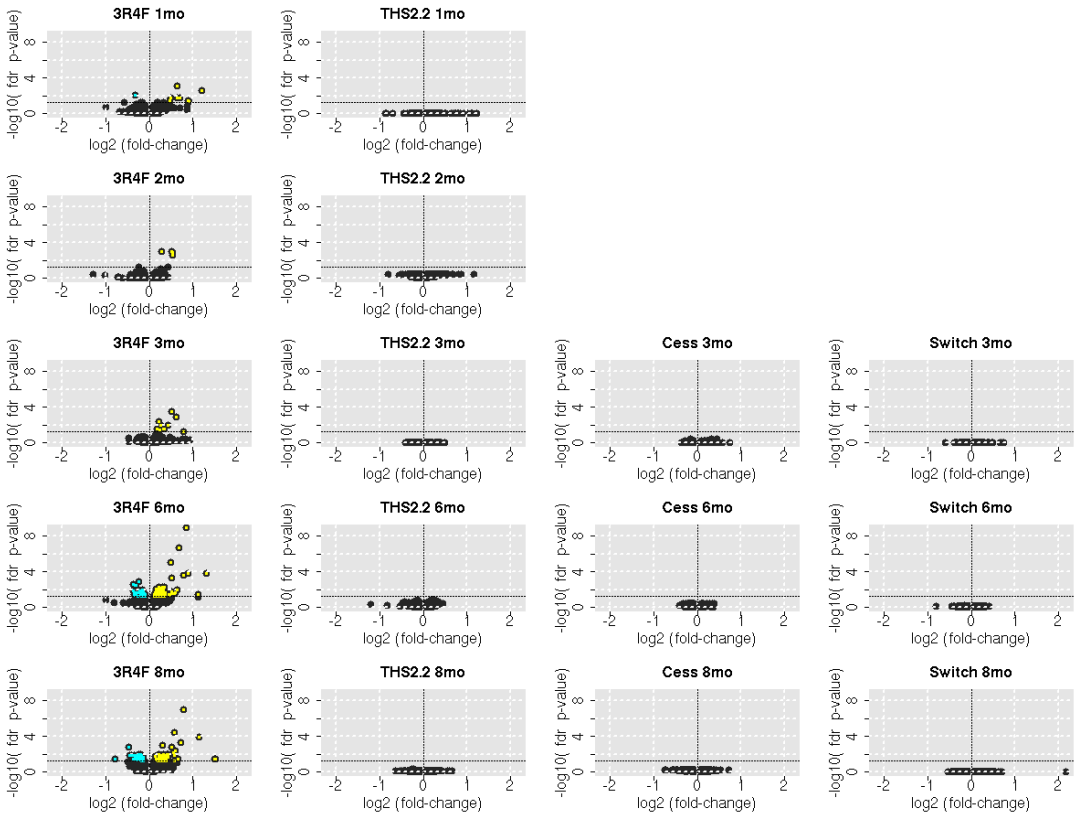


Figure 72: Volcano plots representing the proteome response profiles in the liver

The protein expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that were significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point (BH adjusted *P*-value <0.05).

The numbers of differentially expressed proteins are shown in Figure 73 and Figure 74. The FDR was controlled at the 5% level using the BH adjustment.

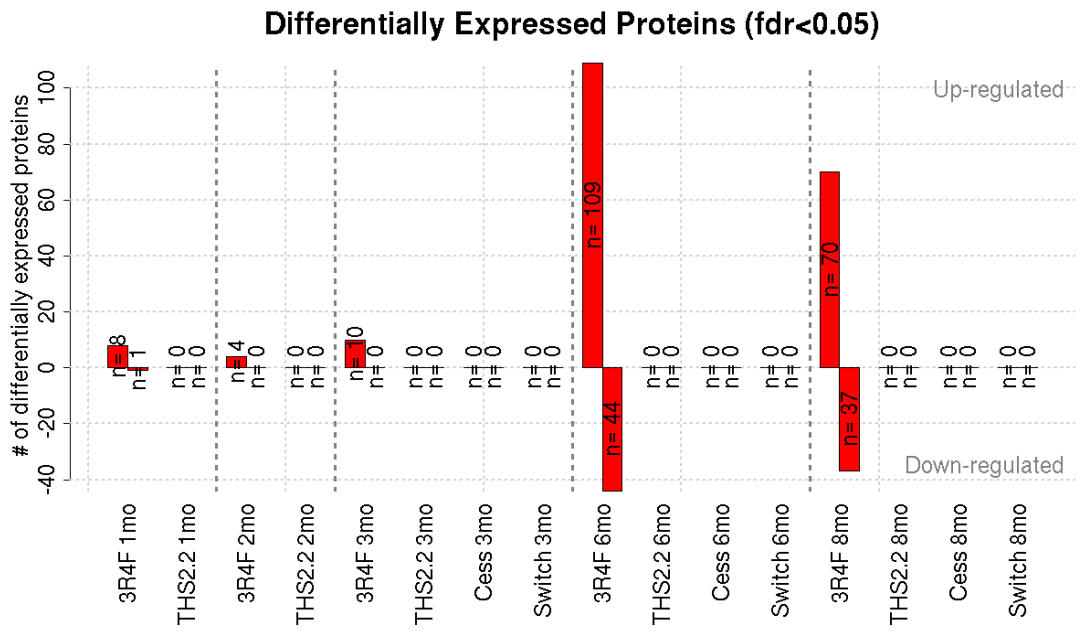


Figure 73: Numbers of differentially expressed proteins for each comparison in the liver
For each comparison with the sham group, the number of differentially up- and down-regulated proteins is shown. The FDR was controlled at the 5% level using the BH approach.

3R4F exposure resulted in only a few significant changes within the first 3 months (4–10 differentially regulated proteins). However, an increased 3R4F exposure effect was observed in months 6 and 8 with a maximum of 153 differentially regulated proteins observed at 6 months. Note that a similar, but less pronounced, time course trend was observed for the lung proteome. By contrast, no differentially regulated proteins were identified for the entire 8-month exposure period for THS2.2 exposure and the cessation and switching groups.

The overlap between the lists of differentially up- or down-regulated proteins was visualized to assess how differentially regulated proteins compared (Figure 75). Whereas only three of eight proteins up-regulated at 1 month were still up-regulated at later (6- and 8-month) time points, most proteins up-regulated after 2 and 3 months of exposure were still up-regulated at later time points. Interestingly, only 42% of the proteins up-regulated after 8 months of 3R4F exposure were up-regulated after 6 months of exposure.

Regulated proteins were also tracked over all conditions using a differential expression heatmap (Figure 76). Again, a time-dependent 3R4F effect was apparent with an increase in effect observed between 3 and 6 months.

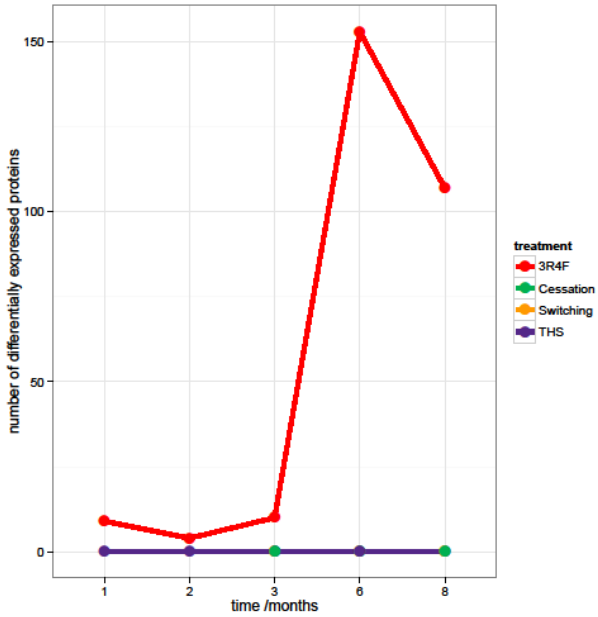


Figure 74: Numbers of differentially expressed proteins in the liver during the exposure time
Summary plot of the number of differentially expressed proteins for different treatment conditions and time points.

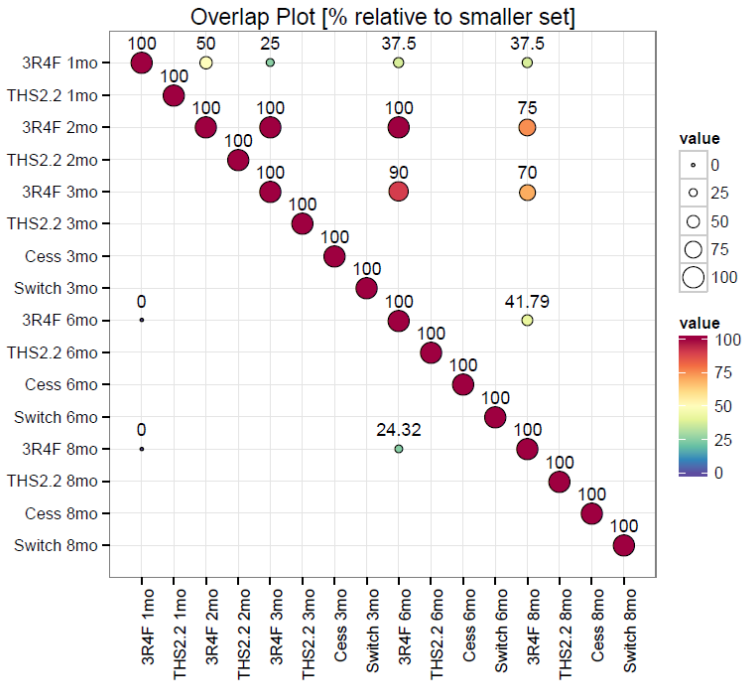


Figure 75: Overlap between sets of differentially expressed proteins in the liver

For each pair of comparisons, the number of shared up- and down-regulated proteins is visualized. The overlap for the up-regulated protein sets is shown in the top right corner, and for the down-regulated proteins in the bottom left corner of the matrix. The percent overlap with respect to the smaller set of differentially regulated proteins is visualized as the node color and size of each compared pair.

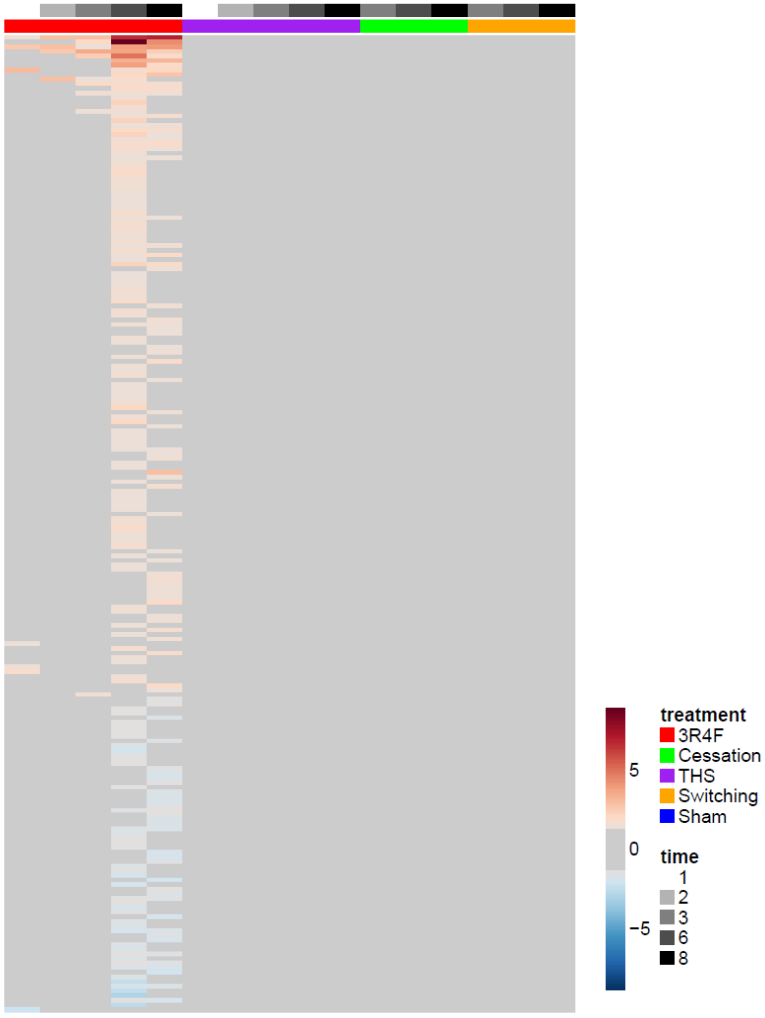


Figure 76: Heatmap of differentially expressed proteins in the liver

Differential protein expression of treatment versus sham at the same exposure duration. Each row represents a differentially expressed protein, each column a sample group comparison (see color key). Only significant changes (FDR<0.05) are colored as the signed –log10 (adjusted *P*-value), with red and blue indicating over- and under-expression, respectively, compared with sham.

5.7.3 Exposure effect comparison for proteins strongly affected by 3R4F exposure

The large number of proteins in the heatmap above made a detailed comparison of their response profiles challenging. Thus, we focused on the protein subset with the most robust 3R4F response. For this, we selected proteins that were significantly regulated by 3R4F exposure (adjusted *P*-value <0.05) with an up- or down-regulation fold-change >1.3, and compared their response across all conditions (Figure 77). We also compared the exposure effects for cytochromes P450 (CYPs) and glutathione-S-transferases (GSTs) using boxplots (Figure 78).

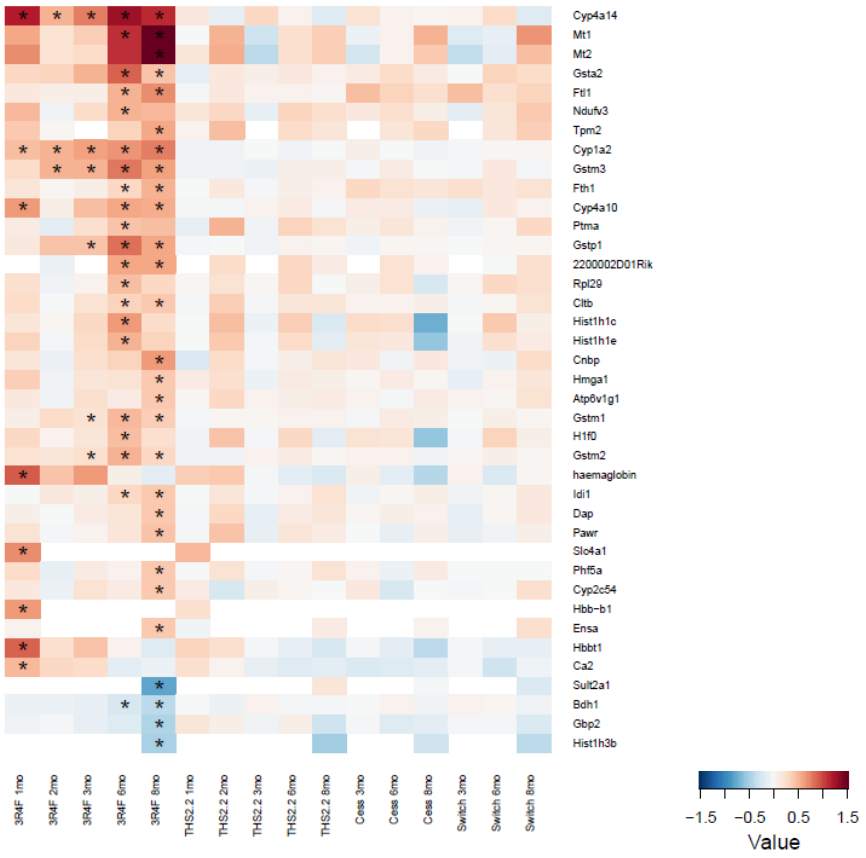


Figure 77: Fold-change expression matrix for proteins strongly affected by 3R4F exposure
Differential protein expression of treatment versus sham at the same exposure duration, visualized as a heatmap. The log2 fold-change is color-coded ("value"). Only proteins significantly regulated by 3R4F exposure with an absolute log2 fold-change > log2(1.3) are shown.

Figure 77 shows the limited overlap between 1 and 2–3 months of 3R4F exposure. Those proteins significantly up-regulated at the 1-month time point were mostly erythrocyte-derived such as hemoglobin, carbonic anhydrase II, and Slc4a1. This could represent variations in whole-body perfusion.

Only a few proteins were differentially up-regulated after 2 and 3 months of 3R4F exposure: Cyp4a14, Cyp1a1, Gstm3, Gstp1, Gstm1, and Gstm2. These proteins also remained up-regulated at later 3R4F exposure time points (6 and 8 months) and all are part of the xenobiotic metabolism. Thus, the main early and persistent responses of the liver proteome appear to comprise xenobiotic metabolism proteins. Figure 78 shows a subset of this response in more detail.

The functional basis for the increased 3R4F response at months 6 and 8 is less clear. Proteins uniquely up-regulated at these time points include ferritin light chain, ferritin heavy chain, metallothionein 1/2, and clathrin light chain B.

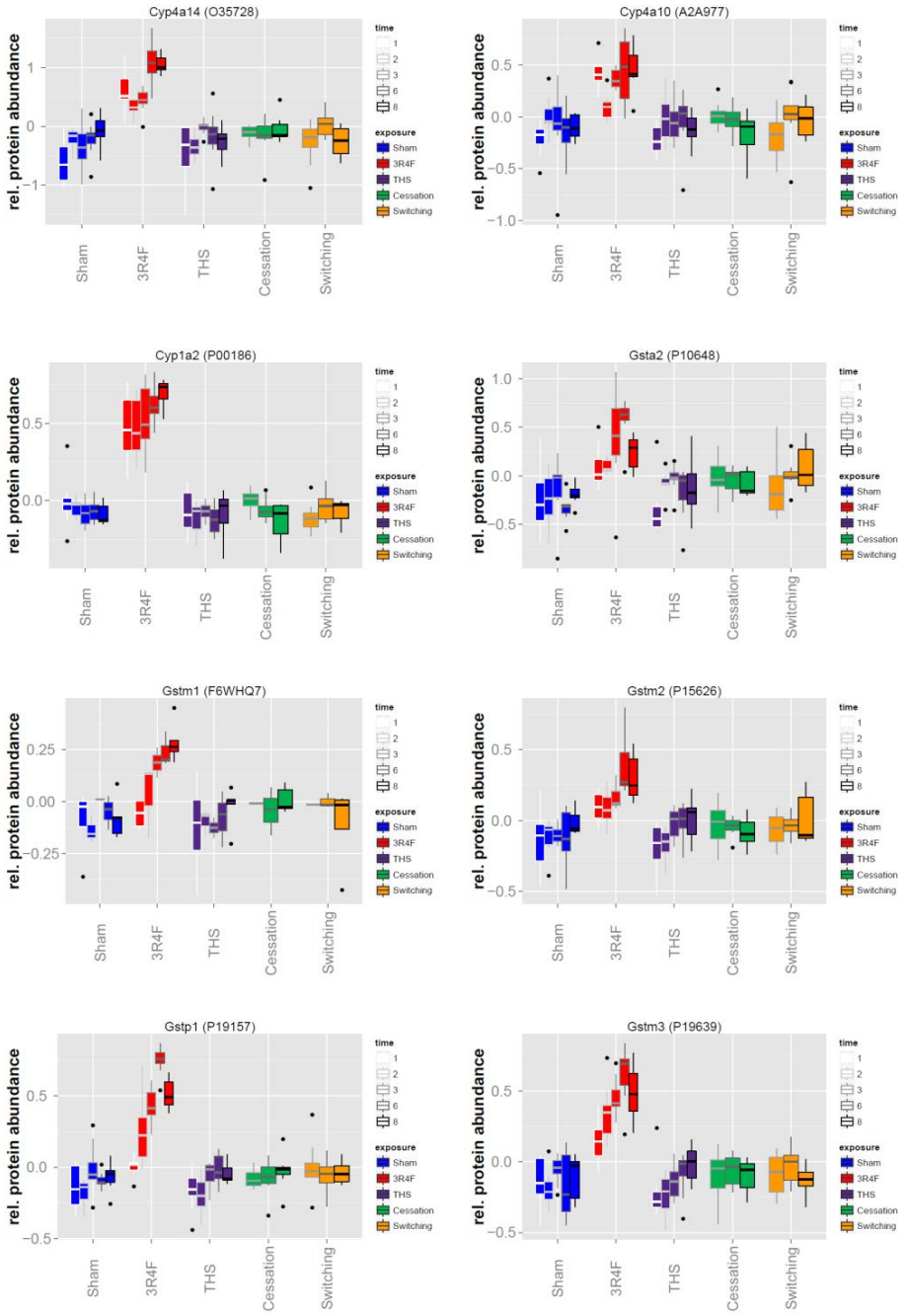


Figure 78: Boxplots of the protein expression response of selected xenobiotic enzymes from Figure 77

5.7.4 Analysis of activated functional protein networks

To better understand and compare the functional impact of 3R4F exposure on the liver proteome, we conducted an activated functional subnetwork analysis. The regulated proteins

in the condition with the largest overall effect (3R4F, 6 months) served as the query set. Functional interactions for these proteins were extracted from the STRING database and functional clusters in the derived network were identified and annotated (Figure 79). Using this network as a template, the regulation of the selected proteins and clusters was compared across all exposure conditions (Figure 80 Figure 81).

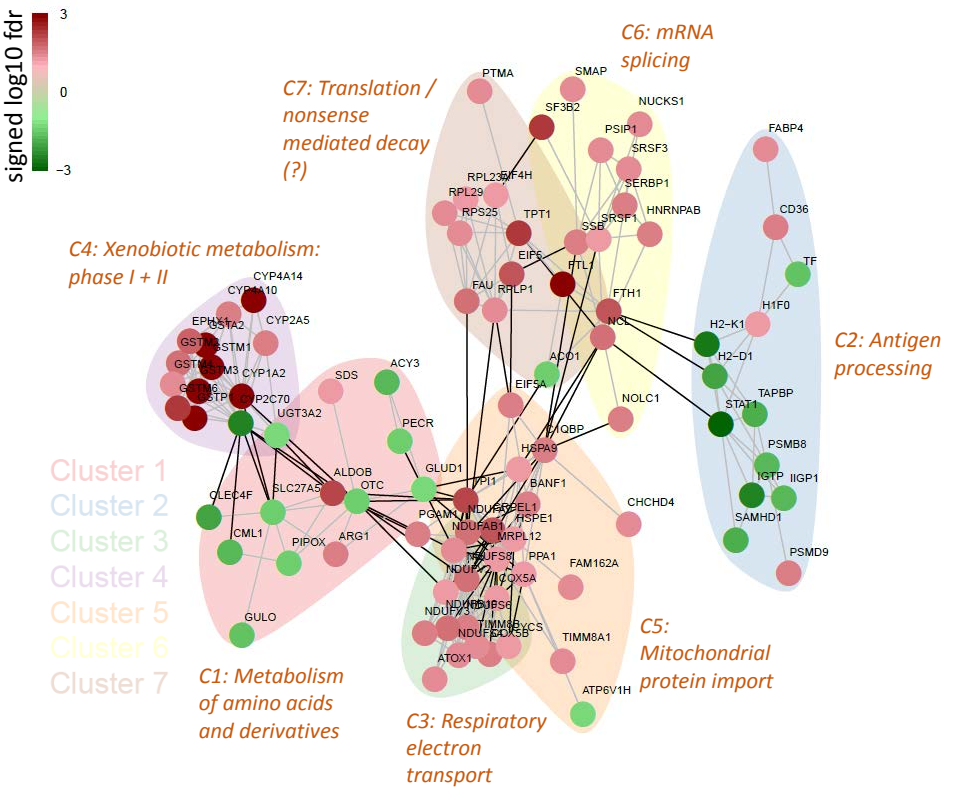


Figure 79: Functional interaction network for proteins significantly regulated in the liver after 6 months of 3R4F exposure

The nodes represent the selected differentially regulated proteins in the 3R4F, 6-month exposure group. The node color reflects the signed log10 adjusted *P*-value of differential expression (see color key). The edges/links represent functional interactions from the STRING database. Clusters in this network were identified and marked by colors (see key). A summary annotation is given next to each cluster. The subnetwork induction method was used to define this template network.

Seven functional clusters for the proteins regulated by 6 months of 3R4F exposure were identified (Figure 79), including “xenobiotic metabolism”, “respiratory electron transport”, “mRNA splicing”, and “translation” clusters, mostly with up-regulated proteins, and “metabolism of amino acids” and “antigen processing” clusters with both up- and down-regulated proteins.

Mapping of the protein alterations onto this functional subnetwork for all exposure conditions underlined the previously observed difference between 3R4F and the other exposure conditions (Figure 80 Figure 81). The "xenobiotic metabolism" cluster showed the most rapid response to 3R4F exposure, with protein induction already apparent after 1 month of exposure (e.g., Cyp4a10, Cyp4a14). Overall, the proteins in other clusters demonstrated increased regulation after 6 and 8 months of 3R4F exposure, but the response profiles were less closely tied to exposure type or time. Therefore, although 3R4F had the greatest effect on xenobiotic metabolism enzymes, further investigations are needed to understand the functional relevance of the other clusters.

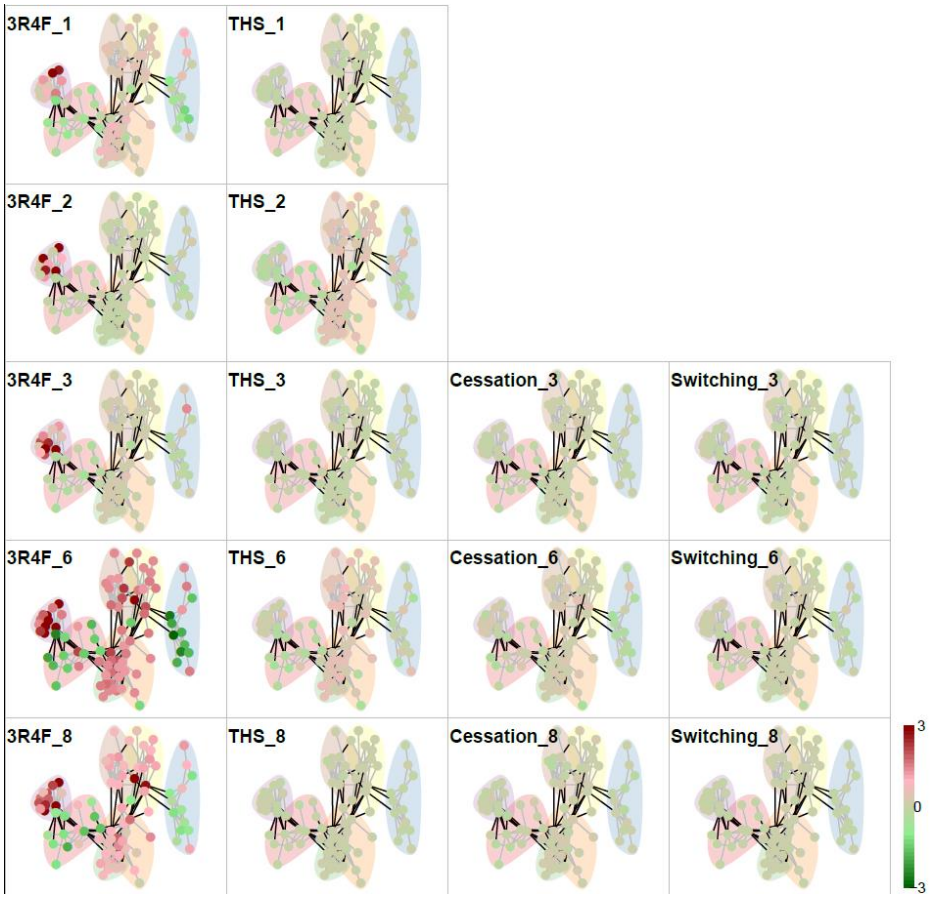


Figure 80: Comparison of exposure effect on the identified functional network across all exposure conditions (signed log10 adjusted *P*-values)

The functional subnetwork from Figure 79 was used as the template and the signed log10 adjusted *P*-value is represented as the protein/node color for each exposure versus sham comparison. A strong and rapid induction of the xenobiotic cluster (C4) following 3R4F exposure is evident in the top left corner.

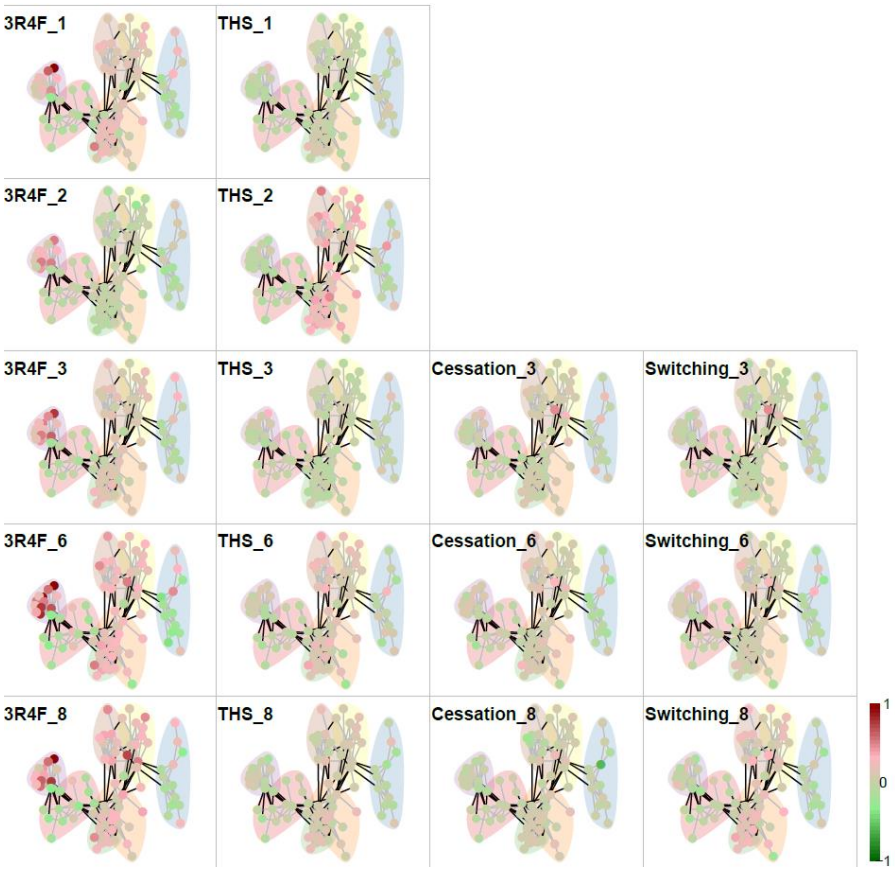


Figure 81: Comparison of exposure effect on the identified functional network across all exposure conditions (log2 fold-changes)

The functional subnetwork from Figure 79 was used as the template and the log2 fold-change value is represented as the protein/node color for each exposure versus sham comparison.

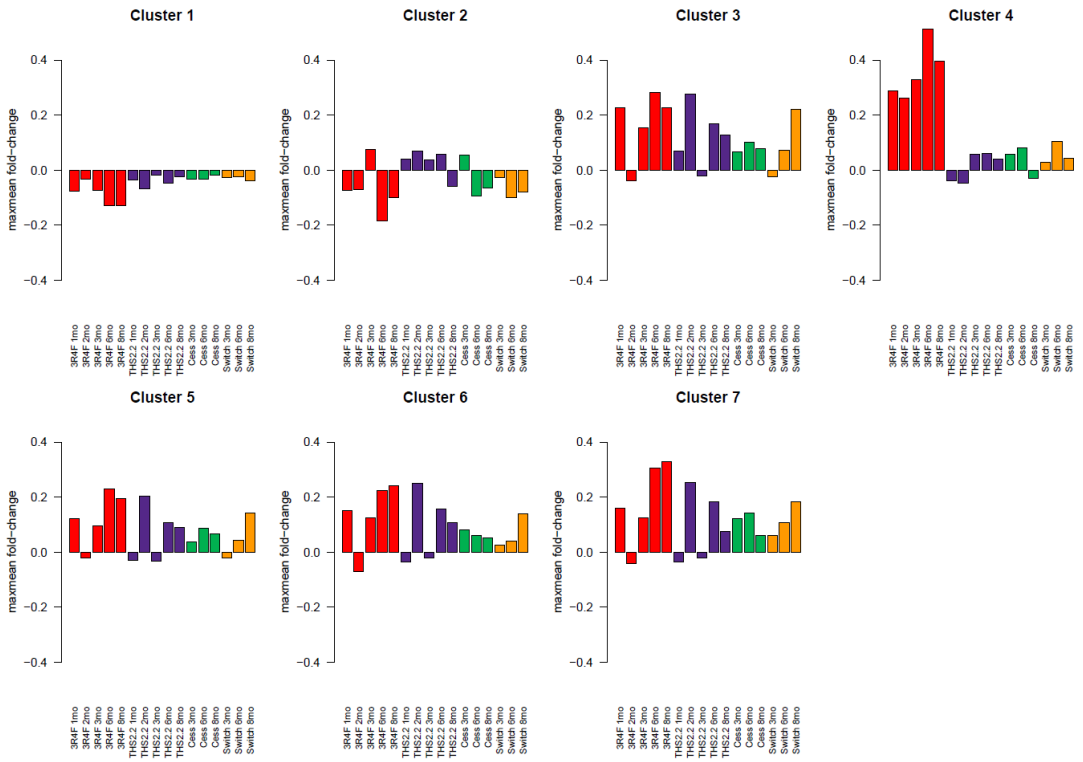


Figure 82: Summary of exposure effects on the identified protein clusters
Maxmean statistics of the fold-changes for each cluster and exposure group.

5.7.5 Gene/protein set analysis

We conducted a gene/protein set analysis to complement the functional insights obtained from the subnetwork analysis. We identified protein sets from a large compendium (MSigDB, canonical pathways, (40)) that demonstrated significant average up- or down-regulation of its protein members.

Overall, the pattern of protein set regulation was more diverse in the liver than in the lung proteome (Figure 46). For example, regulation of translation-related protein sets (e.g., KEGG RIBOSOME) was observed for THS2.2 (8 months) and cessation (8 months) groups, and regulation of oxidative phosphorylation-related protein sets (e.g., KEGG OXIDATIVE PHOSPHORYLATION) was observed for THS2.2 (6 months) and cessation (8 months) groups. While these effects might be associated with later time points, they did not demonstrate an overall consistent exposure- or time-dependency.

Similar to the functional subnetwork analysis (Figure 80 and Figure 82), the dominant 3R4F exposure effect was identified for xenobiotic metabolism-related protein sets (e.g., KEGG METABOLISM OF XENOBIOTICS BY CYTOCHROME P450). For these protein sets, we observe a strong induction by 3R4F exposure over time, a limited effect by THS2.2 exposure, and a rapid loss of the 3R4F exposure effect upon switching or cessation.

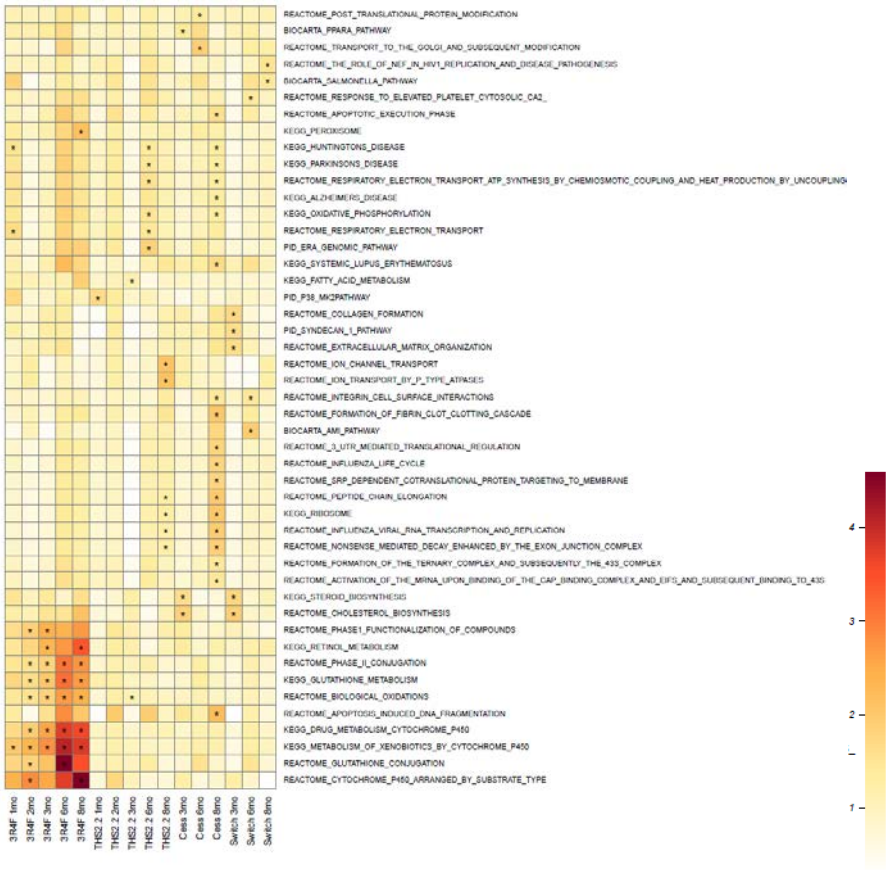


Figure 83: Protein set analysis for the liver

Protein set analysis results based on (absolute) t-scores and the mean as the protein set statistic. Only protein sets that were significantly regulated (BH-adjusted *P*-value <0.05, *) in at least one condition are shown. The color-scale reflects the protein set statistic.

5.7.6 Summary of the observed effects on the liver proteome

The exposure effects on the liver proteome were assessed with a quantitative, mass spectrometry-based approach. A good coverage of the liver proteome was achieved, with a total of 5,213 proteins identified and ~2,000–2,300 consistently quantified proteins for each exposure comparison.

3R4F exposure resulted in only few significantly regulated proteins for months 1–3 (mainly of xenobiotic enzymes), but showed an increased response for months 6 and 8. No significant protein regulation was seen following THS2.2 exposure or in the cessation and switching groups (Figure 72 and Figure 73).

The functional impact of 3R4F on the lung proteome was guided by a functional subnetwork approach. This approach highlighted a “xenobiotic metabolism” cluster as the main component of an early and maintained response to 3R4F exposure (Figure 80, Figure 82). By contrast, THS2.2 showed only a limited to absent effect on “xenobiotic metabolism” proteins, and the

cessation and switching groups demonstrated a rapid loss of the exposure effects (Figure 80, and Figure 82). Proteins in other identified functional clusters were significantly regulated at months 6 and 8 only; these included a "respiratory electron transport", "mRNA splicing", and "translation" cluster. Further investigations will be needed to understand the functional relevance of these protein clusters. Protein set analysis further supported the dominant effect of 3R4F on xenobiotic metabolism proteins in the liver (Figure 83).

In summary, we identified regulation of "xenobiotic metabolism" proteins as the dominant and consistent main effect of cigarette smoke exposure in the liver. THS2.2 exposure resulted only in limited to absent effects on this protein group, and the exposure effects quickly returned to sham group levels upon cessation and switching. Other effects of cigarette smoke exposure were detectable, especially at later time points, but these showed less consistency and require further investigations to fully understand.

5.8 Heart transcriptomics

5.8.1 QC and number of replicates per group

All samples passed QC, so eight replicates were available for each group, as summarized in Table 14.

Table 14: Number of valid replicates of heart transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	8	8	8	8	8
3R4F	8	8	8	8	8
THS2.2	8	8	8	8	8
Cessation	-	-	8	8	8
Switch	-	-	8	8	8

5.8.2 Differentially expressed genes in the heart

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of gene expression changes in the heart for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and asre visualized as volcano plots in Figure 84. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 85.

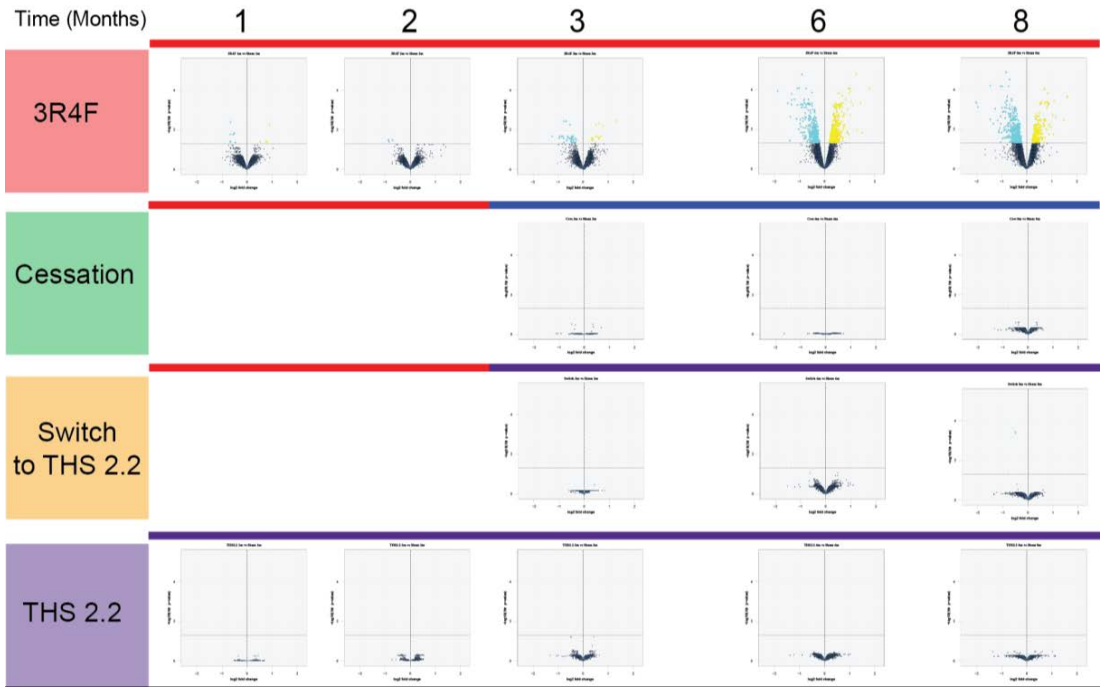


Figure 84: Volcano plots representing the systems response profiles in the heart: FDR

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that were significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

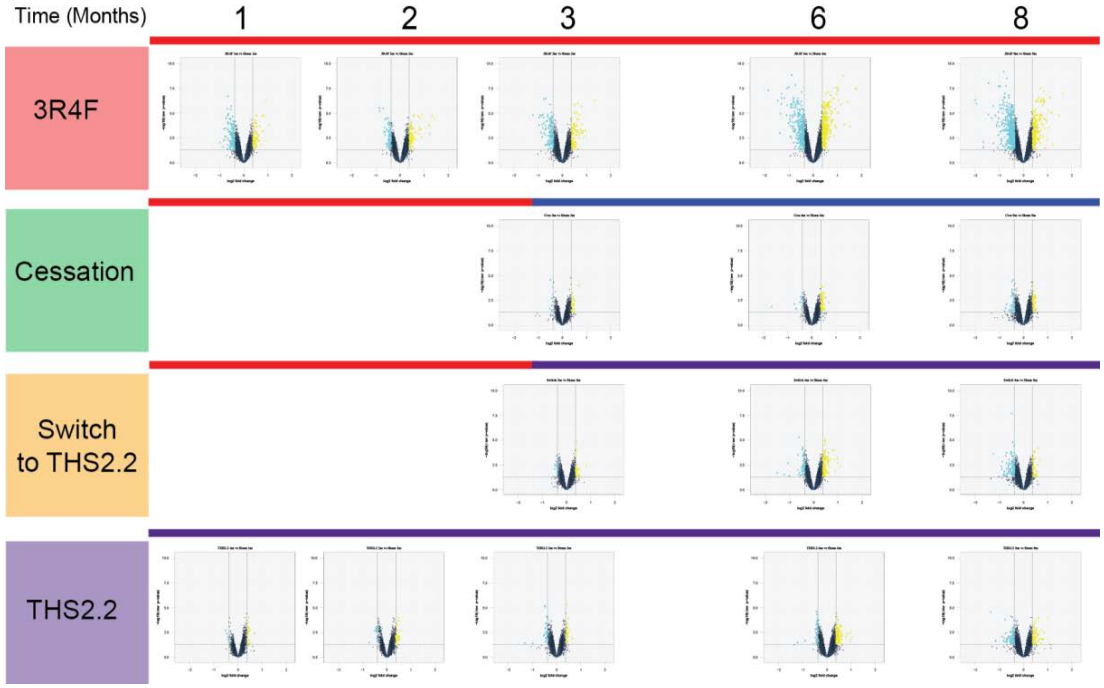


Figure 85: Volcano plots representing the systems response profiles in the heart: raw *P*-value
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold-change >1.2.

The number of differentially expressed genes compared with sham for each exposure group is shown in Table 15. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, a pairwise comparison of the overlap was conducted as shown in Figure 86.

Table 15: Number of differentially expressed genes in the heart (up-/down-regulated compared with sham at each respective time point)

Exposure group/time point (months)	1	2	3	6	8
Sham	REF	REF	REF	REF	REF
3R4F	4/7	0/3	16/37	396/380	264/489
THS2.2	0	0	0	0	0
Cessation	-	-	0	0	0
Switch	-	-	0	0	0/1

REF: reference

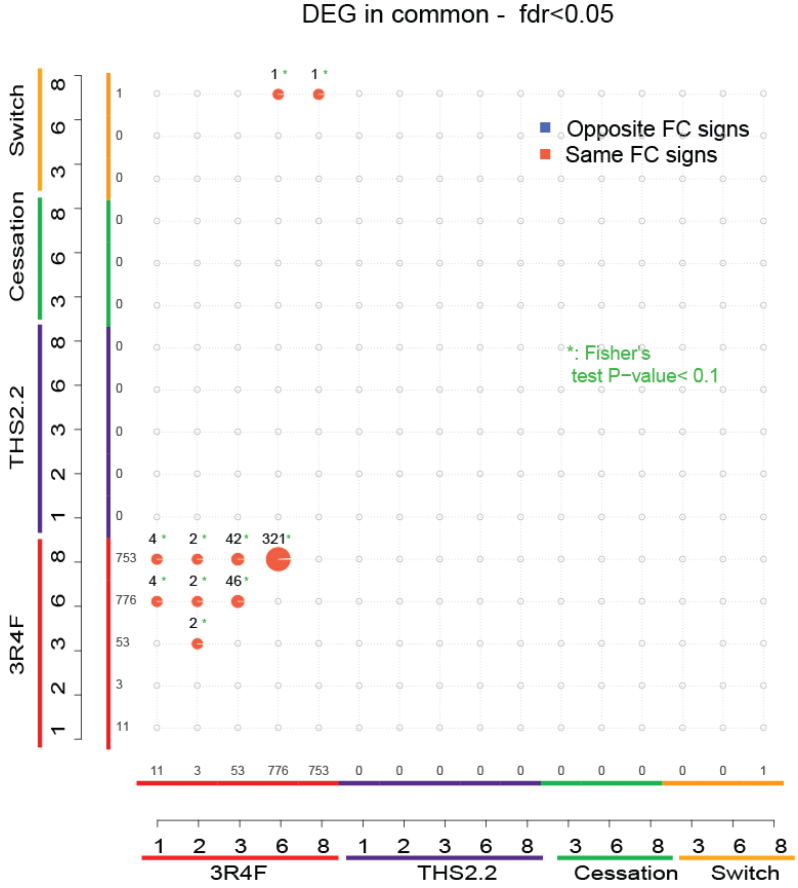


Figure 86: Overlap between sets of differentially expressed genes in the heart

For each pair of comparisons, the number of shared up- and down-regulated genes is visualized. The overlap of genes significantly regulated in the same direction (either both up- or both down-regulated) is indicated in red, while opposite direction changes are shown in blue. A green star signifies that the overlap is significant (Fisher's test *P*<0.1). The number of significantly regulated genes (FDR<0.05) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.8.3 BIF analysis in the heart

The overall impact of the different exposures was estimated from a BIF (32), expressed as an RBIF (Figure 87). It is important to note that not all networks were built for a cardiovascular context. Their primary application was meant for pulmonary tissues, so the results need to be interpreted with caution.

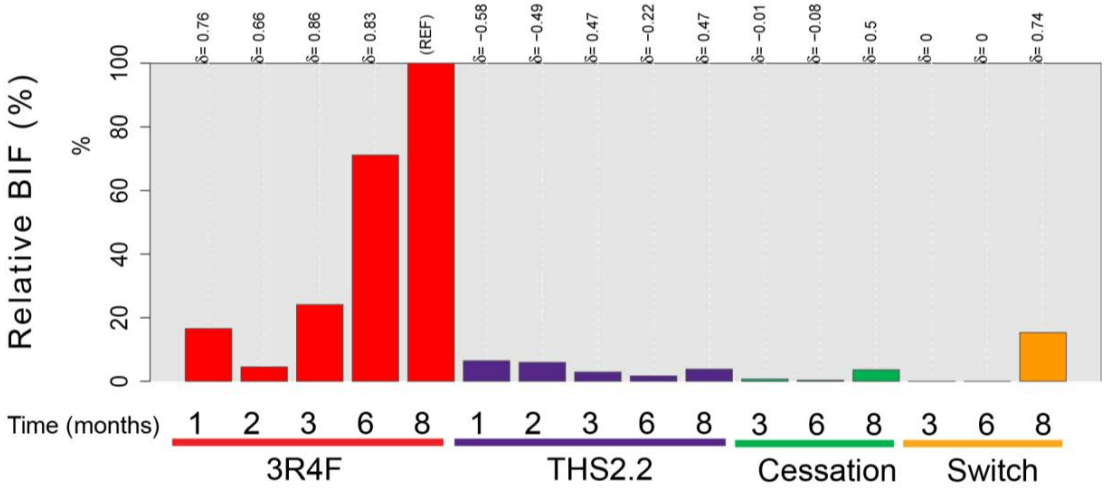


Figure 87: BIF analysis in the heart

The percentages show the relative biological impact, derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 8 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 8 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.8.4 Biological network perturbation amplitude analysis in the heart

The star plots in Figure 88 show the biological processes that most contributed to the overall BIF phenotype at each time point.

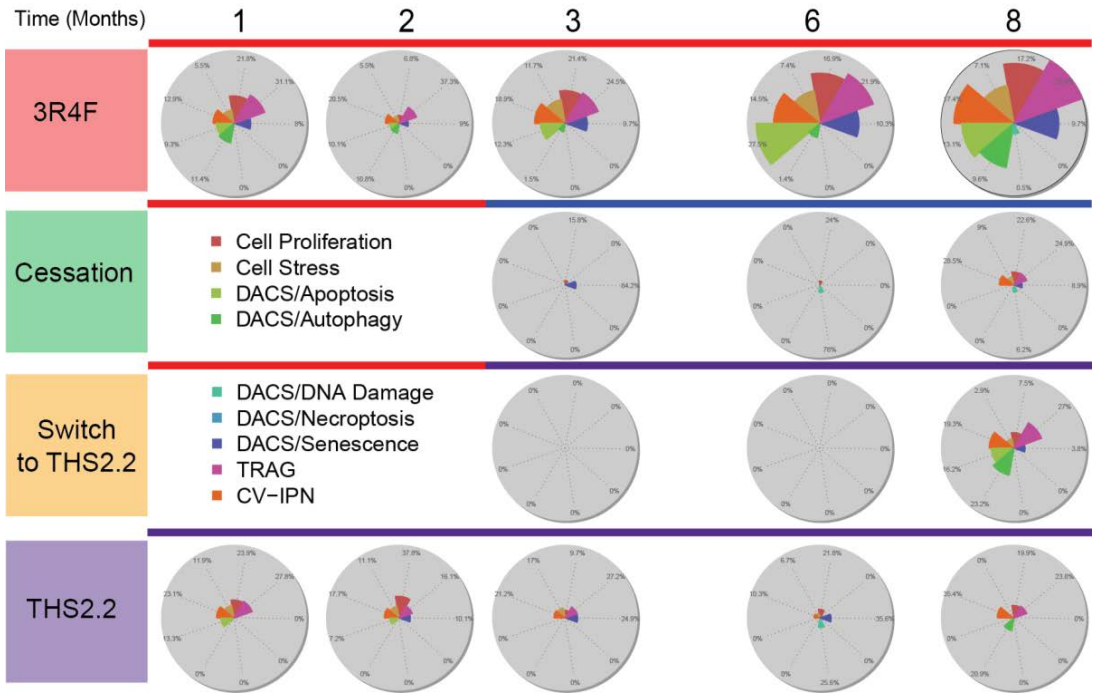


Figure 88: Star plots showing BIF calculated from the heart

Star plots illustrating the decomposition of the overall RBIF (Figure 87) into its eight mechanistic components (from cell proliferation to inflammation, indicated by the colors of the segments) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

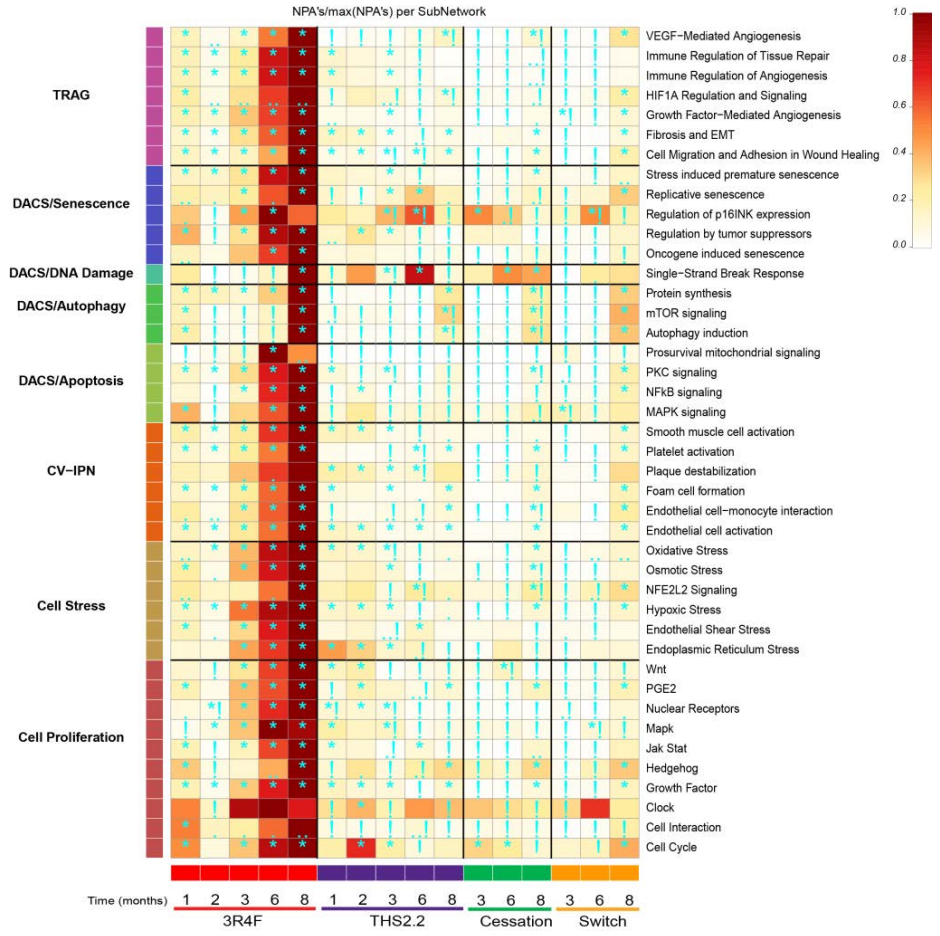


Figure 89: Heatmap of network perturbation amplitude scores in the heart

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network is considered perturbed if, in addition to the significance of the NPA score with respect to the experimental variation, the two companion statistics (O and K) derived to inform on NPA score specificity with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values <0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to the experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrasts comparisons, similar to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 89Figure 90).

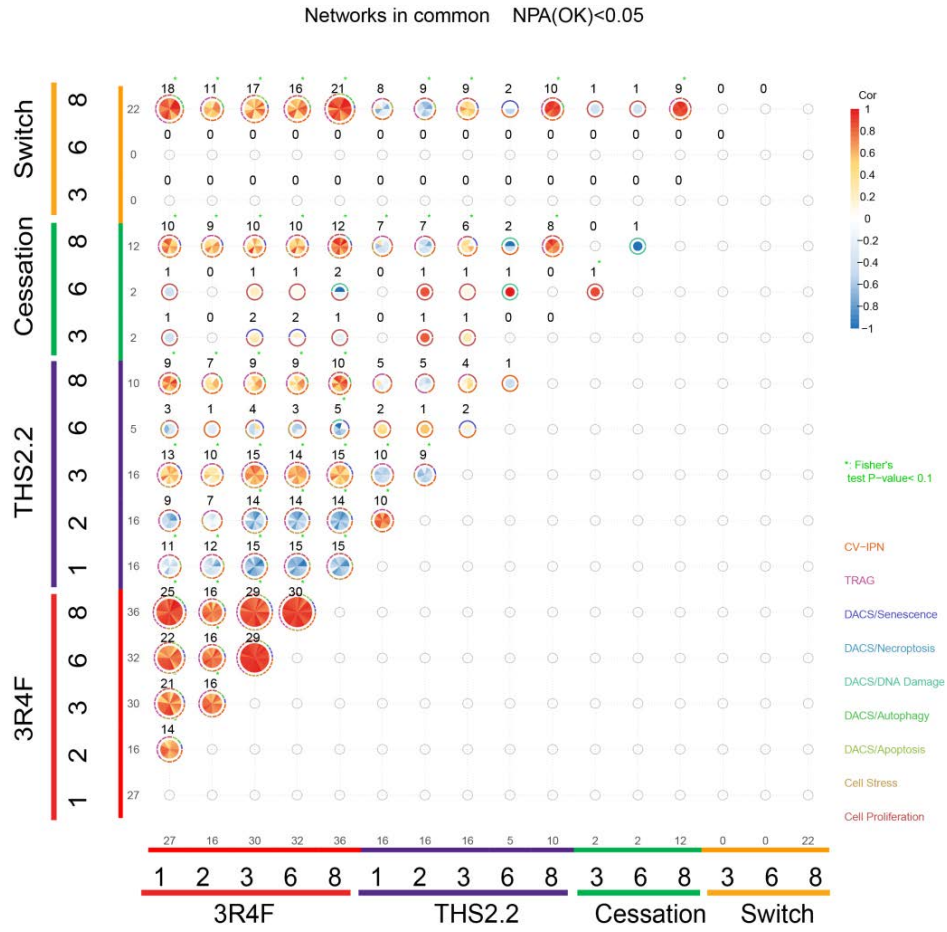


Figure 90: Overlap between sets of significantly perturbed networks in the heart

For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates positive correlation, while blue indicates negative correlation. A green star signifies that the overlap was significant (Fisher's test *P*<0.1). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.8.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by the different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 91).

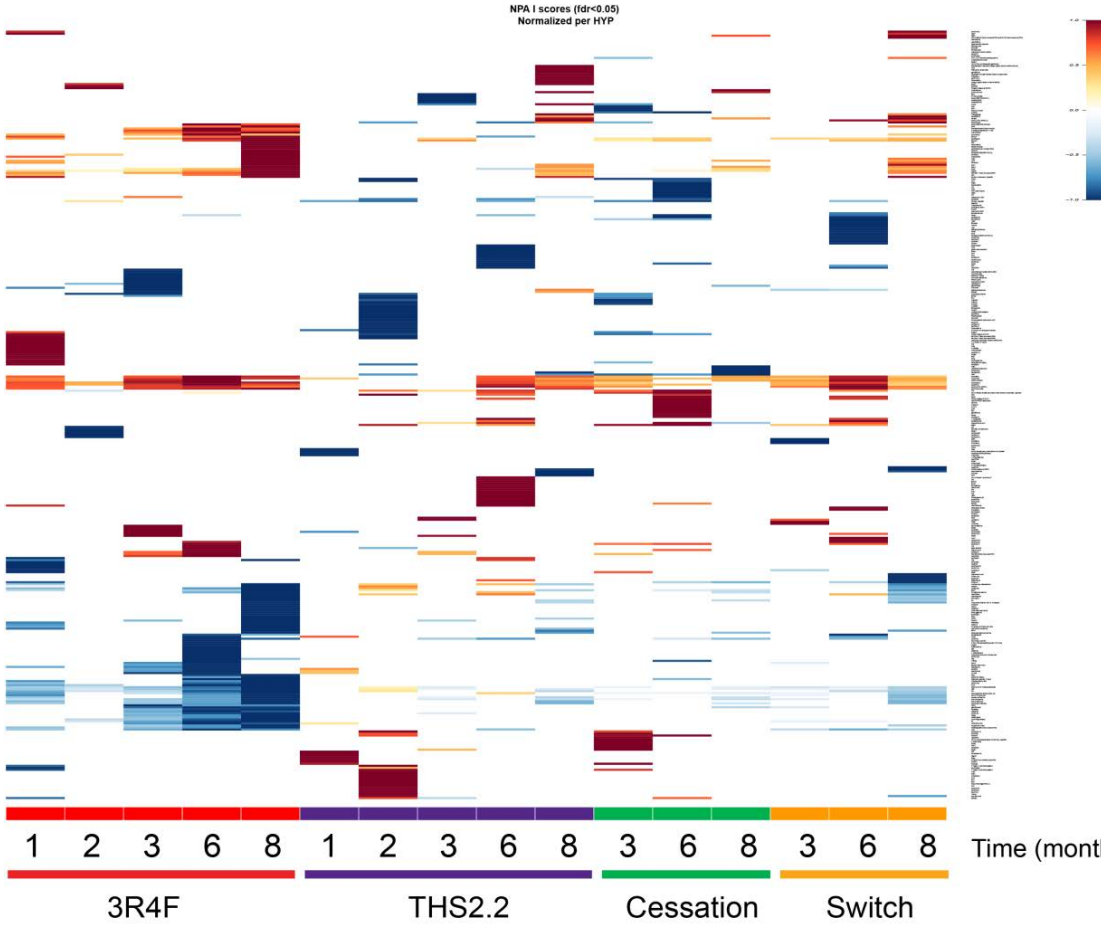


Figure 91: Heatmap of NPAI score for all hyps significant in at least one contrast in the heart

5.8.6 NextBio analysis

5.8.6.1 3R4F exposure effect

To study the effects of 3R4F on the heart, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed hearts at five different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly linked to xenobiotic metabolism and lipid transport (Figure 92).



Figure 92: NextBio meta-analysis of 3R4F effects in the heart

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to 3R4F (compared with sham-exposed control).

5.8.6.2 THS2.2 exposure effects

To study the effects of THS2.2 on the heart, a meta-analysis was performed comparing THS2.2-exposed and sham-exposed hearts at five different time points.

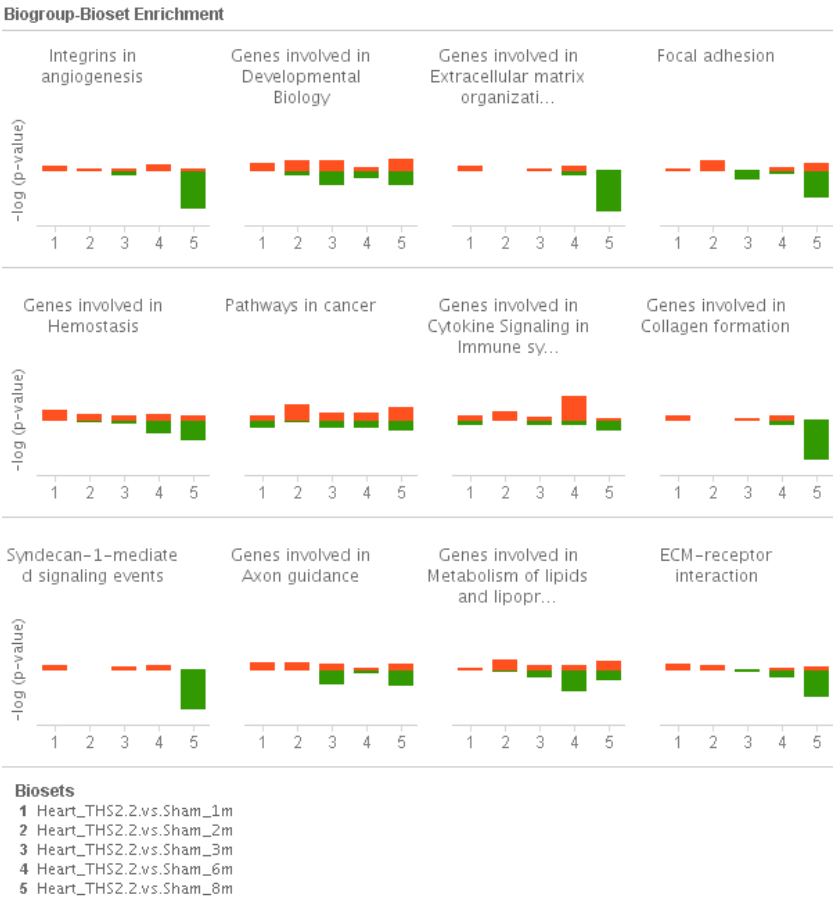


Figure 93: NextBio meta-analysis of THS2.2 effects in the heart

Each of the 12 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to THS2.2 (compared with sham-exposed controls).

5.8.7 Gene set enrichment analysis

5.8.7.1 3R4F effects

To further study the trends within each treatment group over time, gene set enrichment analysis was carried out.

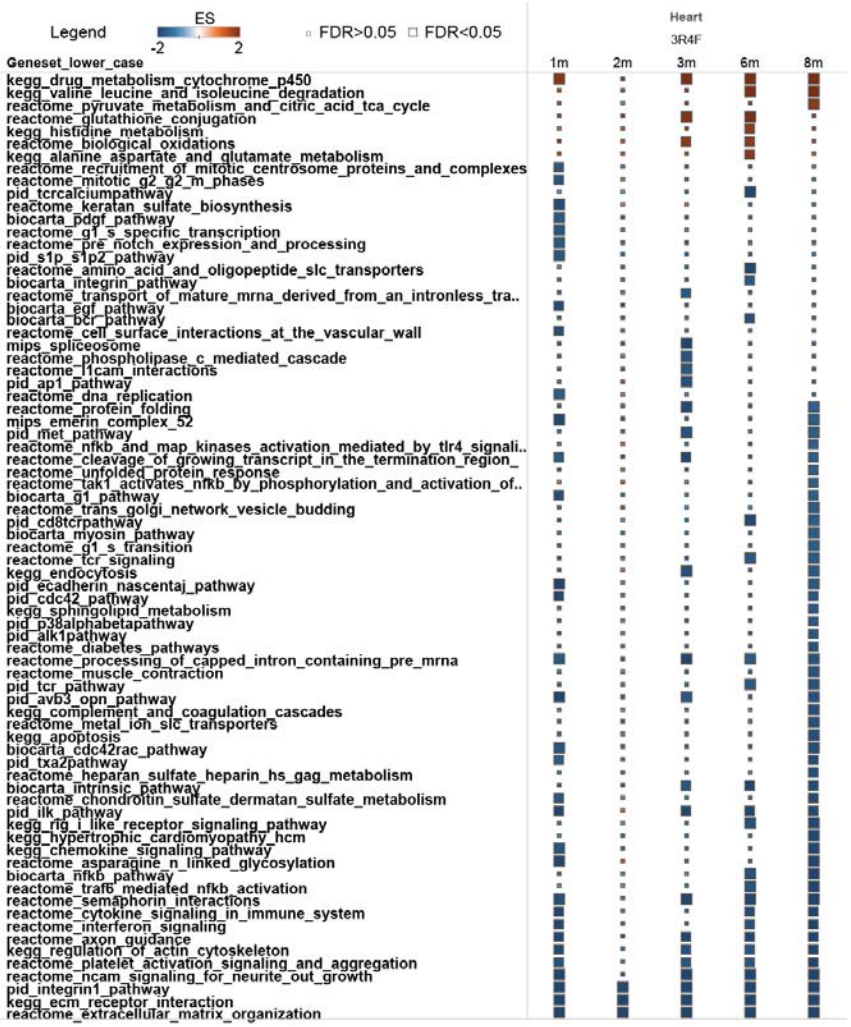


Figure 94: Gene set enrichment analysis in the 3R4F-exposed heart

5.8.7.2 THS2.2 effects

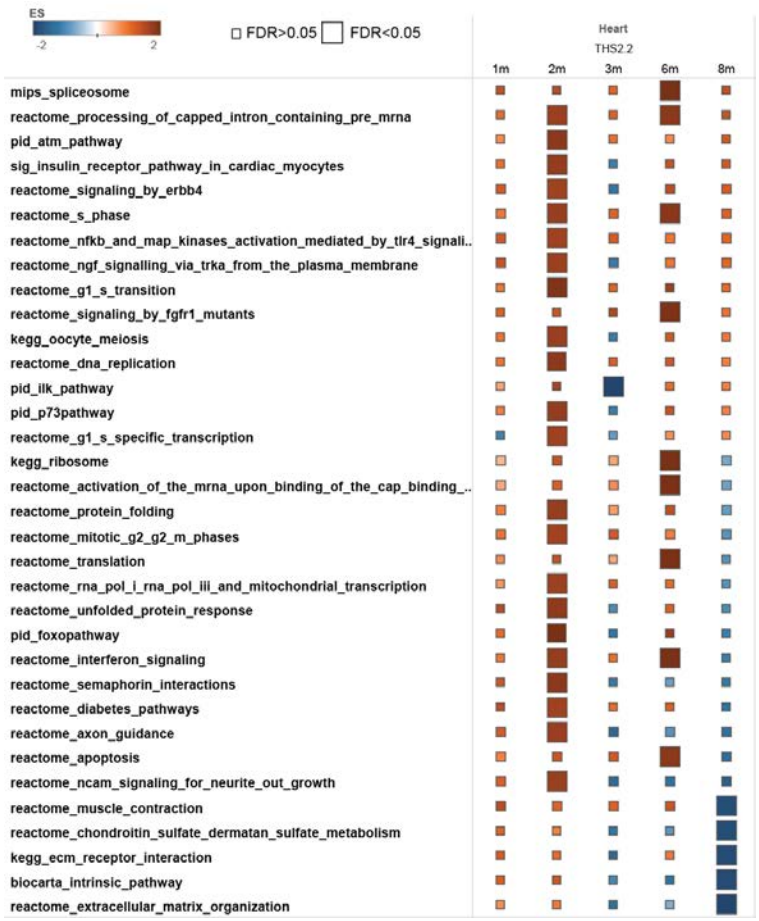


Figure 95: Gene set enrichment analysis in the THS2.2-exposed heart

5.8.7.3 Cessation vs. switch to THS2.2 effect

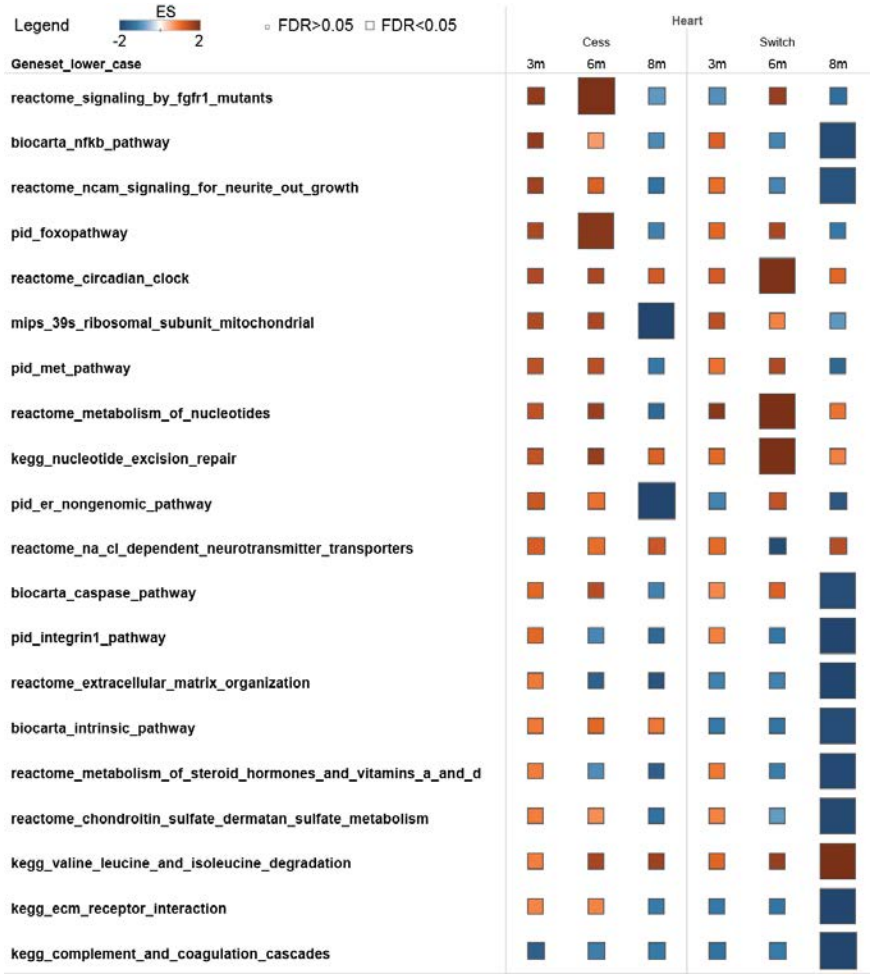


Figure 96: Gene set enrichment analysis in the hearts of animals from the cessation and switch to THS2.2 groups

5.9 Kidney transcriptomics

5.9.1 QC and number of replicates per group

The number of replicates effectively available for each group is summarized in Table 16.

Table 16: Number of valid replicates of kidney transcriptomics profiles

Exposure group/time point (months)	1	2	3	6	8
Sham	-	-	-	7	8
3R4F	-	-	-	8	8
THS2.2	-	-	-	8	8
Cessation	-	-	-	6	8
Switch	-	-	-	8	7

5.9.2 Differentially expressed genes in the kidney

The amplitude (expressed as log-fold change, x-axis) and significance (expressed as –log of the adjusted *P*-value) of gene expression changes in the kidney for 3R4F- and THS2.2-exposed mice were compared with the sham group at each respective time point and are visualized as volcano plots in Figure 97. To visualize more subtle effects, similar volcano plots are shown with the raw *P*-value plotted on the y-axis in Figure 98.

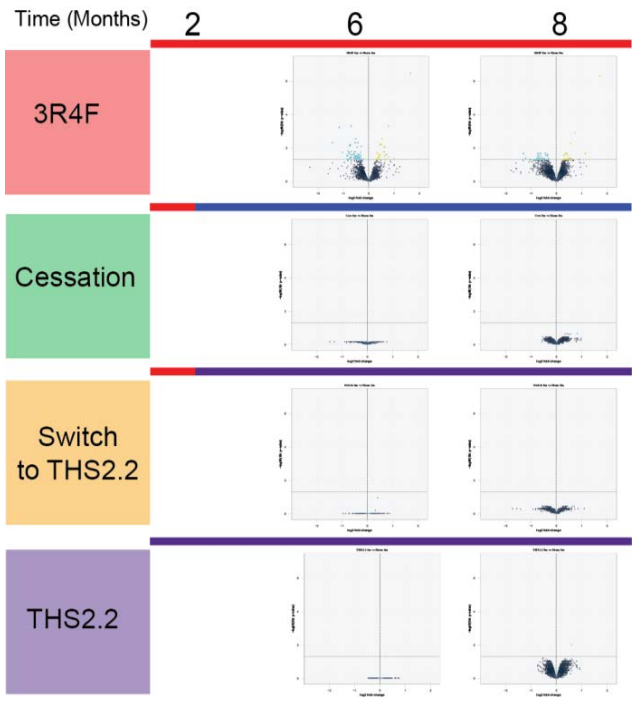


Figure 97: Volcano plots representing the systems response profiles in the kidney: FDR

The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-adjusted *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes that were significantly up- or down-regulated, respectively, compared with those in the sham group at each respective time point.

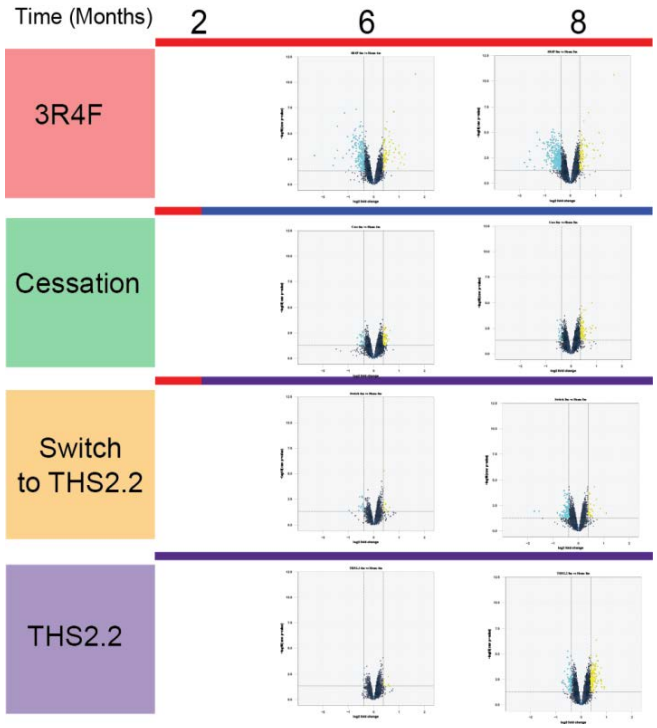


Figure 98: Volcano plots representing the systems response profiles in the kidney: raw *P*-value
The gene expression change, calculated as the log2 fold-change, is plotted on the x-axis and the significance, proportional to the negative log10-raw *P*-value, is plotted on the y-axis. Yellow and blue dots highlight genes for which the raw *P*-value is <0.05, and the absolute fold-change >1.2.

The number of differentially expressed genes compared with sham for each exposure group is shown in Table 17. To obtain a more precise estimation of the overlap of differentially expressed genes between groups, a pairwise comparison of overlap was conducted as shown in Figure 99.

Table 17: Number of differentially expressed genes in the kidney (up-/down-regulated compared with sham at each respective time point)

Exposure group/time point (months)	1	2	3	6	8
Sham	-	-	-	REF	REF
3R4F	-	-	-	24/63	21/32
THS2.2	-	-	-	0	1/0
Cessation	-	-	-	0	0
Switch	-	-	-	0	0

REF: reference

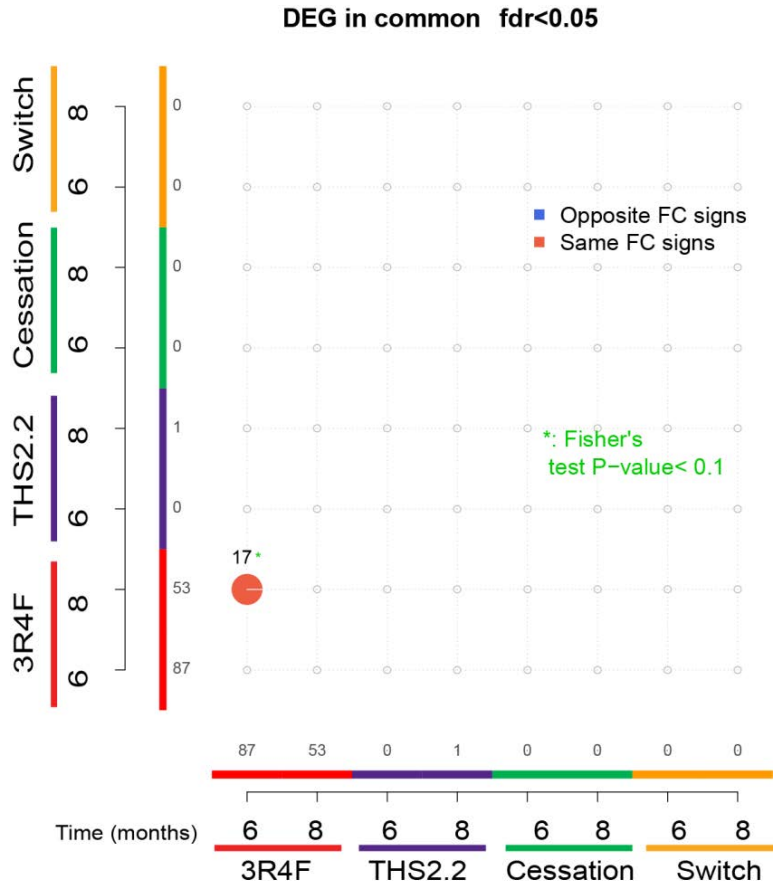


Figure 99: Overlap between sets of differentially expressed genes in the kidney
For each pair of comparisons, the number of shared up- and down-regulated genes is visualized. The overlap of genes significantly regulated in the same direction (either both up- or both down-regulated) is indicated in red, while opposite direction changes are shown in blue. A green star signifies that the overlap is significant (Fisher's test $P<0.1$). The number of significantly regulated genes (FDR<0.05) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.9.3 BIF analysis in the kidney

The overall impact of the different exposures was estimated from a BIF (32), expressed as an RBIF (Figure 100). It is important to note that not all networks were built for a kidney context. Their primary application was meant for pulmonary tissues, so the results need to be interpreted with caution.

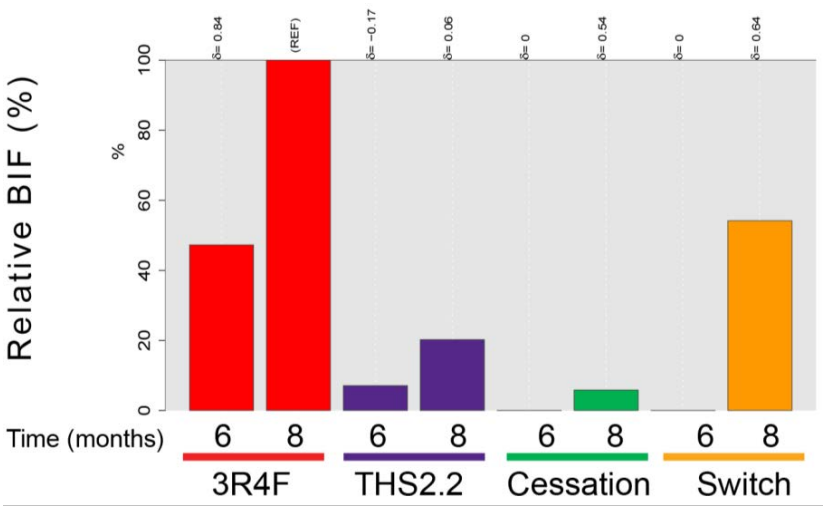


Figure 100: BIF analysis in the kidney

The percentages show the relative biological impact, derived from the cumulated network perturbations caused by the treatment relative to the reference (defined as the treatment comparison showing the highest perturbation, i.e., the highest concentration of 3R4F after 8 months of exposure). For each treatment comparison (e.g., 3R4F at low concentration), the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the reference (i.e., 3R4F at 8 months). A δ value of 1 indicates that all the networks are perturbed by the same mechanisms.

5.9.4 Biological network perturbation amplitude analysis in the kidney

The star plots in Figure 101 show the biological processes that most contributed to the overall BIF phenotype at each time point.

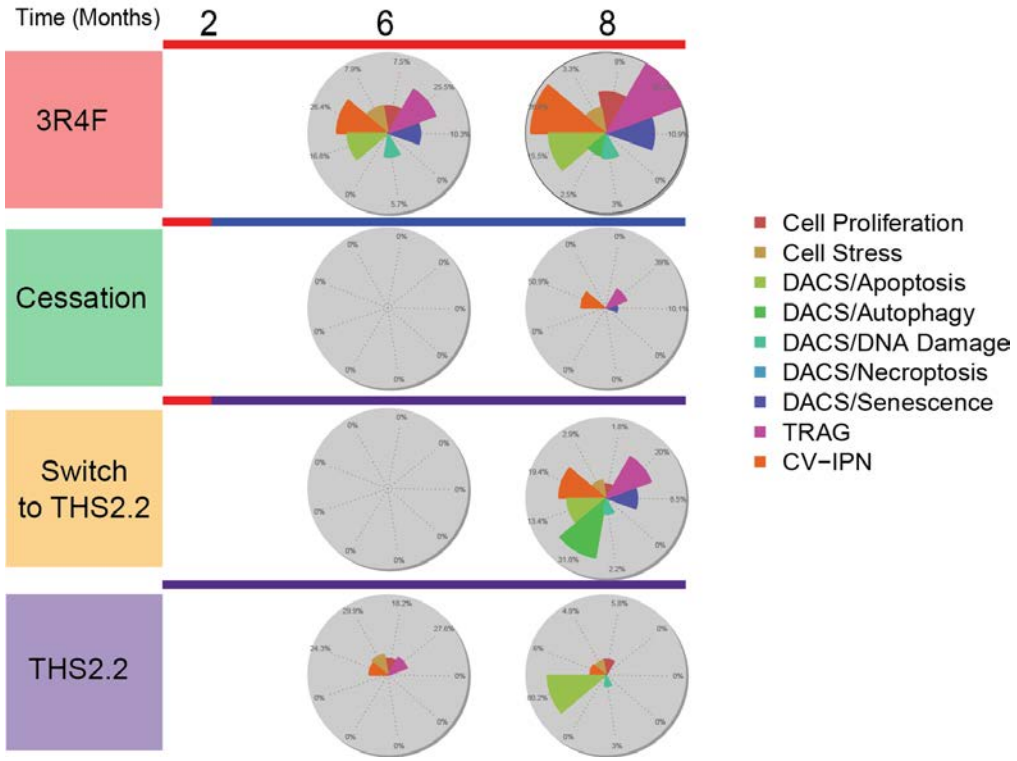


Figure 101: Star plots showing BIF calculated from the kidney

Star plots illustrating the decomposition of the overall RBIF (Figure 7) into its eight mechanistic components (from cell proliferation to inflammation, indicated by the colors of the segments) for each treatment group. The surface area of each segment is proportional to the contribution within a particular treatment of each network perturbation (labeled as %) and is adjusted by the overall relative BIF for the treatment compared with the reference, such that the sum of the segment areas for each treatment equals the RBIF for the treatment.

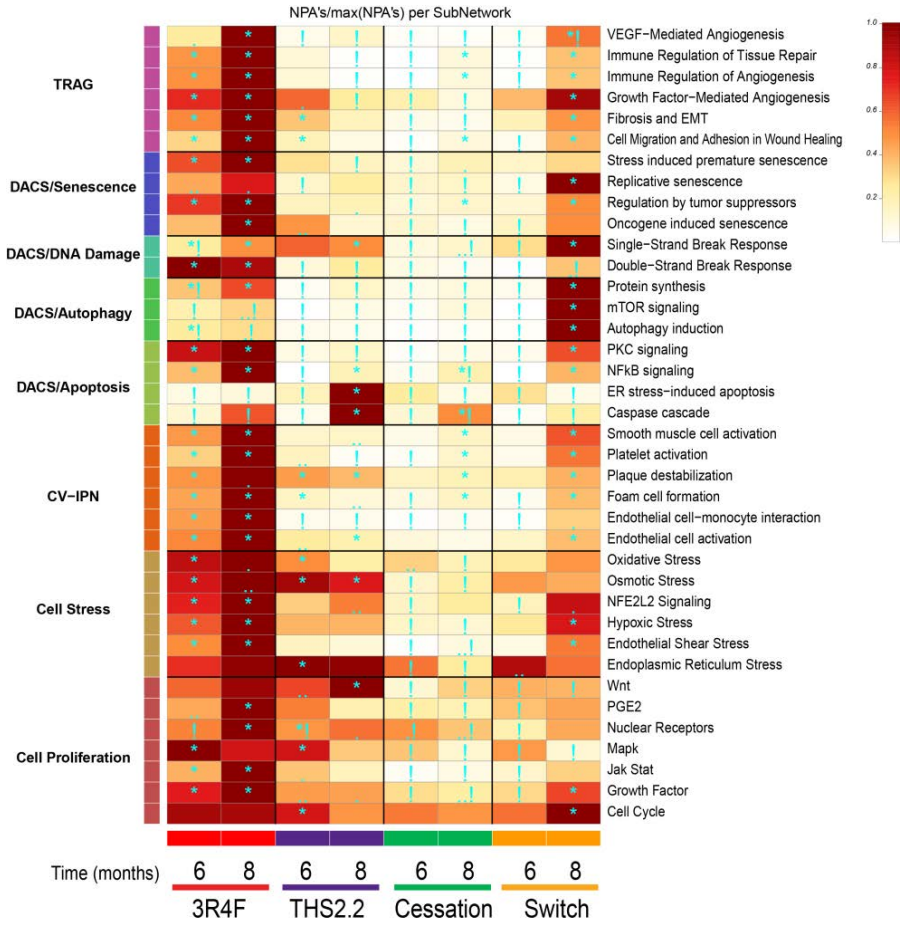


Figure 102: Heatmap of network perturbation amplitude scores in the kidney

Heatmap summarizing the subnetworks significantly perturbed in at least one treatment. A network is considered perturbed if, in addition to the significance of the NPA score with respect to the experimental variation, the two companion statistics (O and K) derived to inform on NPA score specificity with respect to the biology described in the network are significant.

Symbol legend: *: O and K statistic *P*-values below 0.05, :O and K *P*-values 0.05–0.075, .. :O and K *P*-values 0.075–0.1, !: not significant with respect to the experimental variation.

The similitude of NPA scores between contrasts was assessed in pairwise contrasts comparisons, similar to the overlap of differentially expressed genes, but instead for the overlap of significantly perturbed networks (Figure 103).

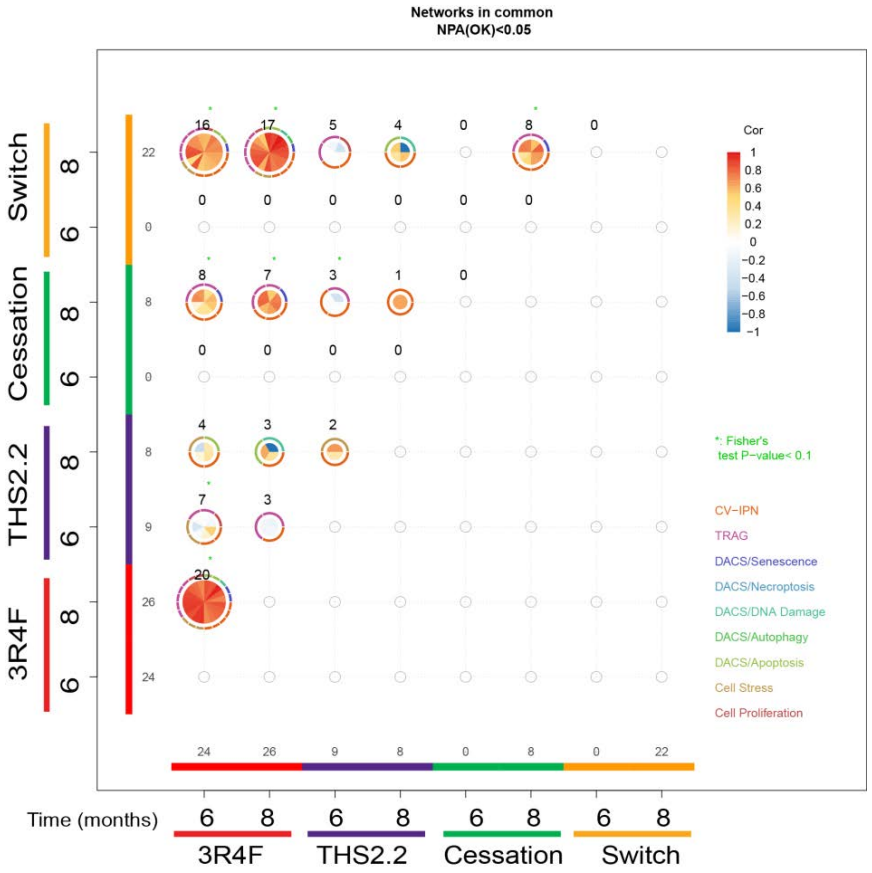


Figure 103: Overlap between sets of significantly perturbed networks in the kidney

For each pair of comparisons, the number of shared significantly perturbed networks is visualized. Colorscaling in the inner pie represents the correlation between backbone scores: red indicates positive correlation, while blue indicates negative correlation. A green star signifies that the overlap is significant (Fisher's test P<0.1). The number of significantly perturbed networks (OK statistics) is indicated in gray for each contrast (treatment versus sham at a given time point).

5.9.5 Hyp-level analysis: NPA1

To obtain a more exhaustive view of the regulation of upstream controllers by the different exposure types, the NPA1 algorithm computed the score for each hyp and contrast in the Selventa knowledgebase (Figure 104).

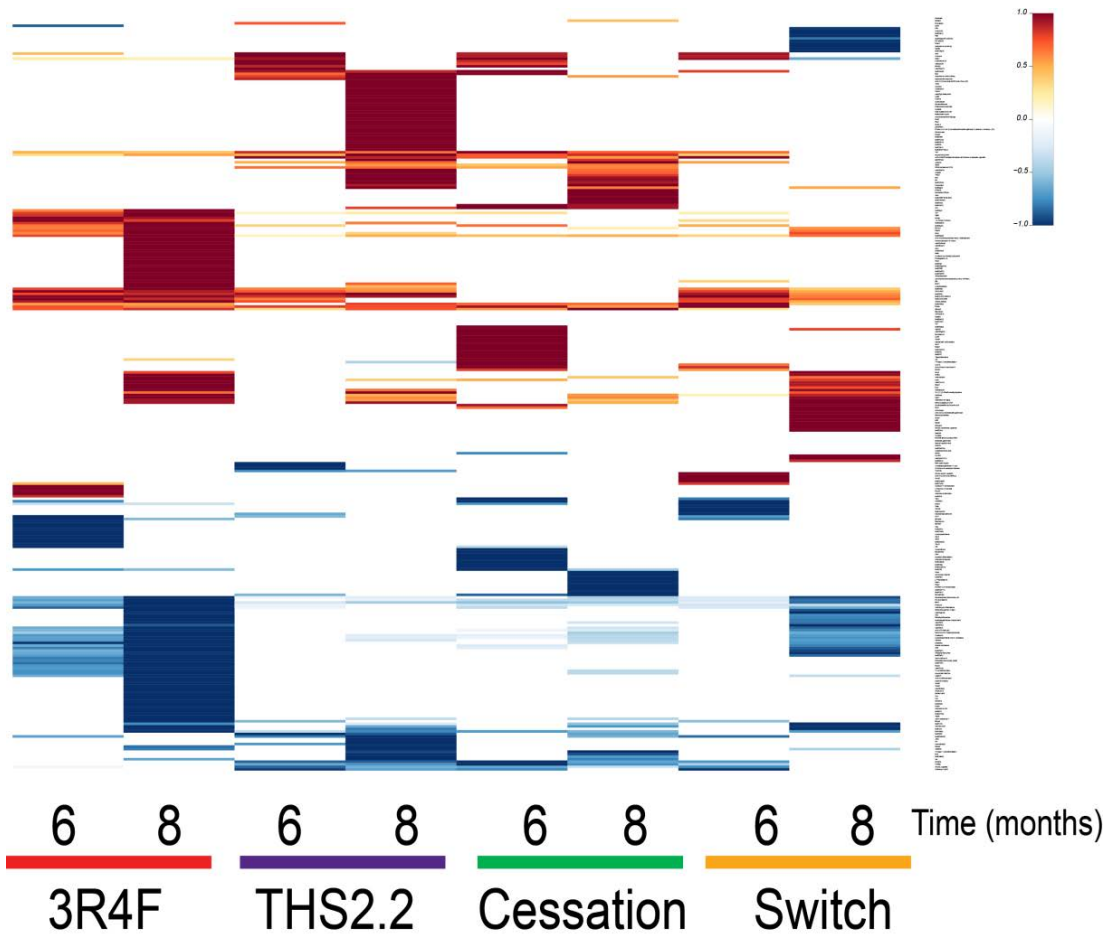


Figure 104: Heatmap of NPAI score for all hyps significant in at least one contrast in the kidney

5.9.6 NextBio analysis

To study the effect of 3R4F on the kidneys, a meta-analysis was performed comparing 3R4F-exposed and sham-exposed kidneys at two time points (6 and 8 months). Most biogroups were related to cell adhesion and mostly linked to the downregulated genes.

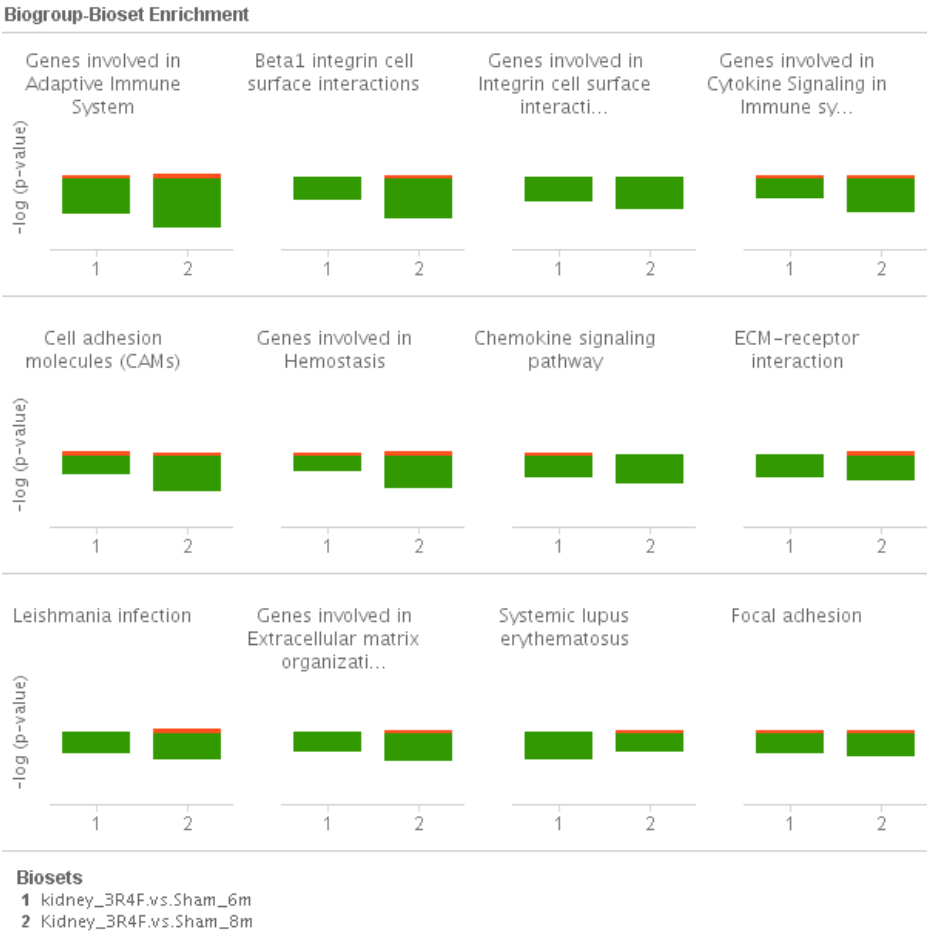


Figure 105: NextBio meta-analysis of the effects of 3R4F in the kidney

Each of the 10 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to 3R4F (compared with sham-exposed control). Each bar is composed of two parts: the red part is based on the strength of the overlap between the biogroup and up-regulated genes, while the green part is based on the strength of the overlap between the biogroup and down-regulated genes.

To study the effects of THS2.2 on the kidney, a meta-analysis was performed comparing THS2.2-exposed and sham-exposed kidneys at two different time points.

The top biogroups (Broad MSigDB canonical pathways) were mostly linked to lipid metabolism (Figure 106).

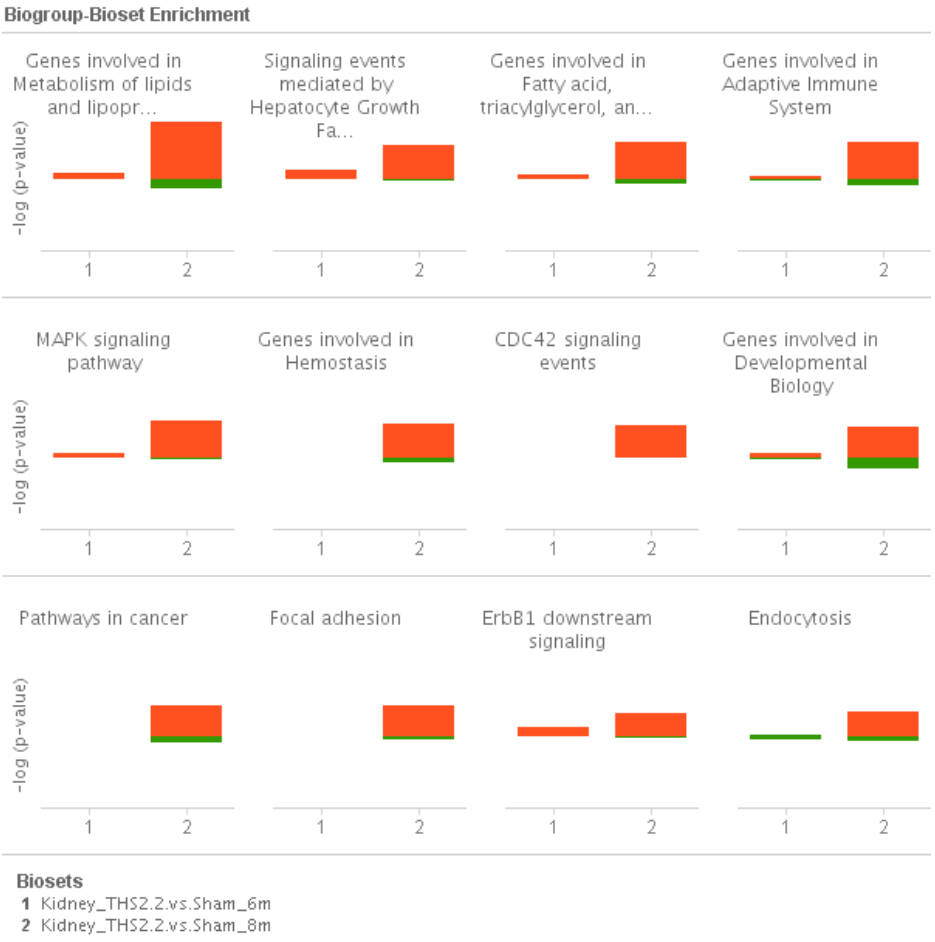


Figure 106: NextBio meta-analysis of the effects of THS2.2 on the kidneys

Each of the 10 graphs represents the score (expressed as –log10 (*P*-value)) for a canonical pathway at different time points of exposure to THS2.2 (compared with sham-exposed control). Each bar is composed of two parts: the red part is based on the strength of the overlap between the biogroup and up-regulated genes, while the green part is based on the strength of the overlap between the biogroup and down-regulated genes.

5.9.7 Gene set enrichment analysis

5.9.7.1 3R4F effects

To further study the trends within each treatment group over time, gene set enrichment analysis was carried out.



Figure 107: Gene set enrichment analysis in the 3R4F-exposed kidneys

5.9.7.2 THS2.2 effects



Figure 108: Gene set enrichment analysis in the THS2.2-exposed kidneys

5.9.7.3 Cessation vs. switch to THS2.2 effect

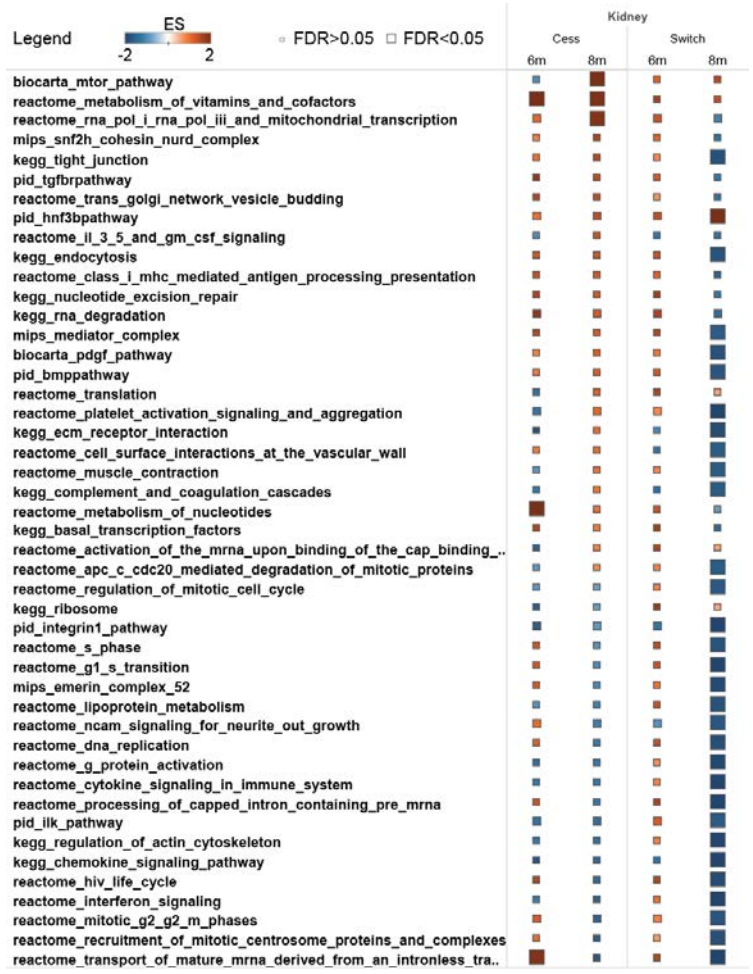


Figure 109: Gene set enrichment analysis in the kidneys of animals from the cessation and switch to THS2.2 groups

6 Conclusion

Overall, these data suggest that the use of THS2.2 will not expose individuals to additional hazards to those presented by smoking CCs.

7 Abbreviations

Abbreviation/Term	
BALF	bronchoalveolar lavage fluid
AGCC	Affymetrix GeneChip® Command Console™
BH	Benjamini–Hochberg
BIF	biological impact factor
CC	conventional cigarette
CYP	cytochrome P450
FDA	Food and Drug Administration
FDR	false discovery rate
GST	glutathione-S-transferase
HPHC	harmful and potentially harmful constituent
ISO	International Organization for Standardization
MRTP	modified risk tobacco product
N	number of individual values (or determinations)
NPA	network perturbation amplitude
NACLAR	National Advisory Committee for Laboratory Animal Research
OECD	Organisation for Economic Co-operation and Development
P	probability of receiving the value of the test statistic or a more extreme value of the test statistic, if the null hypothesis holds
pMRTP	prototypic modified risk tobacco product
QC	quality control
RBIF	relative biological impact factor
RNE	respiratory nasal epithelium
SMAR MD2AD	SMoking ARticle Master Design 2 with Aluminum Disc coating
THS	tobacco heating system
TPM	total particulate matter

For complete definitions, please refer to the PMI OPS Glossary and PMI RD Glossary.

8 Appendices

NA

9 References

1. Patskan, G., and Reininghaus, W. (2003) Toxicological evaluation of an electrically heated cigarette. Part 1: Overview of technical concepts and summary of findings. *Journal of applied toxicology : JAT* **23**, 323-328
2. Wong, E. (2014) Study 15006, 90 day OECD inhalation study on Sprague-Dawley rats to characterize toxicity of mainstream aerosol from zurich 4.2 cigarettes and comparison to 3R4F mainstream aerosol toxicity., Philip Morris International research Laboratories.
3. Phillips, B. (2014) Study 15015: In vivo evaluation of the effects of THS2.2 MRTP (DEL/C3/FR1/RDO) on the respiratory and cardiovascular systems of apolipoprotein E-deficient mice., Philip Morris International Research Laboratories Pte Ltd
4. Boue, S., De Leon, H., Schlage, W. K., Peck, M. J., Weiler, H., Berges, A., Vuillaume, G., Martin, F., Friedrichs, B., Lebrun, S., Meurrens, K., Schracke, N., Moehring, M., Steffen, Y., Schueller, J., Vanscheeuwijck, P., Peitsch, M. C., and Hoeng, J. (2013) Cigarette smoke induces molecular responses in respiratory tissues of ApoE(-/-) mice that are progressively deactivated upon cessation. *Toxicology* **314**, 112-124
5. Boue, S., Tarasov, K., Janis, M., Lebrun, S., Hurme, R., Schlage, W., Lietz, M., Vuillaume, G., Ekroos, K., Steffen, Y., Peitsch, M. C., Laaksonen, R., and Hoeng, J. (2012) Modulation of atherogenic lipidome by cigarette smoke in apolipoprotein E-deficient mice. *Atherosclerosis* **225**, 328-334
6. León, H., Boué, S., Peitsch, M., and Hoeng, J. (2013) Modulation of the Hepatic Lipidome and Transcriptome of Apoe-/-Mice in Response to Smoking Cessation. *J Liver* **2**, 2167-0889.1000132
7. Lietz, M., Berges, A., Lebrun, S., Meurrens, K., Steffen, Y., Stolle, K., Schueller, J., Boue, S., Vuillaume, G., Vanscheeuwijck, P., Moehring, M., Schlage, W., De Leon, H., Hoeng, J., and Peitsch, M. (2013) Cigarette-smoke-induced atherogenic lipid profiles in plasma and vascular tissue of apolipoprotein E-deficient mice are attenuated by smoking cessation. *Atherosclerosis* **229**, 86-93
8. International Organization for Standardization. (2000) ISO 3308:2000 - Routine analytical cigarette-smoking machine -- Definitions and standard conditions.
9. Davies, H. M., and Vaught, A. (2003) The reference cigarette. *Research cigarettes*
10. Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004) affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307-315
11. Gentleman, R., Bates, D., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J., and Zhang, J. (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology* **5**, R80
12. R Development Core Team. (2007) R: A Language and Environment for Statistical Computing.
13. Brettschneider, J., Collin, F., Bolstad, B. M., and Speed, T. P. (2008) Quality assessment for short oligonucleotide microarray data. *Technometrics* **50**, 241-264
14. McCall, M. N., Bolstad, B. M., and Irizarry, R. (2010) Frozen robust multiarray analysis (fRMA). *Biostatistics* **11**, 242-253
15. De Leon, H., Boue, S., Schlage, W. K., Boukharov, N., Westra, J. W., Gebel, S., VanHooser, A., Talikka, M., Fields, R. B., Veljkovic, E., Peck, M. J., Mathis, C., Hoang, V., Poussin, C., Deehan, R., Stolle, K., Hoeng, J., and Peitsch, M. C. (2014) A vascular biology network model

focused on inflammatory processes to investigate atherogenesis and plaque instability. *Journal of Translational Medicine* **12**, 185

16. Gebel, S., Lichtner, R. B., Frushour, B., Schlage, W. K., Hoang, V., Talikka, M., Hengstermann, A., Mathis, C., Veljkovic, E., Peck, M., Peitsch, M. C., Deehan, R., Hoeng, J., and Westra, J. W. (2013) Construction of a computable network model for DNA damage, autophagy, cell death, and senescence. *Bioinform Biol Insights* **7**, 97-117
17. Park, J. S., Schlage, W. K., Frushour, B. P., Talikka, M., Toedter, G., Gebel, S., Deehan, R., Veljkovic, E., Westra, J. W., Peck, M., Boue, S., Kogel, U., Gonzalez-Suarez, I., Hengstermann, A., Peitsch, M. C., and Hoeng, J. (2013) Construction of a Computable Network Model of Tissue Repair and Angiogenesis in the Lung. *Clinical Toxicology* **S12**
18. Schlage, W. K., Westra, J. W., Gebel, S., Catlett, N. L., Mathis, C., Frushour, B. P., Hengstermann, A., Van Hooser, A., Poussin, C., Wong, B., Lietz, M., Park, J., Drubin, D., Veljkovic, E., Peitsch, M. C., Hoeng, J., and Deehan, R. (2011) A computable cellular stress network model for non-diseased pulmonary and cardiovascular tissue. *BMC Syst Biol* **5**, 168
19. Westra, J. W., Schlage, W. K., Frushour, B. P., Gebel, S., Catlett, N. L., Han, W., Eddy, S. F., Hengstermann, A., Matthews, A. L., Mathis, C., Lichtner, R. B., Poussin, C., Talikka, M., Veljkovic, E., Van Hooser, A. A., Wong, B., Maria, M. J., Peitsch, M. C., Deehan, R., and Hoeng, J. (2011) Construction of a computable cell proliferation network focused on non-diseased lung cells. *BMC Syst Biol* **5**, 105
20. Westra, J. W., Schlage, W. K., Hengstermann, A., Gebel, S., Mathis, C., Thomson, T., Wong, B., Hoang, V., Veljkovic, E., Peck, M., Lichtner, R. B., Weisensee, D., Talikka, M., Deehan, R., Hoeng, J., and Peitsch, M. C. (2013) A Modular Cell-Type Focused Inflammatory Process Network Model for Non-Diseased Pulmonary Tissue. *Bioinformatics and biology insights* **7**, 167-192
21. Martin, F., Sewer, A., Talikka, M., Xiang, Y., Hoeng, J., and Peitsch, M. C. (2014) Quantification of biological network perturbations for mechanistic insight and diagnostics using two-layer causal models. *BMC bioinformatics* **15**, 238
22. Martin, F., Thomson, T. M., Sewer, A., Drubin, D. A., Mathis, C., Weisensee, D., Pratt, D., Hoeng, J., and Peitsch, M. C. (2012) Assessment of network perturbation amplitude by applying high-throughput data to causal biological networks. *BMC Syst Biol* **6**, 54
23. Kelstrup, C. D., Young, C., Lavallee, R., Nielsen, M. L., and Olsen, J. V. (2012) Optimized fast and sensitive acquisition methods for shotgun proteomics on a quadrupole orbitrap mass spectrometer. *Journal of proteome research* **11**, 3487-3497
24. Karp, N. A., Huber, W., Sadowski, P. G., Charles, P. D., Hester, S. V., and Lilley, K. S. (2010) Addressing accuracy and precision issues in iTRAQ quantitation. *Molecular & Cellular Proteomics* **9**, 1885-1897
25. Huber, W., Von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**, S96-S104
26. Herbrich, S. M., Cole, R. N., West Jr, K. P., Schulze, K., Yager, J. D., Groopman, J. D., Christian, P., Wu, L., O'Meally, R. N., and May, D. H. (2013) Statistical Inference from Multiple iTRAQ Experiments without Using Common Reference Standards. *Journal of proteome research* **12**, 594-604
27. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L. J. (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* **41**, D808-815
28. Fang, H., and Gough, J. (2014) The 'dnet' approach promotes emerging research on cancer patient survival. *Genome medicine* **6**, 64
29. Chen, J., Bardes, E. E., Aronow, B. J., and Jegga, A. G. (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **37**, W305-311
30. Efron, B., and Tibshirani, R. (2007) On testing the significance of sets of genes. *The annals of applied statistics*, 107-129
31. Väremo, L., Nielsen, J., and Nookaew, I. (2013) Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic acids research* **41**, 4378-4391

32. Thomson, T. M., Sewer, A., Martin, F., Belcastro, V., Frushour, B. P., Gebel, S., Park, J., Schlage, W. K., Talikka, M., Vasilyev, D. M., Westra, J. W., Hoeng, J., and Peitsch, M. C. (2013) Quantitative assessment of biological impact using transcriptomic data and mechanistic network models. *Toxicology and applied pharmacology* **272**, 863-878
33. Gautier, E. L., Shay, T., Miller, J., Greter, M., Jakubzick, C., Ivanov, S., Helft, J., Chow, A., Elpek, K. G., and Gordonov, S. (2012) Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nature immunology* **13**, 1118-1128
34. D'hulst, A., Vermaelen, K., Brusselle, G., Joos, G., and Pauwels, R. (2005) Time course of cigarette smoke-induced pulmonary inflammation in mice. *European Respiratory Journal* **26**, 204-213
35. Agarwal, A. R., Zhao, L., Sancheti, H., Sundar, I. K., Rahman, I., and Cadenas, E. (2012) Short-term cigarette smoke exposure induces reversible changes in energy metabolism and cellular redox status independent of inflammatory responses in mouse lungs. *American journal of physiology. Lung cellular and molecular physiology* **303**, L889-898
36. Noronha-Dutra, A. A., Epperlein, M. M., and Woolf, N. (1993) Effect of cigarette smoking on cultured human endothelial cells. *Cardiovasc Res* **27**, 774-778
37. Jorgensen, E., Stinson, A., Shan, L., Yang, J., Gietl, D., and Albino, A. P. (2008) Cigarette smoke induces endoplasmic reticulum stress and the unfolded protein response in normal and malignant human lung cells. *BMC cancer* **8**, 229
38. Van Straaten, J., Coers, W., Noordhoek, J. A., Huitema, S., Flipsen, J., Kauffman, H., Timens, W., and Postma, D. S. (1999) Proteoglycan changes in the extracellular matrix of lung tissue from patients with pulmonary emphysema. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc* **12**, 697-705
39. Hogg, J., and Senior, R. (2002) Chronic obstructive pulmonary disease c 2: Pathology and biochemistry of emphysema. *Thorax* **57**, 830
40. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011) Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740