

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und
Umwelt

Lehrstuhl für Experimentelle Genetik

Monitoring of volatile organic compounds in mouse breath as a new tool for metabolic phenotyping

Martin Kistler

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:

apl.Prof. Dr. J. Beckers

Prüfer der Dissertation:

1. Univ.-Prof. Dr. M. Hrabě de Angelis

2. Univ.-Prof. Dr. M. Klingenspor

3. Univ.-Prof. Dr. W. Wurst

Die Dissertation wurde am 20.04.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 21.11.2016 angenommen.

Content

Content.....	III
Abbreviations.....	VIII
Zusammenfassung.....	X
Abstract.....	XII
1. Introduction.....	1
1.1. Breath gas analysis as a window to metabolic functions: history and methodological evolvment.....	1
1.2. Human volatile organic compounds in biomedical research and application in clinics	2
1.3. Obesity is a result of disturbed energy homeostasis	3
1.4. Selected mechanisms leading to obesity	4
1.4.1. Mechanisms in neuronal and endocrine regulation of energy homeostasis	4
1.4.2. Environmental obesogenic stimuli.....	5
1.4.3. Obesity-associated shifts in metabolism and related pathologies	5
1.5. Limitations in human VOC analysis and the laboratory mouse as model organism	7
1.6. Aims of this thesis.....	8
2. Methods	9
2.1. Housing conditions	9
2.2. Animals and diets.....	9
2.2.1. Cohorts for effects of diet matrix.....	9
2.2.2. Diet induced obese mice	10
2.2.3. Genetic obese mouse model MC4R (W16X).....	10
2.3. Methods for body mass and body composition monitoring.....	11
2.3.1. Body mass monitoring	11

2.3.2.	Nuclear magnetic resonance (NMR) body composition monitoring	11
2.4.	Analysis of volatile organic compounds in unrestrained mice	11
2.4.1.	Proton transfer mass spectrometry	11
2.4.1.1.	Measurement principle and instrument	11
2.4.1.2.	Operation settings and recording of spectra	13
2.4.2.	A setup to measure volatile organic compounds in mice	14
2.4.3.	Measurement protocol and contamination monitoring	15
2.5.	Primary data analysis	16
2.5.1.	Calculation of source strengths	16
2.5.2.	Profile selection and quality control	16
2.6.	Secondary data analysis	19
2.6.1.	Machine learning protocol for feature selection	19
2.6.2.	Data visualization	19
2.6.3.	Statistical testing	20
3.	Results	22
3.1.	Publication I	22
3.2.	Publication II	25
4.	Discussion	29
4.1.	The effect of diet matrix on the emitted VOC pattern	29
4.1.1.	Change from chow to semi-purified diets	29
4.1.2.	Methanol (pk33B, pk15, pk34B, pk35B, pk51B)	30
4.1.3.	Methyl acetate (pk75, 76B)	30
4.1.4.	Dimethyl sulfone (pk95A)	31
4.2.	Obesity-induced changes in the volatilome	32
4.2.1.	A common VOC signature of obesity	32
4.2.1.1.	Acetic acid (pk61 and pk43A)	32
4.2.1.2.	Methanol (pk33B and 35B)	33

4.2.1.3.	Carbon dioxide*H ₂ O / Dimethyl sulfide (pk63, pk64B, pk65).....	34
4.2.1.4.	(Methylthio)methanethiol (MTMT, pk62).....	35
4.2.2.	Volatiles specifically altered in diet-induced obesity	36
4.2.2.1.	Ammonia (pk18).....	36
4.2.2.2.	Acrolein (pk57B)	37
4.2.2.3.	Methyl acetate (pk75)	37
4.2.2.4.	¹⁸ O ¹⁶ O oxygen (pk34A)	38
4.2.2.5.	H ₃ O ⁺ .(H ₂ O) ₂ water cluster and fragments of aldehydes (pk55B)	38
4.2.2.6.	Acetone and propanol (pk59, pk43B).....	38
4.2.3.	Volatiles specifically altered by MC4R nonsense knock-in	39
4.2.3.1.	Cluster of unknowns 151, 153, 137B and 81B	39
4.2.4.	Unassigned candidate VOCs altered in obesity and strategies for identification.....	40
4.3.	On the methodology of breath gas analysis in rodents	41
4.3.1.	On the evolution of the unrestrained mouse VOC measurement and the corresponding data analysis.....	41
4.3.2.	Advantages and limitations of the established setup in comparison to published literature and implications on human breath analysis	42
5.	Conclusion and Outlook	46
6.	References.....	47
7.	Acknowledgments.....	60
8.	Appendix.....	61
8.1.	Supplementary figure.....	61
8.2.	Publication I.....	62
8.2.1.	Abstract.....	63
8.2.2.	Introduction.....	64
8.2.3.	Methods.....	65
8.2.3.1.	Animals	65

8.2.3.2.	Experimental diets.....	65
8.2.3.3.	Proton-transfer time-of-flight mass spectrometry	66
8.2.3.4.	Setup for real-time breath gas analysis in unrestrained mice.....	67
8.2.3.5.	Monitoring of and controlling for contaminations.....	67
8.2.3.6.	Data analysis	68
8.2.3.7.	Statistics	69
8.2.4.	Results.....	71
8.2.4.1.	Body mass gain in response to feeding the semi-purified diets	71
8.2.4.2.	Identification of discriminating volatile organic compounds for diet matrix	71
8.2.4.3.	Reduced source strengths of volatiles after change from laboratory chow to semi-purified diets	72
8.2.5.	Discussion	75
8.2.6.	Conclusion	77
8.3.	Publication II.....	82
8.3.1.	Introduction.....	84
8.3.2.	Material and methods.....	85
8.3.2.1.	Mice, animal housing and challenge experiments	85
8.3.2.2.	Proton-transfer reaction time-of-flight mass spectrometry and protocol for real-time breath gas analysis in unrestrained mice.....	86
8.3.2.3.	Data analysis and statistics.....	86
8.3.3.	Results.....	89
8.3.3.1.	Obesity state of HFD fed and mono-genetic mice	89
8.3.3.2.	Selection of VOCs relevant for classification.....	90
8.3.3.3.	Visualization of selected source strength data	92
8.3.3.4.	Effects on VOC signature in diet-induced and monogenetic obesity	94
8.3.3.5.	Gaussian graphical modeling as a tool to identify VOCs	98
8.3.4.	Discussion	100
8.3.4.1.	A VOC signature altered in the volatilome both obesity models.....	101

8.3.4.2.	Significant HFD specific peaks.....	103
8.3.4.3.	MC4R-ki specific peaks.....	105
8.3.4.4.	Unassigned volatiles	106
8.3.5.	Conclusion	106

Abbreviations

°C	degree Celsius
adh	alcohol dehydrogenase
ANOVA	analysis of variance
AUC	area under the curve
BMI	body mass index
c_0	starting concentration
c_i	current concentration
E	electric field
en%	energy percent
F	flowrate
FDR	false discovery rate
FELASA	Federation of Laboratory Animal Science Associations
g	gram
GC-MS	gas chromatography – mass spectrometry
GMC	German Mouse Clinic
h	hour
H ₃ O ⁺	Hydronium ion
HFD	high fat diet
HL	high lard diet
hs-PTR-MS	high sensitivity – proton transfer reaction – mass spectrometry
IMS	ion mobility spectrometry
kg	kilogram
kGy	kilo gray
kHz	kilohertz
kJ	kilojoule
l	liter
LL	low lard diet
LFD	low fat diet
mbar	millibar
MC4R	melanocortin 4 receptor
MC4R-ki	melanocortin 4 receptor - knock in
min	minute

ml	milliliter
MTMT	(methylthio)-methanethiol
n	(animal) number
N	buffer gas density / concentration of neutral particles
NMR	nuclear magnetic resonance
PEEK	Polyether ether ketone
p	sample number
PA	proton affinity
pk	peak
ppb	parts per billion, 1×10^{-9}
ppbv	parts per billion, 1×10^{-7} Vol-%
ppm	parts per million, 1×10^{-6}
ppmv	parts per million, 1×10^{-4} Vol-%
ppt	parts per trillion, 1×10^{-12}
pptv	parts per trillion per volume, 1×10^{-10} Vol-%
PTFE	polytetrafluoroethylene
PTR-MS	proton transfer reaction – mass spectrometer <i>or</i> proton transfer reaction – mass spectrometry
PTR-TOF-MS	proton transfer reaction – time of flight – mass spectrometer <i>or</i> proton transfer reaction – time of flight – mass spectrometry
RF	random forest
ROC	receiver operating characteristics
s	second
S_i	source strength
SIFT	selected ion flow tube
SIFT-MS	selected ion flow tube - mass spectrometry
SPME	solid phase micro-extraction
t	time
Td	Townsend
TOF	time of flight
V	Volt
V	chamber volume
VOC(s)	volatile organic compound(s)

Zusammenfassung

Die Nutzung von Biomarkern zur Diagnostik von Umweltexpositionen, Erkrankungswahrscheinlichkeiten und dem Vorhandensein von Erkrankungen ist ein wertvoller Bestandteil des klinischen Alltags, was die Etablierung von eben solchen Markern zu einem ein Hauptziel der biomedizinischen Forschung macht. Neben einer hohen Spezifität und Sensitivität sollten Biomarker auf möglichst nichtinvasivem Wege bestimmbar sein. Die Analytik von volatilen organischen Substanzen (volatile organic compounds, VOCs) aus dem Atem stellt eine nichtinvasive Methodik zur Überwachung von Stoffwechselfunktionen und Erkrankungen dar. Eine Verschiebung des VOC-Musters wurde bereits in verschiedenen Pathologien beschrieben, allerdings besteht selbst in gesunden Menschen eine erhebliche intrapersonelle Varianz und die (biochemische) Herkunft der diagnostisch relevanten VOCs ist oftmals noch ungeklärt. Die Verwendung von Mausmodellen kann zur Beantwortung der in Humanstudien offenen Fragen beitragen.

Die Zielsetzung dieser Dissertation war die Etablierung und Weiterentwicklung einer neuen Methode zur Bestimmung von VOCs in unfixierten und nicht anästhesierten Mäusen zum Zwecke der metabolischen Phänotypisierung. Hierzu wurde eine Analyseplattform entwickelt, welche aus Akkumulationsprofilen der einzelnen VOC Emissionsraten (bezeichnet als Quellstärke) berechnet. Dies basiert auf Daten welche mittels eines Protonen-Transfer-Reaktions Massenspektrometers mit Flugzeitdetektion (PTR-TOF-MS) erhoben wurden. Darauf aufbauend wurden die Auswirkungen von Diätinterventionen und Adipositas in Mäusen untersucht. Die Dissertation wurde in zwei Studien unterteilt.

In der ersten publizierten Studie wurde der Effekt der Futtermatrix auf VOC Quellstärken untersucht, welche als Unterschied zwischen einem Getreide-basierten Standardfutter und einer aufgereinigten Experimentaldiät definiert wurden. Zu diesem Zweck wurden VOCs von männlichen C57BL/6J Mäusen gemessen, welche zunächst mit Standarddiät gefüttert wurden. In der Folge wurden die Tiere auf vier Gruppen aufgeteilt, welche sich in Quantität und Qualität des Fettanteils der aufgereinigten Diäten unterschieden, und wiederholt gemessen. Mittels eines random forest Models wurden diejenigen VOC Quellstärken bestimmt, welche auf die Futterumstellung reagierten. In einer Gruppe von drei VOCs wurde eine drastische Reduzierung der Quellstärke festgestellt; diese wurden vorläufig als Methanol, Methylacetat/Propionat und Dimethylsulfon identifiziert. Die Matrix der Diät stellt somit einen relevanten Einflussfaktor der gemessenen VOC Konzentrationen und des

verbundenen Stoffwechsels dar. Weiterhin ist sie eine potentielle Quelle von Varianz, welche in humanen VOC-Studien beobachtet wird.

In der zweiten publizierten Studie wurde der Effekt von Adipositas auf das gemessene VOC Muster untersucht. Zu diesem Zweck wurden Mäuse auf einer Hochfettdiät (HFD) und mono-genetische adipöse Mäuse (globale knock-in Mutation im Melanokortin-4-Rezeptor, MC4R-ki) sowohl unter *ad libitum* Zugang zu Futter als auch gefastet analysiert. Für jeden Versuch wurde ein random forest-Model erstellt; es ergaben sich 26 VOCs in MC4R-ki und 22 VOCs in HFD Mäusen mit relevanten Änderungen. Die Schnittmenge zwischen den beiden Adipositasmodellen betrug acht VOCs. Interessanterweise stammen die vorläufig identifizierten VOCs aus Prozessen wie Lipidperoxidation, Ketogenese und Pheromonproduktion. Zusätzlich konnte gezeigt werden, dass die Anwendung Gaußscher Graphischer Modelle sowohl die chemische als auch die metabolische Herkunft mehrerer VOCs offenbaren kann. Diese Ergebnisse unterstreichen das Potential von VOCs im Hinblick auf Identifikation und Überwachung von Stoffwechsel assoziierten Erkrankungen.

Zusammenfassend demonstriert der deutliche Einfluss von Diät und Adipositas auf eine Reihe von VOCs dass Rückschlüsse auf den Energiestoffwechsel aus VOC Muster gezogen werden können. Auch müssen Diät und Adipositas als Einflussfaktoren in Studien am Menschen Beachtung finden. Weiterhin existiert eine Gruppe an VOCs welche langfristig die Möglichkeit zur nichtinvasiven Beurteilung von Adipositas assoziierten Erkrankungen verspricht.

Abstract

The use of biomarkers associated with environmental exposure, disease presence or risk is a common strategy in clinical diagnosis and thus the establishment of such markers a primary goal for biomedical research. Besides high specificity and sensitivity, biomarkers should preferably be measurable applying a minimal invasive technique. The analysis of volatile organic compounds (VOCs) from breath is a non-invasive method to measure exhaled metabolites in order to monitor metabolic functions and disease states. Changes in VOC signature have been described in various diseases, however even healthy humans show considerable inter-individual variability and (biochemical) origins of changed VOCs are not understood in many cases. These limitations in human studies can be addressed using mouse models.

The aim of this thesis was the establishment and further development of a novel method to analyze VOCs in un-restrained and non-anaesthetized mice as a tool for metabolic phenotyping. An analysis platform was developed to calculate emission rates (so-called source strengths) from VOC accumulation profiles measured using a proton transfer reaction mass spectrometer with time-of-flight detection (PTR-TOF-MS). Based on this, the effects of dietary interventions and obesity in mice on the VOC signature was studied. The thesis was divided in two studies.

In publication I, the effect of the diet-matrix, defined as the difference between a grain-based chow diet and semi-purified intervention diets, was analyzed. Therefore, VOCs of male C57BL/6J mice were quantified while having *ad libitum* access to chow diet and after diet switch to four semi-purified diets varying in fat quantity and quality. A random forest model was created to identify VOC source strengths responsive to this diet matrix change. Interestingly, a set of three volatiles dropped drastically after the diet switch and were tentatively identified as methanol, methyl acetate/propionate and dimethyl sulfone. Thus, diet matrix is a relevant modulator of measured VOC levels as well as underlying metabolic functions and is one potential source of variability in exhaled VOCs observed in human studies.

In publication II, the aim was to determine the effect of obesity on the VOC signature. For this, both high fat diet-induced obese (HFD) and mono-genetic obese mice (global knock-in mutation in melanocortin-4 receptor MC4R-ki) were analysed in *ad libitum* fed as well as fasted state. After building random forest models for both experiments, 26 candidate VOCs in MC4R-ki versus 22 candidate VOCs in HFD mice were detected. In overlap, eight VOCs were found to be altered in both mouse models. Interestingly, among tentatively identified VOCs processes like lipid peroxidation,

ketogenesis and pheromone production were detected. In addition, it could be shown that by applying a gaussian graphical model both chemical and metabolic origins of several VOCs can be revealed. These findings show the potential of VOCs to identify and monitor metabolic disease states.

In conclusion, the clear impact of both diet and obesity on a variety of VOCs underline that conclusions about the status of energy metabolism can be drawn from VOC signatures and that diet and obesity need to be accounted for in human studies. Furthermore, a set of VOCs offers the opportunity to be used to non-invasively assess obesity associated disease risks.

1. Introduction

1.1. Breath gas analysis as a window to metabolic functions: history and methodological evolution

An important goal in biomedical research is the discovery of biomarkers, which can be used to assess current health or future disease risks of a person. A range of biomarkers can only be determined invasively, e.g. with assays requiring blood or tissue samples. The sampling of such material is often inconvenient or even painful. The analysis of volatile organic compounds (VOCs) from breath is a rapidly emerging field, which by the mode of sampling is non-invasive and thus patient-friendly. Furthermore, the time necessary for breath analysis is typically low, in the range of seconds to hours depending on the methodology. This allows for fast diagnostic read outs as well as for monitoring of long term changes. Volatile organic compounds originating from different tissues are transported by the blood and also transferred to alveolar gaseous phase (Herbig and Amann, 2009). Thus, breath VOCs can be a window to whole body metabolic processes which are typically not measurable in traditional (liquid) metabolomics techniques due to their volatility.

Volatile organic compounds are legally defined as any organic compound having an initial boiling point less than or equal to 250 °C (482 °F) measured at a standard atmospheric pressure of 101.3 kPa (European Parliament, Council of the European Union, 2004). Although this includes only organic, thus carbon containing compounds, other volatiles like e.g. volatile sulfur compounds (VSCs) not covered by this definition are typically also analyzed from breath with the same methodologies and in parallel. Interestingly, VOCs contain a variety of chemical heterogeneous substances, including but not limited to alcohols, aldehydes, ketones, acids, esters or hydrocarbons.

Historically, such compounds have been detected by smell and were known to be informative since ancient Greece, when Hippocrates of Kos first mentioned *fetor oris* and *fetor hepaticus* (Amann et al., 2014a). A major step forward in breath analysis was the work of Antoine Lavoisier, who discovered carbon dioxide in breath in the late 18th century (Donovan, 1996). His discoveries led to the development of indirect calorimetry, still one of the most widely-used breath-based methods. Following in the late 19th century, first organic volatiles like acetone could be described and quantified from breath (Müller, 1898). However, the modern era of VOC analysis from breath was initiated from Linus Pauling's article which highlighted 250 substances in a sample of breath (Pauling et al., 1971), with now up to 1849 known and further 1765 identified volatiles from various biomaterials of the body (Costello et al., 2014).

This scientific gain in knowledge was only possible due to technical progress in volatile compound analysis. For the analysis of VOCs in gaseous phases, a variety of technologies have been established. In many respects, the “gold-standard” for the analysis of VOCs is still the gas chromatography – mass spectrometry technique with optional additional fragmentation. Herein, a pre-concentrated gaseous sample is injected into a carrier gas flowing over a capillary column to chemically separate volatiles and to obtain a retention time. In an additional step, mass spectra of the fragments of a certain volatile can be recorded to identify the molecule structure using established databases. However, despite the high specificity, GC-MS is not the ideal technique for every application. GC-MS time requirements are rather high due to the necessary pre-concentration step as well as the measurement time, which can be in the range of hours for several samples (Cikach and Dweik, 2012). In addition, the sensitivity for different chemical classes of volatiles in the system is affected by pre-concentration. The selected column material is crucial for sensitivity, which is further reduced due to dilution of VOCs in the carrier gas. Also high water content in the sample as it is present in breath can be problematic (Miekisch and Schubert, 2006). Other analytical methods were developed to overcome these limitations in breath matrices, as Ion Mobility Spectrometry (IMS), selected ion flow tube mass spectrometry (SIFT-MS) and proton-transfer reaction mass spectrometry (PTR-MS). The later technique was specifically designed to detect volatile organic molecules from air in a concentration smaller than ppm and down to ppt range (Lindinger et al., 1998). In contrast to GC-MS, especially in PTR-MS a high time resolution is combined with high sensitivity due to the fact that the exhaled air itself is used as carrier gas for VOCs. Although some of the molecule specificity of GC-MS is lost, PTR-MS is highly useful to follow dynamics in VOC responses in the range of seconds (Biasioli et al., 2011).

1.2. Human volatile organic compounds in biomedical research and application in clinics

Along with technical progress, VOCs have been used in various ways in medical research and diagnosis over the past decades (Amann et al., 2014b). There are now several stable-isotope based challenge tests where compounds are metabolized and degradation products as $^{13}\text{C}\text{O}_2$ are measured. Probably known best are those targeting microbial degradation like ^{13}C -Urea test diagnosing gastric *Helicobacter pylori* infections, the use of breath hydrogen for certain carbohydrate intolerances or intestinal overgrowth (Eisenmann et al., 2008; Klein et al., 1996). In addition, challenges targeting endogenous metabolism like the ^{13}C -methacetine or the ^{13}C -aminopyrine breath tests to estimate hepatic enzyme activity (e.g. CYP1A2) are actively used in clinics (Giannini et al., 2005; Lalazar et

al., 2009). Further interests include the measurement of the human exposome, e.g. environmental noxes or the cultivation or infection of VOC-emitting bacteria or fungi as reviewed in (Amann et al., 2014b). Another major approach in VOC research is the analysis of breath in certain diseases to detect volatile biomarkers for non-invasive screening and monitoring. Such biomarker discovery studies are performed in pathologies of liver (Chen et al., 1970; Morisco et al., 2013; Van den Velde et al., 2008), kidney (Kohl et al., 2013; Meinardi et al., 2013) and respiratory diseases like asthma (Paredi et al., 2000; Simpson and Wark, 2008), cystic fibrosis (Barker et al., 2006) and chronic obstructive pulmonary disease (McCurdy et al., 2011). Furthermore, cancers in lung (Bajtarevic et al., 2009; Hakim et al., 2012), liver (Qin et al., 2010), breast (Li et al., 2014) and other organs (Peng et al., 2010) are known to change VOC signature. However, the outcome of these mostly small trials is typically a set of volatile marker substances associated with a certain disease. The field has just started to attribute parts of those disease associations to general pathophysiological states e.g. of oxidative stress, inflammation and carcinogenesis (Boots et al., 2012), but many are not described in origin and metabolism yet. Another research area in which breath gas analysis can be used is the field of energy homeostasis and associated diseases.

1.3. Obesity is a result of disturbed energy homeostasis

The etiology of the most common diseases has changed in the industrialized western world. Historically, the main causes for death were infectious diseases, nowadays taken over by non-communicable diseases like cancer, cardiovascular disease or type 2 diabetes (Mendis and World Health Organization, 2014). Obesity is a risk factor for all of the mentioned diseases. Normal weight is defined as a body mass index (BMI, height/ mass²) between 18.5 and 25 kg/m², whereas overweight starts from 25 kg/m² and obesity is specified with a BMI of greater than 30 kg/m², with class II from 35 kg/m² and class III from 40 kg/m² on. According to the World Health Organization, the obesity epidemic is no longer only taking place in industrialized but also in emerging nations leading to a doubling in obesity rates from 1980 and to a prevalence 11% in male and 15% of female adults worldwide (Mendis and World Health Organization, 2014).

Obesity as a state of extensive systemic energy storage is a result of a disturbed energy homeostasis by altered in- and effluxes to the system. A positive energy balance is fundamental for the change of body mass and the development of obesity, thus either elevated energy flux into the system or reduced energy expenditure must be present (Hall et al., 2012). Energy intake consists predominantly of the metabolizable energy from three macronutrients carbohydrates, proteins and fat

out of beverages and food. The proportion of energy which can be assimilated and metabolized from a diet varies depending on diet composition or food processing. The metabolized energy is either utilized to maintain ongoing metabolic processes or stored to enable the organism to endure periods of limited energy intake. Energy is predominantly stored in adipose tissue in the form of triglycerides, although stored carbohydrates like glycogen are utilized for short term energy supply. In addition, amino acids can be mobilized from body protein for energy demand. The energy expenditure, on the other side, consists of basal metabolic rate, dietary induced thermogenesis, thermoregulation, physical activity, growth and sexual reproduction.

1.4. Selected mechanisms leading to obesity

Basically, mechanisms leading to extension of stored energy and finally obesity can be found within all physiological processes associated with energy balance regulation. Over the past decades, multiple mechanisms and underlying physiological systems leading to obesity have been identified.

1.4.1. Mechanisms in neuronal and endocrine regulation of energy homeostasis

The energy homeostasis needs to be regulated in a very precise manner to maintain body weight at a certain set point. In humans, for example, there are estimates that for becoming obese at the age of 50+, a 25 year old person with a BMI of 25 needs an over-consumption or under-expenditure of 680 kcal/day (Katan and Ludwig, 2010). Smaller deviations, which occur from day to day, are typically balanced by endogenous regulatory systems. For the control of energy balance, the hypothalamus plays a major role in integrating central and peripheral signals of short and long term energy storage and demands. A major afferent signal for adipose energy storage is leptin (Zhang et al., 1994). This peptide hormone is secreted to bloodstream by white adipose tissue proportional to the amount of fat stored. From circulation, leptin activates a subpopulation of hypothalamic neurons in the arcuate nucleus carrying leptin receptors. Those neurons express proopiomelanocortin (POMC) and project to melanocortin-receptor 4 positive cells. In POMC expressing neurons, α -melanocyte-stimulating hormone (α -MSH) is cleaved from POMC and activates MC4R-positive cells to inhibit food uptake and increase energy expenditure (Myers et al., 2009). In states of low adipose fat storage or fasting, leptin levels are reduced (Ahima et al., 1996). Along with that, other signals like ghrelin activate neurons projecting to agouti-related peptide (AgRP) expressing neurons. AgRP positive neurons are

activated and inhibit both MC4R-positive as well as POMC neurons mediated by γ -aminobutyric acid (GABA) release. Several well established mouse models for obesity knock out genes relevant in this system, as the ob/ob (Leptin receptor), db/db (leptin) or mc4r^{-/-} mice. A variant of the later mouse mutant using a knock-in of a nonsense mutation (W16X) found in obese human patients (Marti et al., 2003) was studied in this thesis.

1.4.2. Environmental obesogenic stimuli

BMI is reported to feature a heritability of 40-70 %, of which common variations found in GWAS are estimated to explain up to 20% (Locke et al., 2015). The vast majority of the obesity heritability cannot be explained in genome-wide association studies, thus other mechanisms like epigenetic inheritance are discussed to drive this phenomenon. Environmental factors and lifestyle are known major modulators of both epigenetics and obesity risk, in interaction with genetic predisposition (Temelkova-Kurktschiev and Stefanov, 2012). On population level, both increased availability of energy dense diets and sedentary lifestyle are observed and contribute to the obesity epidemic. The so-called “western diet” includes high intake of animal protein, saturated and ω -6 fatty acids combined with simple carbohydrates, of which a considerable amount is consumed in beverages. In addition such a diet is typically rich in sodium chloride, low in other micronutrients as e.g. potassium, features only small amounts of unprocessed fruit and vegetables and is thus low in fiber (Cordain et al., 2005). Typically, to study the effects of such dietary challenge rodent models are used and exposed to variants of either high sugar and/or high fat containing diets. This leads, at least in diet-induced obesity susceptible mice as the C57Bl/6J strain used in this thesis, to an obese phenotype in comparison to control diets.

1.4.3. Obesity-associated shifts in metabolism and related pathologies

The storage of triglycerides in adipose tissue is a physiological necessity. In obesity, the most associated adverse effects on health do not originate in excessive energy storage per se, but are due to a set of processes increasing risks for comorbidities. The pathophysiological state often found to be associated with obesity includes low-grade chronic inflammation and oxidative stress, insulin resistance and resulting elevated (fasting) glucose levels, elevated triglycerides, low high density lipoprotein (HDL) levels, high low density lipoprotein levels, ectopic fat deposition and elevated blood pressure (Després and Lemieux, 2006). In consequence, the prevalence of cardio-vascular

diseases, type II diabetes, musculoskeletal disorders like osteoarthritis and several types of cancer is elevated in obese populations (Guh et al., 2009). Given that in obesity a multitude of risk factors arise, the potential benefit of monitoring obese people to prevent development of comorbidities is obvious. Therefore, biomarkers for such personalized risk assessment need to be established which are robust, cheap to measure and ideally non-invasive to ensure subject compliance and practical feasibility. All these characteristics can potentially be met by analysis of VOCs from breath.

1.4.4. Energy metabolism and diet relevant findings in VOC analysis

Interestingly, a number of studies focusing on the diagnosis and monitoring of nutrition, energy metabolism and associated diseases has been conducted. Regarding the influence of diet on breath VOC signatures, multiple mechanisms of contribution are known (Ajibola et al., 2013). Volatiles can be released directly from foods and diet and, after systemic passage and potential modification in metabolism can be found in breath. The time range of such compounds found in breath is known to be from minutes up to days for more lipophilic substances (Pellizzari et al., 1992). Furthermore, diet can induce physiological changes and adaptations, which transfer to altered VOC emission. Diets lacking carbohydrates, for example, induce gluconeogenesis as well ketogenesis. The ketone body acetone is highly correlated to both acetoacetate and β -hydroxybutyrate and can be used to monitor systemic ketone body levels from breath (Musa-Veloso et al., 2002). In addition, diet is known to be the major modulator for microbial composition in the gut (Cotillard et al., 2013; David et al., 2014; Zarrinpar et al., 2014). Microbial degradation is another mechanism by which diet can contribute volatiles in breath. This is well described for fiber consumption, for example, where methane and hydrogen are released and can be measured in breath (Rumessen, 1992), two volatile substances which interestingly have been associated with increased BMI when elevated after lactulose breath test (Mathur et al., 2013).

In the area of energy metabolism and associated diseases, a limited number of VOC studies in human subjects can be found. In a study to address sleep apnea in obesity, no clear separation according to sleep parameters could be found, although obese patients showed elevated VOCs for inflammatory processes (Dragonieri et al., 2015). In obese children, elevation in four volatiles, namely isoprene, 1-decene, 1-octene, ammonia and hydrogen sulfide could be found (Alkhoury et al., 2015). Regarding obesity-associated diseases, some work was performed on hepatic pathologies like non-alcoholic fatty liver and liver cirrhosis (Alkhoury et al., 2014; Morisco et al., 2013), showing changes in the volatilome. The detection and monitoring of a diabetic state has drawn some interest, probably

originally from the fact that acetone is generally found elevated in diabetic groups but highly variable between persons (Španěl et al., 2011; Tassopoulos et al., 1969). In several studies, the identification of gestational, type 1 and type 2 diabetes was possible using not a single but a set of volatiles (Greiter et al., 2010; Halbritter et al., 2012; Novak et al., 2007). There are even efforts to model glucose (Lee et al., 2009; Minh et al., 2011) or plasma triglycerides and fatty acids (Minh et al., 2012) from GC-MS measured VOCs. Therefore, there is potential to use breath analysis to non-invasively diagnose and monitor diseases in the area of energy metabolism despite the variation and confounders found in human studies.

1.5. Limitations in human VOC analysis and the laboratory mouse as model organism

Despite several findings showing interesting associations between disease and the volatilome, there are limitations and questions which are hardly addressable in human studies. Human subjects are exposed to a complex environment contributing to the volatilome, e.g. by the so-called “wash-in” of airborne VOCs, which are distributed in the system and exhaled over time (Beauchamp, 2011). Furthermore, diet and in consequence also microbiota in mouth and gastrointestinal tract do vary in humans and are a source of - (potentially useful) - variation which add to the fact of a diverse genetic background in human populations (Le Chatelier et al., 2013). Depending on methodology, peak identification is still an issue, especially in e-NOSE and PTR-MS studies without confirmation using further - (but probably less-sensitive) - methods. In addition to that, discussions on sampling standardization and mouth or nose breathing are ongoing in the field (Beauchamp et al., 2008; Beauchamp and Pleil, 2013; Thekedar et al., 2009). One way to complement and fill several gaps of knowledge evident from human VOC analysis is the use of model organisms. For the study of volatiles, cell culture and in the last years also more rodent models are discussed and used (Boots et al., 2015). Especially rodents are a promising model for broadening the understanding of volatile biology. The issues of environmental contributions to emitted volatiles can be controlled for multiple ways. First of all, a defined diet can be applied to rodents, and defined diet switches as well as food restriction are possible. In addition, light cycle and housing temperature can be kept constant or, if necessary, be varied according to the question to study. An environmental “wash-in” of volatiles via inhaled air is a lot less variable compared to the human situation. A controlled microbial status is further reducing variance in rodent measurements. And, especially in mice, the genetic homogeneity of inbred strains in combination with the broad tools of genetic manipulation available can address open questions in volatile metabolism.

1.6. Aims of this thesis

In this thesis, the non-invasive method of VOC analysis in un-restrained mice was established (Szymczak et al., 2014) and further developed by using a time-of-flight PTR-MS. This includes establishment of a software routine for processing and analysis of high dimensional time-of-flight mass spectrometric data. Based on this methodological development, further aims could be set and answered in two publications. One aim of this thesis was the characterization of effects a change in dietary matrix from a chow based diet to diets typically used in obesity studies on emitted VOCs. Results on this question are referred to as publication I. Furthermore, the existence of VOCs with diagnostic potential in the fields of energy metabolism and obesity associated diseases as e.g. type II diabetes mellitus is hypothesized and needs to be verified. Results regarding this question are referred to as publication II.

2. Methods

2.1. Housing conditions

All mice were housed at the German Mouse Clinic (Fuchs et al., 2009) in a S1 facility. A room temperature of 24 +/- 1 °C and humidity of 50 - 60 % were maintained. Light regime was a 12:12 hour light-dark cycle (lights on at 6:00 am). Mice were kept in isolated ventilated Type II long cages (Tecniplast Deutschland GmbH, Deutschland). Housing was under specific pathogen free conditions according to FELASA recommendations (Nicklas et al., 2002). Per cage, individual mice (if ordered externally) or groups of 2 – 5 mice were housed on wood shavings as bedding material (LIGNOCEL ¾-S –Fichte, Rettenmaier GmbH, Germany). A plastic house was supplied to all cages as an opportunity to retreat, additionally tissue paper was provided to individually housed mice as nesting material (Mouse Igloo red, Plexx, Netherlands; Kimtech Science Labortücher 7216, Kimberly-Clark Professional, Germany). All animals had ad libitum access to water and chow diet (altromin 1314, Total Pathogenfreies Zucht- und Haltungsfutter, Altromin Spezialfutter GmbH & Co. KG, Germany) if not specified otherwise. All experiments were performed following animal welfare regulations with permission from the district government of Upper Bavaria (Regierung von Oberbayern).

2.2. Animals and diets

2.2.1. Cohorts for effects of diet matrix

For the analysis of diet matrix effects, two cohorts of male C57BL/6J mice (n=40, age 12 weeks) were either ordered from Charles River Laboratories (Germany; Lard groups) or obtained from in-house breeding from C57BL/6J founder animals (tallow/soy based diet groups). Mice had ad libitum access to chow diet from weaning/ arrival on until the VOC analysis; afterwards mice were assigned to four groups with different semi-purified diets (Table 1). The purified diets contained low amount of soy oil versus a high amount of soy oil and tallow fat (LF, HF) or low versus high amount of soy oil and lard (LL, HL). Both LF and HF diets were pelleted (LF, E 15000-04, Ssniff, Germany; HF, E 15741-347, Ssniff, Germany). LL diet was based on a purchased pre-mixture (S5741-E761, Ssniff). HL diet was based on different a purchased pre-mixture (S5741-E762, Ssniff) to match ingredients excluding macronutrients. Lard was molten at 50 °C and added to both lard diet-premixes to achieve 10 en% (LL) or 60 en% (HL) from lard. Before import into the facility, all diets and the respective separate ingredients were sterilized with 25 kGy/min γ -radiation (Isotron Deutschland, Germany). Diets were replaced weekly (pelleted) or twice a week (self-mixed).

Table 1: Diet composition and energy content shown in metabolizable energy percent (calculated using Atwater factors).

Diet	Manufacturer identification	Carbohydrates	Protein	Fat	Energy content [kJ/g]	Used in publication
Laboratory chow (chow)	Altromin 1314	60 en%	27 en%	13 en%	12.5	1, 2
Low soy oil diet (LF)	Ssniff, E 15000-04	66 en%	23 en%	11 en%	15.0	1, 2
High soy oil/tallow fat diet (HF)	Ssnif, E 15741-347	21 en%	19 en%	60 en%	21.4	1, 2
Low soy oil/lard diet (LL)	Self-mixed, based on S5741-E761, Ssnif	74 en%	16 en%	10 en%	15.6	1
High soy oil/lard diet (HL)	Self-mixed, based on S5741-E762, Ssniff	25 en %	15 en%	60 en%	22.3	1

2.2.2. Diet induced obese mice

For the generation of the diet-induced obesity model, 20 male BL6/J mice from in house breeding were randomly assigned to two diet groups. The two mouse groups were switched from chow to either pelleted semi-purified low fat or high fat diet (low fat: E 15000-04; high fat: E 15741-347; both: Ssniff) at the age of twelve weeks to the time of the VOC measurement (24± 2 weeks) on.

2.2.3. Genetic obese mouse model MC4R (W16X)

A melanocortin-4-receptor nonsense allele W16X knock-in mouse line as published by (Bolze et al., 2011) was used to address the question of the impact of (mono-)genetic obesity on emitted VOC pattern. Mice carrying the mc4r knock-in sequence (MC4R-ki) were on C57BL/6J background. 15 knock in - mice as well as 15 controls of both sexes had ad libitum access to drinking water and chow from week 3 on.

2.3. Methods for body mass and body composition monitoring

2.3.1. Body mass monitoring

Body mass of mice was monitored to the nearest 0.1 g weekly and before every breath gas measurement (balance: 440-47N, Kern & Sohn GmbH, Germany). Mice were placed onto an elevated holder to prevent escaping. Balances were calibrated regularly to minimize systemic measurement error.

2.3.2. Nuclear magnetic resonance (NMR) body composition monitoring

The body composition expressed as lean mass and fat mass was determined using a time domain (TD)-NMR scanner (Minispec LF50 body composition analyzer, Bruker Optic GmbH, Germany). Calibration was performed using varying amounts of chicken muscle and pork fat, so that the whole body mass of lean and fat tissue can be measured non-invasively. Mice were weighted before measurement (balance: Kern 440-47N, Kern & Sohn GmbH), placed into a restraining container and measured in the NMR for 2:46 min. Estimation of lean and fat mass in gram was performed by MinispecPlus (Bruker Optic GmbH) und OPUS (OPUS Version 5.0, Bruker Optic GmbH) software.

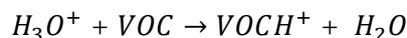
2.4. Analysis of volatile organic compounds in unrestrained mice

2.4.1. Proton transfer mass spectrometry

2.4.1.1. Measurement principle and instrument

A proton transfer reaction mass spectrometer with time of flight detection (PTR-TOF-MS) was used to analyze volatile organic compounds (VOCs) emitted from mice. The PTR-TOF-MS consists of two functional parts, one relevant for ionization of molecules and the other for the detection of those ions. For the chemical ionization as originally described in the 1960s (Munson and Field, 1966), the ion-source hollow cathode charges water to hydronium ions. Hydronium ions as primary ions are most suitable for the analysis of volatile molecules in air, as the most abundant air components have a lower proton affinity (PA) and are therefore not reacting with and consuming primary ions. In contrast most VOCs do have higher PA and can be protonated using hydronium ions. This “soft”, low energy ionization of VOCs is low in fragmentation compared to other methods and therefore easier to interpret in samples with a mixture of VOCs. Unlike other methods, which need a buffer or carrier gas

for transfer of the VOCs, in PTR-MS the air itself is used for that purpose, leading to an outstanding sensitivity down to pptv range (Lindinger et al., 1998). The ionization of volatiles is taking place in the drift tube part of the instrument.



The protonated volatiles are accelerated by an electrical field E from the drift tube through a lens system to detection part of the instrument. The electrical field and especially its ratio with the density of the buffer gas N is furthermore relevant for water cluster formation (low E/N) and fragmentation due to collisions with neutrals (high E/N) of the drifting ions. In the second part of the instrument, the time of flight t of the accelerated ions is detected to determine the mass m of the ion per charge q , which is a recent addition to PTR technology (Herbig et al., 2009). Ions are deflected by voltage pulses from the constant ion stream with a known potential U and a known distance d . A reflector reverses the flight direction of the extracted ions and focuses the kinetic energy distribution of ions to enhance mass resolution. Following reflection, ions are detected in a chevron-configured set of detector plates. After signal multiplication by secondary electrons, the time of the voltage pulse in combination with the extractor time is used to determine the time of flight according to

$$t_{TOF} = \frac{d}{\sqrt{2U}} \sqrt{\frac{m}{q}}$$

The fact that complete spectra can be recorded in the range of milliseconds allows for monitoring of highly dynamic signals and is thus suitable for breath gas analysis (Herbig et al., 2009). Mass resolution is defined as the ratio of mass and the half-maximal width of a peak. In the spectrometer used, a resolution of up to 2000 $m/\Delta m$ can be achieved. This allows differentiation of isobaric compounds unlike in the classical quadrupole PTR-MS while maintaining higher sensitivity compared to PTR-TOF instruments with mass resolution of up to 8000 $m/\Delta m$.

The format in which the spectrometry data is recorded is counts per second [cps]. In order to express these signal intensities as a fraction of the analyzed gas volume e.g. as parts per billion, further conversions have to be made. As the signal intensity is dependent on the specific collision rate constant k and the amount of primary ions H_3O^+ , it needs to be corrected for both parameters as following

$$[VOCH^+] = [H_3O^+]_0(1 - e^{-k[VOC]t}) \approx [H_3O^+]_0[VOC]kt$$

This simplified relationship is fulfilled only for $[VOCH^+] < [H_3O^+]$, thus an excess of primary ions (Lindinger et al., 1998). k is specific for a volatile and will determine the amount of $VOCH^+$ ions which can be produced in the reaction time t . As k -rates are not known for unidentified VOCs, a constant k -rate of $2 \times 10^{-9} \text{ cm}^3/\text{s}$ was used. In addition, raw data needs to be corrected for the transmission, thus the proportion of ions entering the spectrometer later on being detected in the instrument.

2.4.1.2. Operation settings and recording of spectra

PTR settings and measurements were controlled using the software PTR-Manager (Ionicon Analytik GmbH). The ion source was operated with a current of 5 mA. A water flow in the range of 5.0 – 5.5 sccm to supply the ion generation was maintained. Hydronium ions (H_3O^+) were kept in the range of 3000 – 10.000 cps at pk21 (m/z 21.0221, $H_3^{18}O^+$, 1/500 of $H_3^{16}O^+$ which is not measurable due to detector saturation). This was achieved by stepwise increasing the detector voltage between 2100 and 2800 V to compensate detector ageing. A drift tube pressure of 2.3 mbar and voltage of 600 V were maintained, resulting in an E/N ratio of 140 Td. Drift tube temperature was 80 °C. Water cluster formation was kept lower than 10 percent (as monitored by the ratio $H_3O^+ \cdot H_2O / H_3O^+$). Ammonia was kept at approximately 2% but lower than 5%. Protonated oxygen levels were kept lower than 2% of primary hydronium ions. NO^+ levels were minimized to approximately 1% of primary ions. Resolution of the instrument was routinely controlled for peaks 21, 30 and 59 and found to be in the range of 1000 – 2000 $m/\Delta m$. Transmission was determined by measuring a mixture of volatiles with varying masses using a gas calibration unit to set steps of defined concentrations (Transmission gas standard; GCU-a; both: Ionicon Analytik GmbH). A transmission curve was generated to correct for transmission losses of unknown volatiles of known masses over a range of concentrations (Figure 1). Transmission measurements were performed regularly, before and after experimental campaigns. A mass range from m/z 0 to 349.5 was recorded with a repetition rate of 77 kHz and the sum spectra integrated over 3 s stored (TOF-DAQ, Tofwerk AG, Switzerland). Spectra were calibrated using the known peaks NO^+ (m/z 29.9971) and protonated acetone (m/z 59.0491, $C_3H_6O \cdot H^+$) in publication I and $H_3^{18}O^+$ (m/z 21.0221), NO^+ and protonated acetone in publication II. In publication I, a manual selection of peaks using TOFViewer (Version 3.0.1, Ionicon Analytik GmbH, Austria) yielded 164 peaks. 306 peaks were selected manually from spectra in publication II using

PTR-MS Viewer (Version 3.2.6, Ionicon Analytic GmbH). Concentrations were calculated from cps data using PTR-MS Viewer ($k=2 \times 10^{-9} \text{ cm}^3/\text{s}$, transmission curves see Figure 1).

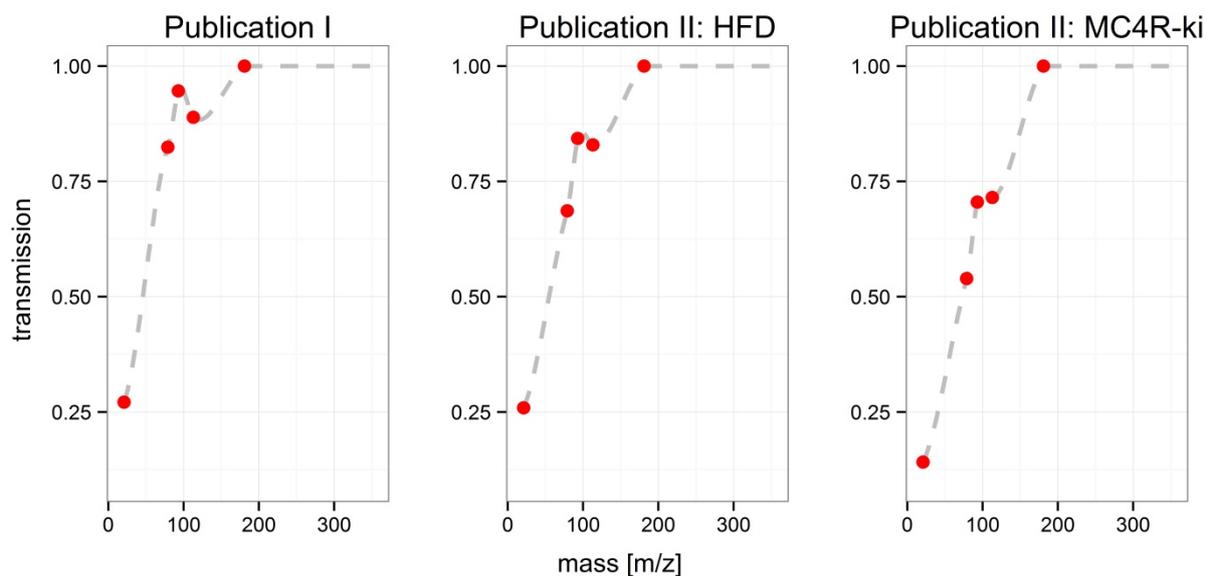


Figure 1: Transmission curves determined for publications I and II (individually for cohorts HFD and MC4R)

2.4.2. A setup to measure volatile organic compounds in mice

The measurement setup for analysis of volatile organic compounds emitted from unrestrained and unanaesthetized mice is used as previously published (Szymczak et al., 2014). A polypropylene box with a volume of 600 ml was utilized as respiratory chamber (Lock&Lock GmbH, Germany). The chamber was connected to a supply tube for synthetic air (20 % oxygen, 80 % nitrogen, concentration of hydrocarbons < 0.1 ppm, Linde AG, Germany) on one end and to the PTR-MS from the top cover (constant flow 60 mL min^{-1} controlled by PTR). A PTFE membrane filter (pore size 2.0 μm , PALL Corporation, USA) was installed between chamber and PTR-MS to avoid particular contaminations. Connections to the PTR-MS were heated to avoid condensation of humid breath and trapping of VOCs (before filter > 40 °C, after filter 80 °C). The chamber itself was not actively temperature controlled; the repeated flushing of the chamber and continuous diluting of the headspace in the presence of a mouse kept the temperature within the chamber roughly at ambient temperature (20 – 25 °C) comparable to home cage conditions. The supply of synthetic air could be switched by a PTFE-coated valve (Series 1 &2, Parker, USA) between a PTFE air reservoir (Welch Fluorocarbon Inc., USA) during measurement or a flow of 3 l min^{-1} during flushing of the chamber. Custom made feedthroughs out of PEEK were deployed at both the gas inlet and outlet in the chamber. The

feedthroughs were designed to split into 6 paths rectangular to original gas flow. This provided an indirect air supply during flushing the chamber with higher flow rates at the inlet, thus lower disturbance of the mouse in comparison to a direct air jet. In addition, the multiple-path outlet reduced concentration fluctuations due to direct sniffing at the feed-through by the mouse. A one-way valve (Bürkle GmbH, Germany) was mounted to the top to prevent over-pressure during high-flow flushing of the chamber. Flow meters for both measurement and flushing flow (TD9411M, Aalborg, USA) were included and monitored for drops indicating system leakage. Valves for controlling gas flow through the system are operated using an USB-relay (ABACOM Ingenieurbüro GbR, Germany) and custom made software interface (based on ProfiLab, ABACOM Ingenieurbüro GbR).

2.4.3. Measurement protocol and contamination monitoring

The basal setup and modifications have been described in (Kistler et al., 2016, 2014; Szymczak et al., 2014). Before the measurement, mice were placed in an acclimatization chamber for several minutes. This chamber is equal to the measurement chamber with a supply of laboratory room air. In addition, a soft tissue paper (Kimtech Science, Kimberly-Clark) was placed in the chamber to reduce fur contamination in case of urine or feces events. On top of the chamber, a cover was placed over half the chamber to allow the mouse to retreat and reduce potential stress of being restrained in the transparent box.

A measurement chamber is connected to the mass spectrometer and flushed 2 minutes with a flow rate 3 l/min to dilute enclosed laboratory room air. The recording of mass spectrometry data is started right after start of flushing. After flushing, VOCs from the empty respiratory chamber are measured (5 min, flow 60 mL min⁻¹) to detect system leakage and background VOC concentrations and being able to correct for them. Signals monitored for leakage from laboratory air are acetone (m/z = 59.05) and propanol (m/z = 41.06) concentration, which are typically several orders of magnitude higher concentrated compared to synthetic air. Following this blank measurement, the system is switched to flushing state and the mouse is placed into the chamber. Entered laboratory air is replaced by flushing 2 min. After that, measurement phases alternate with flushing of the chamber, allowing the volatiles to accumulate in the gaseous phase. Generally, three or more valid accumulation phases are measured per mouse.

During accumulation of VOCs in the headspace of unrestrained and non-anaesthetized mice, contaminations can arise e.g. from urination and defecation and add variance and bias to the data. Thus, mice were cleaned prior to the measurement to remove fur contaminations if necessary. The

measurement chamber was monitored for macroscopic signs of contamination repeatedly during measurements. In addition, several volatiles were used as marker substances for urine or feces events. Urine was detected due to sudden changes in humidity (determined as water-cluster $(\text{H}_2^{18}\text{O})_2\text{H}^+$, $m/z = 39.05$), concentration of trimethylamine ($m/z = 60.07$) and pk127B (tentatively dimethyl trisulfide, $m/z = 127.02$). Concentration of methanethiol ($m/z = 49.02$) indicated presence of feces. In case of feces contamination, feces were removed and accumulation phase was skipped. If urine was present, the respirometry chamber was replaced, the mouse was gently cleaned using soft tissue paper (Kimtech Science, Kimberly-Clark) and the measurement was restarted.

2.5. Primary data analysis

2.5.1. Calculation of source strengths

In the respiratory chamber, the emission of volatiles from the mouse and the dilution due to gas flow to the PTR-MS and replacement of this volume by VOC-reduced synthetic air contribute to a measured count rate per second [cps]. An equilibrium will be reached after more than 15 minutes at a flow of 90 ml min^{-1} and will vary with background and starting concentration of volatiles (Szymczak et al., 2014). Therefore, the quantification of VOC emission was based on dynamic changes during the saturation curve using a compartment model. The concentration of certain VOC c_i at time t was described with the following equation (Szymczak et al., 2014):

$$c_i(t) = c_0 * e^{-\frac{F}{V}t} + \frac{S_i}{F} \left(1 - e^{-\frac{F}{V}t}\right)$$

Herein, the flow F in the chamber as well as the volume V of the chamber are known and kept constant. c_0 represents the starting concentration of the volatile, which is kept minimal by flushing with synthetic air. S_i , which was termed source strength, can then be derived from a nonlinear least-square fitting and utilized to quantify the emissions of specific VOCs in exhaled breath.

2.5.2. Profile selection and quality control

For publication 1, a custom matlab script was used to detect blank measurements and saturation profiles from PTR-MS concentration data in a fully automated way (Matlab Version R2012a, The Mathworks Inc., MA, USA). Herein, signals from multiple VOCs (ethanol, acetone, water clusters at pk39 and pk55) were used to determine local minima and maxima of the data as well as trend-

breaking events in source strengths. These events were used to segment saturation profiles. Nonlinear regression fitting with both gaussian and poisson noise modeling was then performed in the segments. Goodness of fit was evaluated by comparison of mean and variance of both fits. Identification and flagging of contaminations and disturbances of the measurements were automated. This included urine and feces contamination, box opening and time scale misalignment (“dejavu” data). Data was carefully checked manually and discarded if false positive flagging and miss-segmentation of profiles was present. For urine and feces, measurements with a source strength higher than $S_{i(pk127B)} > 300$ ppb mL min⁻¹ and $S_{i(pk49)} > 100$ ppb mL min⁻¹ thresholds were excluded.

As part of this thesis project, a custom server-based application based on R and the shiny package (Chang et al., 2015; R Core Team, 2015) was developed and applied on PTR-MS concentration data in publication II (Figure 2A). Compared to the previous data analysis protocol, this web-application has several improvements. Concentration is visualized over time instantly after upload of data and according to a selectable chosen signal, blank measurements and segments can be selected manually after automatic pre-detection. For the analysis acetone and propanol peaks were used for profile selection.

Furthermore, profiles which need to be segmented due to urine or feces events can be chosen manually. Thus, no discarding of information is necessary if a single segmentation criterion cannot achieve correct segmentation. For the evaluation of contamination events, visualization of marker volatiles directly next to other VOCs is possible. In addition, protocolled events can be read in and displayed right next to the data allowing correct and reproducible evaluation of measurements. In publication II, concentrations greater 1 ppb and sudden increases in source strengths of methanethiol ($m/z = 49.02$), trimethylamine ($m/z = 60.07$) and pk127B ($m/z = 127.02$) were used to discard profiles.

For the analysis of source strengths, nonlinear regression fitting with gaussian noise modeling was then performed in the segments and shown instantly right next to the data. Additionally, further models e.g. without noise modelling can be calculated simultaneously and overlayed on the same plot. Using the AIC criterion, goodness of fit of the models can be compared directly. In publication II, nonlinear regression fitting with gaussian noise modeling was applied.

Unlike previous static reports from matlab scripts, data can be zoomed and segmented in real-time during analysis. This is especially useful in prolonged experiments e.g. glucose tolerance testing or to visualize VOC correlations within a certain profile (Figure 2B). Maximum and minimum data thresholds can be set to exclude e.g. high propanol levels from room air while inserting the mouse.

For reproducible and transparent data analysis, metadata can be entered and are exported with the calculated source strengths. Furthermore, first capabilities to analyze data on cohort level (e.g. plotting, mixed effects modelling) and export a reduced dataset for MausDB upload have been implemented.

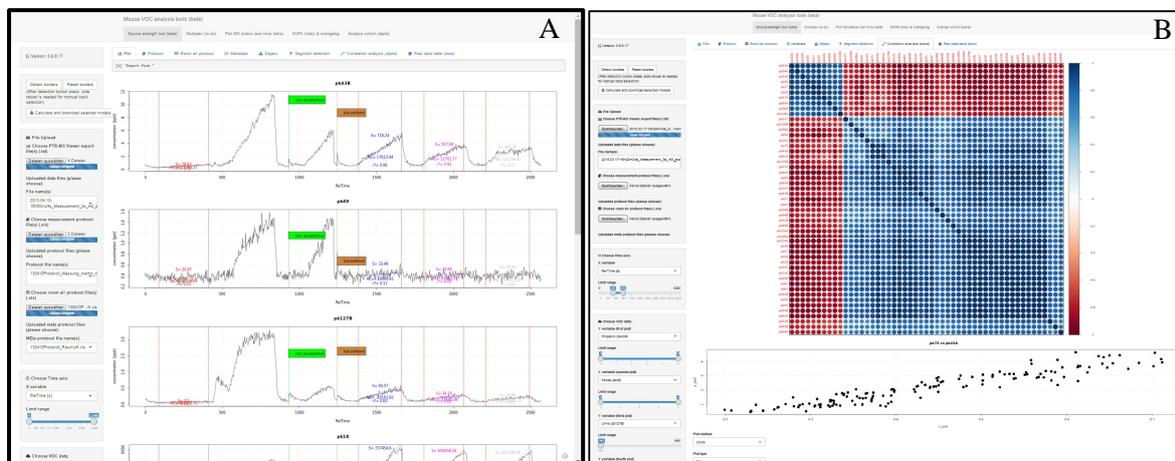


Figure 2: Web application for processing and evaluating PTR-MS concentration data from mouse headspace measurements. A: Data analysis tool with source strength calculation and protocol event visualization. B: Correlation matrix plot with visualization of individual peak correlations from selected data segments.

In both publications, VOC source strengths measured during blanks were subtracted from VOC source strengths of mouse accumulation profiles. This was performed to ensure no undetected micro-leakage in respiratory chambers could affect mouse source strengths. In addition, peaks which had a non-different or smaller source strength during measurement compared to blank source strength were removed from data set in publication II. Negative source strength was determined using differences of mean source strengths. Significant increases in source strengths were determined using linear regression modelling. A $p < 0.1$ threshold was set to ensure that potential very low signal marker volatiles were not filtered out.

2.6. Secondary data analysis

2.6.1. Machine learning protocol for feature selection

The analysis of volatiles from gaseous phases using PTR-MS results in data sets of with a high number of variables, typically several hundreds and therefore typically around 10 times more variables p than individual mice n are part of the experiment. The arising issues regarding the correction for multiple hypothesis testing, especially in a screening approach, is therefore commonly found in methods dealing with $p \gg n$ datasets like metabolomics or transcriptomics (Lay Jr. et al., 2006). One approach in narrowing down the number of candidate volatiles is applying machine learning algorithms to rank importance for feature selection.

In publication I, a decision tree and a bagging based algorithm, namely random forest ++ (RF++, (Karpievitch et al., 2009)), was applied. In this algorithm, subject-based bootstrapping allows to extend the power of random forest by circumventing using means over repeated-measures data. For the classes chow and semi-purified, baseline and one week data was analyzed (5000 trees, 13 splitting variables per tree). The top three independent VOCs were further analyzed.

In publication II, another supervised feature selection approach, namely the AUC-RF algorithm, was implemented (Urrea and Calle, 2012). Here, a receiver operating characteristic curve and ranking of predicting volatiles are computed from an initial random forest. Less important variables are then removed in an elimination process until an optimal ROC-AUC is found. For both HFD-fed and MC4R-ki datasets strata were set to allow only one measurement per mouse and fasting status in every decision tree. As over-fitting of the resulting model and reproducibility are possible in a single model, a five-fold cross validation was applied with 20 iterations. Only peaks with a selection probability of more than 70% in the cross-validation were further analyzed. In addition, the complete analytical pipeline was repeated using randomly permuted class labels to ensure robustness of the classification approach.

2.6.2. Data visualization

Source strength data was visualized using r package ggplot2 with default settings (Wickham, 2009). The top of each box in the boxplots indicates the 75% percentile, the bottom the 25% percentile, and the thick bar inside the box is the median. Whiskers are defined as 25% percentile – 1.5 * Interquartile

range and 75% percentile + 1.5 * Interquartile range, respectively. Data points higher or lower than Whiskers are defined as outliers.

For both heatmaps and gaussian graphical modelling in publication II, missing values were imputed using chained equations (mice R package (Buuren and Groothuis-Oudshoorn, 2011)).

Heatmaps in publication 2 were performed using the Heatplus (Ploner, 2014) package from the Bioconductor project (Gentleman et al., 2004). Mean ad libitum fed as well as mean fasted source strength data per mouse was used and shown individually. Data has been scaled and centered. Annotation according to genotype, diet, fasting status and mouse weight was performed on the right side. Hierarchical clustering was performed on both volatiles (columns) and mice in fasted and ad libitum state (rows), sub clusters of mice were color-coded.

Recently, gaussian graphical models were shown to reconstruct metabolic pathways without a-priori information (Krumsiek et al., 2012, 2011). These models rely on a partial correlation matrix, in which every two variables correlation is corrected for contribution of all other variables from the dataset. For the creation of a gaussian graphical model, source strengths were log-transformed to achieve normality. A shrinkage approach to estimate a partial correlation matrix was applied as this is $n < p$ data set (Schäfer and Strimmer, 2005). Additionally, the data was of longitudinal structure, thus a dynamic partial correlation was estimated (Opgen-Rhein and Strimmer, 2006) using the GeneNet R package (Schaefer et al., 2015). A local FDR of 3% was used to extract a network from the partial correlation matrix. As two cohorts of mice were measured individually, a “dummy” variable correcting inter-experimental differences was introduced but not plotted. The percentaged coefficients from linear mixed effects models are shown in the nodes as a pie-chart for every peak with significant intervention or fasting effects. Subnetworks of significant peaks and direct positive correlated neighbors were highlighted. Top 20% of connections are shown with bold lines; minor 20% with grey lines, negative partial correlations with dotted lines. Peaks identified in feature selection but without significant connections were included in the graphical model for illustration purposes.

2.6.3. Statistical testing

Linear mixed effects models with repeated measures as random effects were applied to test for intervention effects (publication I: diet matrix & diet quality; publication II: fasting state & diet or fasting state and genotype). This was performed using the nlme R package (Pinheiro et al., 2015). Source strength data was logarithmized to achieve normality (as tested visually in qq-plotting). In

publication I, models were reduced using AIC to remove irrelevant parameters (stepAIC function, R). In publication II, mixed effects models were maintained. In addition, post-hoc test for individual groups were performed on significant interaction. Control of false discovery rate after Benjamini and Hochberg (Benjamini and Hochberg, 1995) was applied to avoid cumulation of the Type I – Error. All p-values were adjusted according to a false discovery rate of 10% in publication II.

3. Results

3.1. Publication I

Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice

M. Kistler, W. Szymczak, M. Fedrigo, J. Fiamoncini, V. Höllriegl, C. Hoeschen, M. Klingenspor, M. Hrabě de Angelis, J. Rozman

Journal of Breath Research, Volume 8, Number 1 (2014)

Abstract

Breath gas analysis in humans proved successful in identifying disease states and assessing metabolic functions in a non-invasive way. While many studies report diagnostic capability using volatile organic compounds (VOC) in breath, the inter-individual variability even in healthy human cohorts is rather large and not completely understood in its biochemical origin. Laboratory mice are the predominant animal model system for human disorders and are analyzed under highly standardized and controlled conditions. We established a novel setup to monitor VOCs as biomarkers for disease in the breath gas of non-anesthetized, non-restrained mice using a proton transfer reaction mass spectrometer with time of flight detection. In this study, we implemented breath gas analysis in a dietary intervention study in C57BL/6J mice with the aim to assess the variability in VOC signatures due to a change in the diet matrix. Mice were fed a standard laboratory chow and then exposed to four semi-purified low- or high-fat diets for four weeks. Random forest (RF++) was used to identify VOCs that specifically respond to the diet matrix change. Interestingly, we found that the change from a chow diet to semi-purified diets resulted in a considerable drop of several VOC levels (*Figure 3*). Our results suggest that the diet matrix impacts VOC signatures and the underlying metabolic functions and may be one source of variability in exhaled volatiles.

Key figures

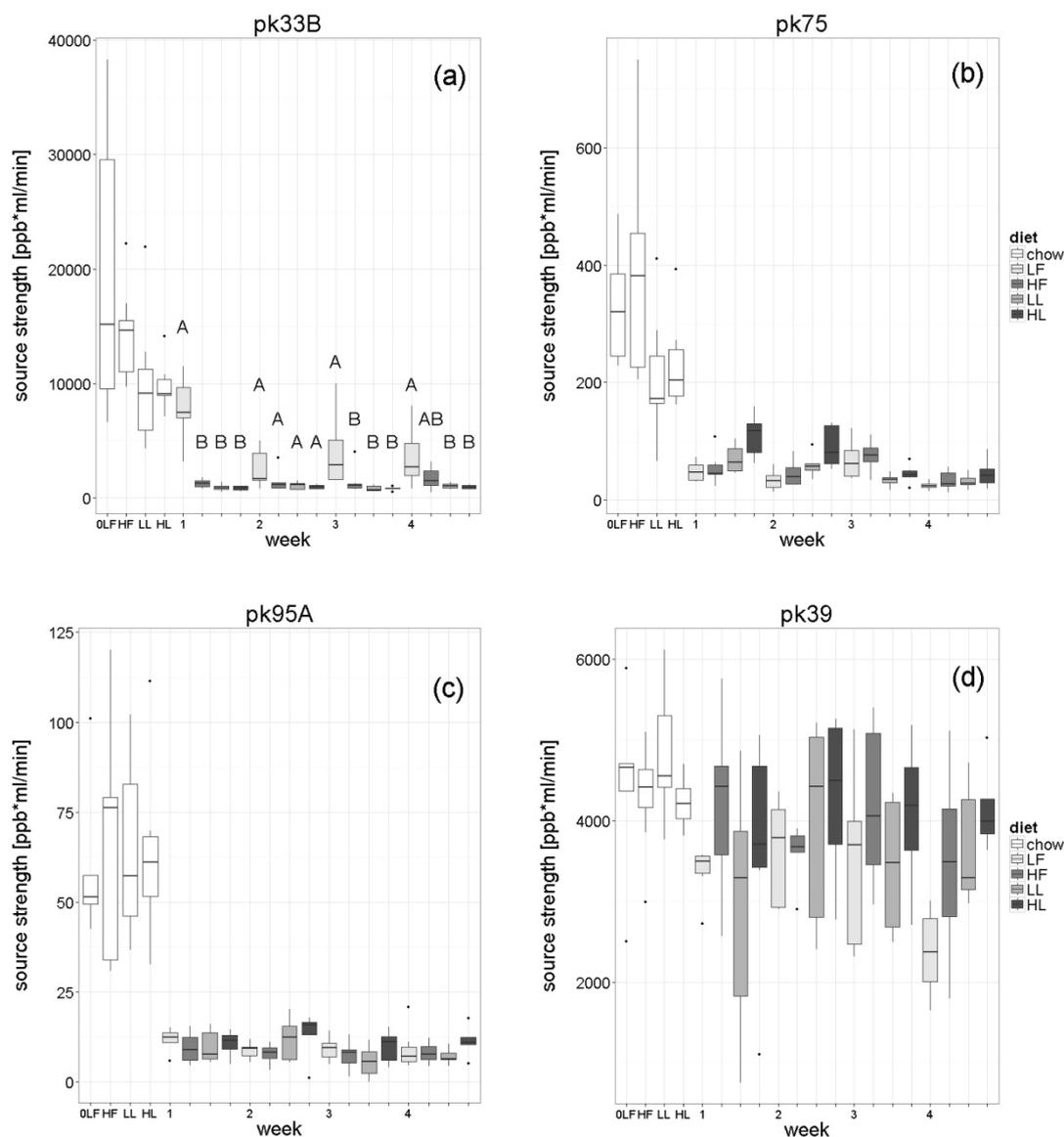


Figure 3: Effects of a switch from chow (week 0) to purified diets (weeks 1–4) on emitted breath VOCs (a)–(c) are shown (5–10 mice per group and date). If an interaction in diet quantity and quality was detected, different capital letters were used to show significant differences between diet groups tested with pairwise Student’s t-test at a time point (significance level: $p < 0.05$). For pk95A, no differences at a time point were found in pairwise t-testing. Water isotope cluster pk39 (d) is shown as a measure for humidity to estimate activity and breath frequency.

Candidates own contribution:

Idea of the manuscript, together with supervisor J. Rozman

Design and execution of experiments

Processing and analyzing all data in R statistics

Interpretation of data and drafting manuscript

Writing and finalizing paper in consultation with co-authors

3.2. Publication II

Diet-induced and mono-genetic obesity alter volatile organic compound signature in mice

Kistler M., Muntean A., Szymczak W., Rink N., Fuchs H., Gailus-Durner V., Wurst W., Hoeschen C., Klingenspor M., Hrabě de Angelis M., Rozman J.

Journal of Breath Research, Volume 10, Number 1 (2016)

Abstract

The prevalence of obesity is still rising in many countries of the world resulting in an increased risk for associated metabolic diseases. In this study we aimed to describe the volatile organic compound (VOC) patterns symptomatic for obesity. We analyzed high fat diet-induced obese (HFD) and mono-genetic obese mice (global knock-in mutation in melanocortin-4 receptor MC4R-ki). The source strengths of 208 VOCs were analyzed in ad libitum fed mice and after overnight food restriction. Volatiles relevant for a random forest-based separation of obese mice were detected (26 in MC4R-ki, 22 in HFD mice). Eight volatiles were found to be important in both obesity models. Interestingly, by creating a partial correlation network of the volatile metabolites, the chemical and metabolic origins of several volatiles were identified. HFD-induced obese mice showed an elevation in the ketone body acetone and acrolein, a marker of lipid peroxidation, and several unidentified volatiles. In MC4R-ki mice, several yet-unidentified VOCs were found to be altered. Remarkably, the pheromone (methylthio)methanethiol was found to be reduced, linking metabolic dysfunction and reproduction. The signature of volatile metabolites can be instrumental to identify and monitor metabolic disease states, as shown in this screening of two obese mouse models. Our findings show the potential of breath gas analysis to non-invasively assess metabolic alterations for personalized diagnosis.

Key figures

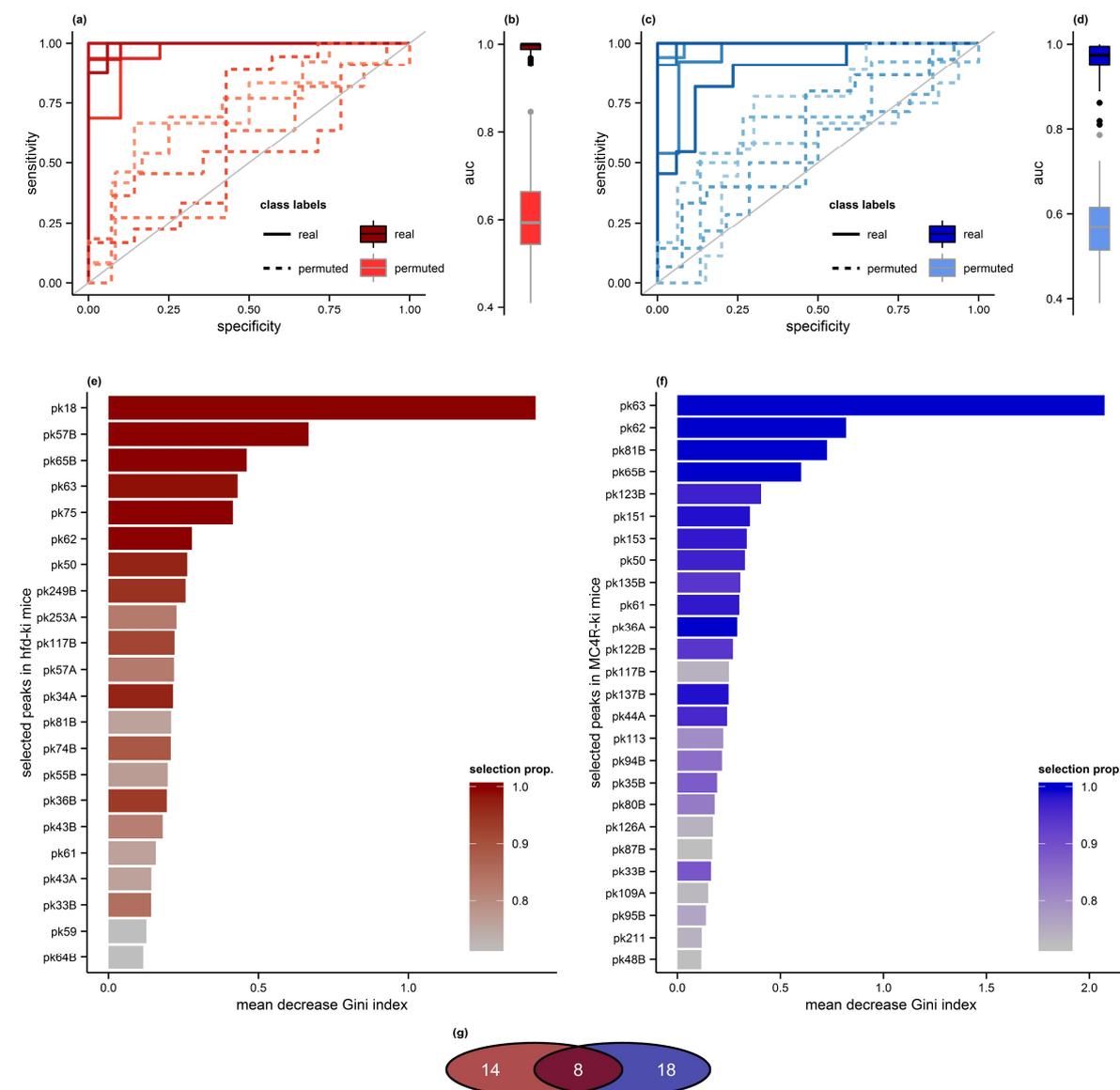


Figure 4: Machine learning strategy for selection of obesity relevant peaks. The performance of classification models was evaluated using ROC curves. For high fat diet fed mice (HFD, (a)) and melanocortin-4-receptor W16X knock-in mice (MC4R-ki, (c)), ROC curves are shown for a single 5-fold cross-validation of the recursive feature selection using AUC-RF algorithm. Results for real class labels are shown in dark solid lines. In comparison, light dashed lines represent ROC curves from using randomly permuted class labels in the classification procedure. The variation in area under the curves (AUCs) under ROC curves is shown in boxplots (HFD (b), MC4R-ki (d)). For this, 20 different 5-fold cross-validation sets were analysed. Dark boxplot fills represent real class labels while light colors represent permuted labels. For further analysis, peaks with more than 70% selection probability in the repeated cross-validation procedure were selected. The variable

importance of selected peaks is shown (HFD, (e); MC4R-ki, (f)). Colour gradients indicate selection probability after 20 iterations of a five-fold cross validation procedure. Overlapping of selected peaks is shown as Venn diagram (g).

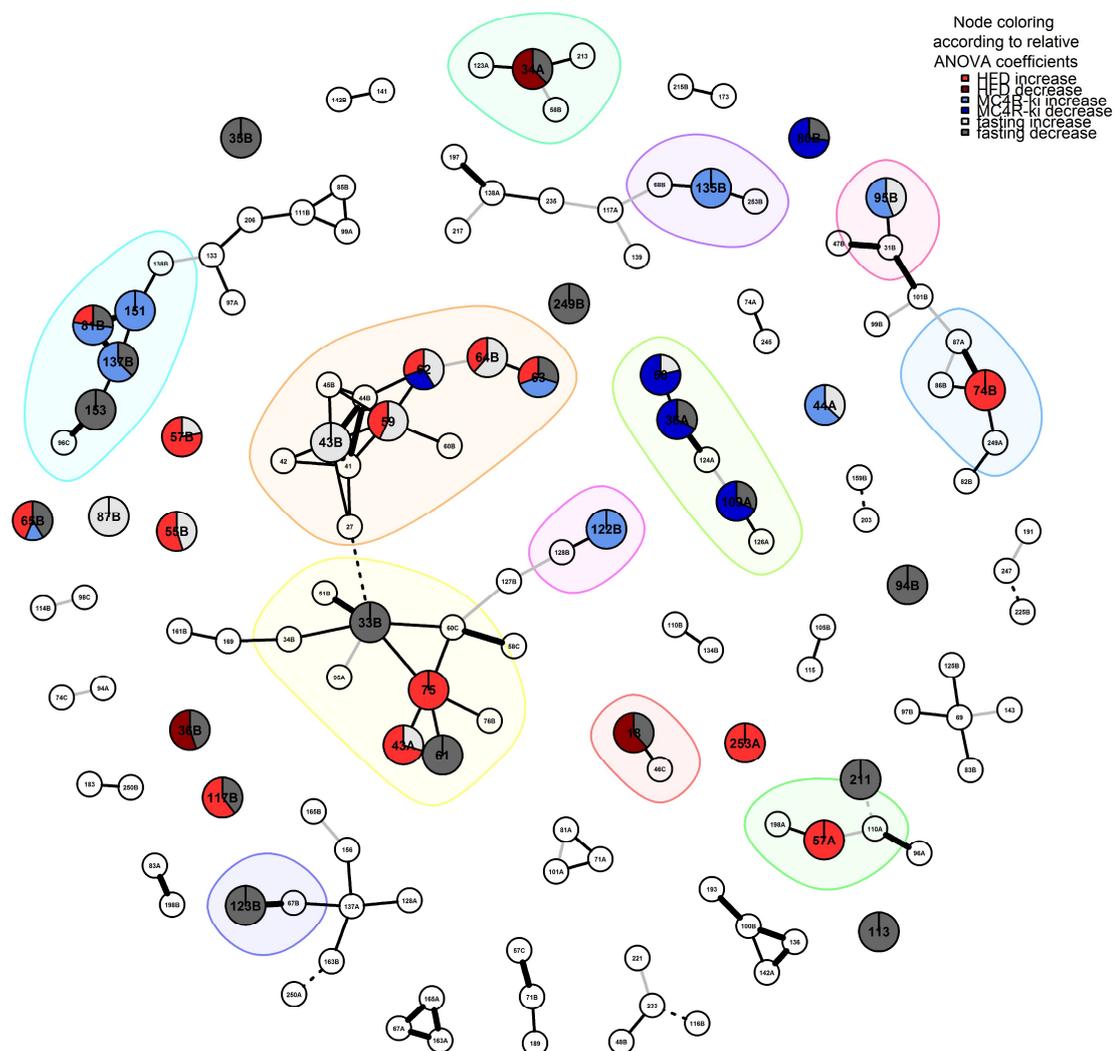


Figure 5: Gaussian graphical model for VOC identification. Gaussian graphical model with nodes corresponding to peaks (labelled with nominal mass and letter for multiple peaks at the same nominal mass) and edges corresponding to shrinkage estimated partial correlation. Highest and lowest 20 percent of correlations are highlighted (bold black/ thin grey). Dotted edges indicate negative partial correlation. Peaks with significant mixed effects model main effects (as seen in Fig. 4 and Fig. 5) are shown as pie charts in nodes. Coloring of model coefficients is according to diet increase/ decrease (red/ dark red), genotype increase/ decrease (blue/ dark blue) and fasting increase/ decrease (grey/ dark grey). Mean fasting coefficients are shown if significant in both obesity models. Subnetworks of peaks with significant effects and directly connected nodes with positive partial correlation are plotted on colored background. Peaks without significant edges selected in AUC-RF were added for illustration purposes.

Candidates own contribution:

Idea of the manuscript, together with supervisor J. Rozman

Design and execution of experiments

Processing and analyzing all data in R statistics

Interpretation of data and drafting manuscript

Writing and finalizing paper in consultation with co-authors

4. Discussion

Fundamental for the analysis of volatile organic compounds emitted from unrestrained and non-anaesthetized mice was the establishment of the methodology (Szymczak et al., 2014). Herein, VOCs from mice are measured using a small box with supply of synthetic air and a PTR-Quad-MS. The obtained saturation curves are described by applying a compartment model with the assumption of constant source strength of a particular VOC during a curve. In a prove-of-principle experiment, source strengths of several volatiles were affected by changing from a chow to a purified high-fat diet feeding. However, this diet switch consists of a variety of changed experimental parameters, namely diet energy content, diet matrix (chow versus purified), progression in age and obesity of mice. To dissect and answer those questions individually, additional experiments were performed and published. These experiments resulted in a novel catalogue of candidate VOCs associated with diet matrix changes or obesity and will be discussed in the following sections.

4.1. The effect of diet matrix on the emitted VOC pattern

4.1.1. Change from chow to semi-purified diets

In obesity research, a change from a grain-based chow diet to breed and maintain mice to a semi-purified or synthetic experimental diet is a common procedure. However, the effects of the diet-matrix change per-se are not always considered. Typically, such a shift towards a purified diet comparable in macronutrient composition leads to various physiological effects, such as microbial change (Daniel et al., 2014; Sonnenburg et al., 2016) or higher body weight gain in rats on a synthetic diet (Moraal et al., 2012). In addition, the generally elevated but also highly variable content of phytoestrogens like isoflavones are known to be present in chow diets. For example, for Altromin 1324 isoflavone levels vary from ~ 220 to 530 $\mu\text{g/g}$ chow (Jensen and Ritskes-Hoitinga, 2007) and can lead to activation of hepatic phase I and II metabolism (Rudolf et al., 2008).

Due to such profound alterations it was hypothesized that the VOC emission pattern could be altered as well. Therefore, the shift from a chow diet to four different semi-purified diets was analyzed in respect to emitted volatiles. Interestingly, a massive drop in three volatiles was observed, which were tentatively identified as methanol, methyl acetate or propionate and dimethyl sulfone.

4.1.2. Methanol (pk33B, pk15, pk34B, pk35B, pk51B)

In chow fed mice, high amounts of methanol in breath are present, which are massively reduced on semi-purified diets. The presence of methanol in human breath was detected in the 1960s (Eriksen and Kulkarni, 1963) and can originate from various precursors. One major source is the consumption of pectin and microbial cleavage of methyl groups. In addition, other sources like aspartame, demethylation of proteins and S-Adenosylmethionine can contribute to methanol levels (Axelrod and Daly, 1965; Dorokhov et al., 2012; Lindinger et al., 1997; Morin and Liss, 1973; Siragusa et al., 1988). Pectic polysaccharides from soy, which is a major ingredient in chow diets, might be the dominant source of systemic methanol in chow fed mice. Interestingly, not only a switch from chow to a semi-purified diets but also overnight food restriction leads to massive drop in methanol as shown in chow fed MC4R-Ki mice. This might be an effect of retained pectin supply to microbiota and indicates that breath methanol is massively and acutely influenced by diet supply.

Notably, mice on LFD remained higher compared to other semi-purified diets for the first four weeks after diet change. According to the manufacturer, raw fiber content is 5.0% compared to 6.0% of the corresponding HFD and 5.0% cellulose in lard-based diets. It is likely that cellulose will not be metabolized to methanol but to short-chain fatty acids by the gut microbiota (Peng et al., 2013). Comparing LFD and HFD, fiber composition might vary and a higher content of pectins might be present in LFD. Another possibility is that the HFD induced modification of the gut microbiota is interacting with pectin metabolism and leading to lower methanol release.

4.1.3. Methyl acetate (pk75, 76B)

Another finding in publication I was the massive drop in a peak of the mass 75.02 (pk75), tentatively identified as either methyl acetate or propionate. Interestingly, after fasting challenge in chow fed MC4R-ki mice in manuscript II, pk75 was also reduced. In an approach to reconstruct biochemical pathways as well as fragmentation, isotopes and chemical reactions, a gaussian graphical model indicated methanol and acetic acid with a significant partial correlation. Thus, the assumption that methyl acetate is dominantly contributing to pk75 is further substantiated. The origin of methyl acetate in breath can be endogenously, as shown e.g. in human bronchial epithelial primary cells in vitro (Filipiak et al., 2010). As shown for longer lipid esters, it could possibly be synthesized out of methanol and acetate by liver, duodenal mucosa, and pancreas (Diczfalusy et al., 2001). In addition, the gut microbial production of methyl acetate has been shown and can be modulated by uptake of symbiotic food (fructooligosaccharides and the probiotic strains) as measured in feces (Vitali et al.,

2010). As indicated by the graphical model, breath levels of methyl acetate might be driven by a synthesis from high methanol levels, either endogenously using acetate or acetyl-CoA or by the microbiota using acetate out of fiber fermentation. Thus, upon lack of methanol supply, both by diet matrix change or food restriction, methyl acetate levels decrease.

In humans, breath methyl acetate is acutely responsive to exercise, with a ~2-5 fold rise and plateauing after 5 minutes. This exercise-induced rise in concentration is closely accompanied by elevations in the cardiac output or blood pressure (King et al., 2010). Regarding the setup of unrestrained and non-anaesthetized mice, varying locomotor activity could modify methyl acetate source strengths. This might be especially true for mice within a measurement, beginning with an active exploring phase and calming down after several repetitions. However, as shown in manuscript I, humidity in the box is not affected, which is assumed to be driven by increased activity (Szymczak et al., 2014). Hence, acute locomotive induced changes in cardiac output are unlikely to contribute relevantly to the drop of methyl acetate levels after diet switch in this mouse setup using a small respiratory chamber.

4.1.4. Dimethyl sulfone (pk95A)

Another volatile compound affected by the diet matrix change was found at a mass of 94.97. In rodent breath, three candidate VOCs with a nominal mass of m/z 95 have been described, namely dimethyl sulfone (DMSO₂, also methylsulfonylmethane), phenol and substances with the sum formula C₇H₁₀ (Aprea et al., 2012). According to the exact mass of 94.97, pk95A was tentatively assigned to DMSO₂ ($m/z = 94.998479$). Interestingly, in publication I again a massive drop in source strength was identified when switching from chow to semi-purified diets. The same was true after overnight fasting in chow fed MC4R cohort (data not shown in publication II/ see Supplementary figure 1). DMSO₂ is known to be widely available in human diet, especially from vegetables, fruits, beverages and grains (Buško et al., 2010; Pearson et al., 1981). Hence, a direct contribution from the grain-based diet is a likely origin. Another suggested source of DMSO₂ is the microbial degradation of methionine, which is metabolized via methanethiol to dimethyl sulfide and further to DMSO₂ by the host (He and Slupsky, 2014). However, both methanethiol (m/z 49.01) and dimethyl sulfide (m/z 63.03) are volatile and could be detected in breath if present systemically. Although it is possible that oxidation of both precursors is extremely efficient and no alterations in systemic and breath concentrations are measureable, this seems rather unlikely. In conclusion, in case of DMSO₂ the majority of the determined source strengths can likely be attributed to a direct dietary effect.

Notably, as emitted DMSO₂ levels vary massively between chow and semi-purified diets, it is of high importance to specify dietary composition in models of obesity and choose an appropriate control. DMSO₂ has a history of use as a diet supplement for its anti-oxidative, anti-inflammatory and anti-apoptotic properties (Amirshahrokhi and Bohlooli, 2013; Gregory et al., 2008; Kamel and Morsy, 2013). These effects might be mediated by inhibition of the NLRP3 inflammasome activation, thus attenuating the transcriptional expression of IL-1 α , IL-1 β , IL-6 and reducing production of mitochondrial reactive oxygen species (ROS) (Ahn et al., 2015). Therefore, the level of DMSO₂ in breath is an interesting candidate to monitor for information on the oxidative state, which might modulate obesity associated pathologies like a diabetic phenotype. Interestingly, DMSO₂ was found to be reduced in rats with NASH by feeding a HFD compared to a LFD (Aprea et al., 2012). However, the control diet was referred to as “typical pellet diet”, but not clarified if chow or purified. Thus it is possible that the reduction might at least partly reflect the change in diet matrix and grain-derived dietary DMSO₂. This is further substantiated as in HFD fed mice no such reduction in DMSO₂ was observed when compared to a semi-purified LFD (Kistler et al., 2016).

4.2. Obesity-induced changes in the volatilome

In publication II, the aim was to identify common and model-specific changes in the emission of VOC using two obesity mouse models. Therefore, both a model of diet induced obesity as well as monogenetic obesity was analyzed.

4.2.1. A common VOC signature of obesity

4.2.1.1. Acetic acid (pk61 and pk43A)

Acetic acid was identified in both obesity models as potentially relevant by random forest models. Although adding fasting as a factor to the mixed effects modelling only the acetic acid fragment pk43A in HFD fed mice is significant, acetic acid still seems an interesting volatile for metabolic readouts. Acetic acid is central in energy metabolism via its metabolite acetyl-CoA which can be derived from carbohydrates, amino acids and fatty acids. The enzymes acetyl-CoA synthetase and acetyl-CoA hydrolase can regulate free acetic acid levels, in addition exogenous sources like gut fiber fermentation contribute to serum levels predominantly after food intake (Wolever et al., 1997). Serum acetic acid levels were reported to be dependent on glucose tolerance status (Wolever et al., 1997) and inversely correlated to insulin levels in mice and humans (Layden et al., 2012; Sakakibara et al.,

2009). Glucose-induced insulin secretion from rat pancreas is attenuated by acetic acid (Tiengo et al., 1981). This inhibition is mediated by pancreatic FFAR2 and FFAR3. These receptors can be activated both by overall systemic acetic acid levels as well as by pancreas secreted acetate, which is produced from glucose as a negative feedback (Tang et al., 2015).

However, effects of HFD on systemic acetic acid levels in mice have been reported to be either elevated or decreased. These conflicting results are probably due to the fact that dietary fiber content was not reported or not accounted for. Lower serum levels in HFD fed mice without fiber were measured compared to chow diet (Shearer et al., 2008). In addition, mice fed an HFD with unspecified fiber content have lower portal acetate levels compared to chow fed mice, which can be restored after antibiotic intervention (Carvalho et al., 2012). By feeding a HFD with 5.0% fiber content, a higher systemic level of acetic acid has been found in obese mice compared to control chow (Tang et al., 2015). This finding of elevated acetate is reproduced in a D12492 HFD with 6.5% fiber content compared to standard chow (McNelis et al., 2015). The later findings match the trend of increase in HFD fed mice within this thesis (D12492 mod., 5.0% fiber content).

Notably, hepatic release of acetic acid was reported to be essential for energy supply to extra-hepatic tissues in rats and mice, especially in ketogenic state (Sakakibara et al., 2009; Yamashita et al., 2001). Serum levels in mice were reported to be reduced from 0.20 to 0.15 mM after 24h of fasting (McNelis et al., 2015), but to rise in plasma compared to ad libitum state from ~0.20 to 1.50 mM after 48h fasting (Sakakibara et al., 2009). This indicates that upon food restriction, dietary and microbial influx of acetate is reduced, which is then reversed and overcompensated to supply energy upon prolonged fasting. Hence, after overnight fasting a small reduction of plasma acetic acid level might be present. In breath acetic acid, however, only a slight reduction in fasted MC4R-ki mice but no effect in the HFD mice on was detected.

4.2.1.2. **Methanol (pk33B and 35B)**

In addition to the effects of chow diet matrix on methanol, no obesity effect in MC4R-ki mice could be detected. In contrast to this mono-genetic model, mice fed a semi-purified LF or HFD diet showed only a minor increase after fasting. A small reduction in short term HFD fed (publication I) and no significant effect in long term HFD fed mice (publication II) could be detected. In obese humans, methanol levels were reported to be reduced (Halbritter et al., 2012; Turner et al., 2006) but increased in liver cirrhosis (Morisco et al., 2013). Methanol levels, as discussed previously, are strongly diet and microbiota dependent, thus altered diet or gut microbiome composition in obese patients could be

responsible for decreased levels. Elevations in cirrhotic patients can be explained by reduced hepatic alcohol dehydrogenase (adh1) activity, which is the primary detoxification mechanism in humans (Dorokhov et al., 2015). So regarding publications I and II it can be concluded that diet derived methanol can be the dominant source in terms of effect size, while obesity per se, at least in the used models, has only a minor effect on emitted methanol levels. However, it has to be noted that detoxification of methanol in humans is primarily adh1 driven while in model organisms like rodents and other mammals, peroxidative activity of catalase is more relevant for degradation (Dorokhov et al., 2015; Karinje and Ogata, 1990). This alternative methanol metabolism needs to be considered when comparing human and mouse breath methanol findings.

4.2.1.3. **Carbon dioxide*H₂O / Dimethyl sulfide (pk63, pk64B, pk65)**

In both obesity models peaks 63 (m/z 63.01) and 65B (m/z 65.01) were elevated; in HFD fed mice additionally also 64B (m/z 64.01). These peaks can be assigned to a both carbon dioxide water cluster (CO₂.H₃O⁺, m/z 63.00) and Dimethyl sulfide (DMS, C₂H₆S.H⁺, m/z 63.02) and its respective most common carbon (+ 1.003 m/z) and oxygen/sulfur (+ 2.004 m/z) isotope peaks. Although carbon dioxide has a proton affinity of 548 kJ/mol, which is lower than water (697 kJ/mol), it is present in a concentration several dimensions higher than typical VOCs. Thus, although the reaction [CO₂] + [H₂O.H⁺] \ll \rightarrow [CO₂.H⁺] + [H₂O] is strongly in favor of the back reaction, some protonated carbon dioxide can be detected. Carbon dioxide is the terminal oxidation production of most energy containing molecules and therefore used together with oxygen consumption to determine metabolic rate. Metabolic rate in various species is known to depend on ambient temperature and body mass (Gillooly et al., 2001). Indeed the used obese models do have an increased overall amount of metabolic active tissue. This increase consists not only of fat mass, which is considered to have a lower but not negligible metabolic activity per gram (Kaiyala et al., 2010), but also highly active lean mass is elevated. Thus, higher emission of carbon dioxide in heavier mice is not surprising (Butler and Kozak, 2010; Tschöp et al., 2012).

Dimethyl sulfide is the other volatile potentially contributing to peaks 63, 64B and 65. DMS was found to be elevated in rats fed a HFD, liver cirrhotic patients, obese children and is part of the fetor hepaticus (Alkhoury et al., 2014; Aprea et al., 2012; Morisco et al., 2013; Van den Velde et al., 2008). After methionine ingestion, DMS can be found in breath, indicating methionine (and possibly other sulfur-containing amino acids) as potential precursor (Kaji et al., 1979). DMS was also observed to be produced by rat skeletal muscle cells, possibly by the transamination pathway out of methionine

and cysteine (Mochalski et al., 2014). Thus, after further insight in its metabolism DMS could be used a non-invasive biomarker of altered systemic or hepatic metabolism of sulfur containing amino acids.

4.2.1.4. (Methylthio)methanethiol (MTMT, pk62)

A volatile which was identified in both obesity models is (methylthio)methanethiol (MTMT, fragment at pk62). Unexpectedly, the MTMT response in both male obesity mouse models is opposed. MC4R-ki mice show a marked decrease in MTMT source strength, while in HFD fed mice, an increase in ad libitum fed state but not in fasted state could be detected. MTMT was originally found in male mouse urine while screening for responses in mitral cells (Da Yu Lin et al., 2005). As Lin and colleagues showed, MTMT is attractive to females, but needs the context of other urinary odors to be fully effective. If detected mass spectrometrically, a fragment with m/z 61 after loss thiol group was observed, which is in this setup protonated using a PTR-MS and detected as pk62. Interestingly, they also found that MTMT is already present in bladder urine (Da Yu Lin et al., 2005); hence a contribution of fecal contamination, local bacterial metabolism or glandular secretion is not likely. Besides the identification of the receptor MOR244-3 in olfactory bulb (Duan et al., 2012), no further studies on MTMT are available. Thus, no organ or pathway of synthesis is known. The link between sexual reproduction and the energy regulatory system, namely α -MSH and MC4R was shown previously (Faulkner et al., 2015; Van der Ploeg et al., 2002). Probably, part of the reported diminished sexual motivation in male MC4R knockout mice could be due to reduced MTMT levels. In addition to the expression in the brain, MC4R has also been shown to be expressed in gut, kidney, liver and testis with a sex dependent expression level (Panaro et al., 2014; Qu et al., 2014). As MTMT is at least ~100 times higher in males than in females, the testis might be a potential location of MTMT synthesis.

In contrast to MC4R-ki mice, ad libitum fed HFD fed mice show increased MTMT source strengths. After overnight fasting, MTMT source strengths are further elevated but no longer different from LFD mice. HFD fed mice have elevated systemic leptin levels both acutely after lipid feeding and due to increased fat mass (Butler et al., 2001). Inverse to the global knockout in MC4R-ki mice, leptin in HFD-fed mice can lead to α -MSH mediated MC4R activation. Thus, the same mechanism (not) present in MC4R-ki mice can lead to modulation of MTMT emission. However, central resistance to leptin signals following a HFD feeding was described and is present already in 12 week HFD fed mice (Lin et al., 2000). Thus, it is still to answer if hypothalamic neurons expressing MC4R modulate MTMT emission or peripheral MC4R is of relevance.

In addition to obesity effects, a strong increase after overnight fasting was observed in both models. It is not clear if MTMT is induced by fasting or a circadian rhythmicity is present. In fasting, leptin levels have been reported to drop massively and out of proportion to remaining fat mass (Ahima et al., 1996). If MTMT levels would be singularly leptin signaling dependent, a reduction but not an increase in source strengths would be expected. However, if MTMT levels show a circadian rhythmicity, the disturbance of internal clocks induced by high fat diet provides another hypothesis for explaining elevated source strengths (Eckel-Mahan et al., 2013; Kohsaka et al., 2007).

4.2.2. Volatiles specifically altered in diet-induced obesity

4.2.2.1. Ammonia (pk18)

Regarding the literature about breath ammonia, elevated concentrations could be found in liver pathologies (Adeva et al., 2012) and in obese children (Alkhoury et al., 2015). In a rat study using PTR-TOF-MS, animals fed a purified HFD versus low fat “standard” diet had increased breath concentrations (Aprea et al., 2012). In contrast to that mice fed a purified HFD had lower ammonia source strengths compared to purified LFD fed animals. This seems also true for a comparison of HFD mice to both groups in the MC4R experiment. The reproducibility of breath ammonia in human studies is under debate as it is influenced by a variety of factors (Blanco Vela and Bosques Padilla, 2011). Ammonia breath concentration in addition to blood concentration varies due to mode of breathing, airway or mouth pH (Solga et al., 2013), exercise and subject training status (Solga et al., 2014) and mouth bacteria expressing urease (Chen et al., 2014). The preference of nasal breathing is very high in mice, thus mouth contaminations seem unlikely in the present setting. However, HFD fed mice do tend to have reduced spontaneous activity (Wong et al., 2015), although it is unclear if this could contribute to the difference in a setting were also LFD fed mice settle down quickly.

Systemic ammonia levels underlie a complex regulation and are dependent on endogenous and exogenous sources as well as acid-base status. Generally, ammonia is removed from circulation via hepatic metabolism in the urea cycle (Adeva et al., 2012). Mice fed a high fat diet do show higher rates of ketogenesis, as visible in acetone source strengths. In addition, HFD contains more relatively more protein (24.1 % in HFD versus 20.8 % in LFD). Both facts might contribute to a metabolic acidosis, a status which has recently discussed to be relevant for insulin resistance and type II diabetes (Akteer et al., 2015; Fagherazzi et al., 2014). As pH is a known modulator of ammonia volatility, the massive differences in ammonia source strengths in ad libitum fed and in a smaller extend also in fasted mice might be attributed to changes in the acid-base-system. Furthermore, substrate

competition may also play a role. The supply of fatty acids as an alternative energy source to renal tubular cells can lead to reduce glutamine metabolism, thus freeing less ammonia (Abate et al., 2004).

4.2.2.2. **Acrolein (pk57B)**

Acrolein (also: propenal) is, in comparison to many other volatile organic compounds, extensively studied. As a highly reactive environmental pollutant, exogenous acrolein exposure via respiration (cigarette smoke, automobile exhaust, biocide use) and dermal routes has been identified. In addition to that, oral exposures include fried food, alcoholic beverages and charred meat; furthermore endogenous production via lipid peroxidation, anti-cancer drugs, polyamines and threonine contributes to acrolein levels (Moghe et al., 2015). In HFD fed mice, acrolein levels are increased compared to LFD fed mice. HFD feeding results in a pro-oxidative state and increased lipid peroxidation, thus creating more acrolein (Matsuzawa-Nagata et al., 2008).

Interestingly, the effect on breath acrolein levels is more pronounced in ad libitum state. In ad libitum state, both dietary acrolein from oxidation of the HFD and endogenously produced acrolein due to oxidative stress and ROS production add up. Notably, acrolein is considered both a product and an initiator of lipid peroxidation (Adams Jr. and Klaidman, 1993; Uchida et al., 1998), therefore potentiating its contribution to systemic oxidative stress. Breath acrolein might thus be used as a non-invasive marker of redox-status and oxidative stress levels.

4.2.2.3. **Methyl acetate (pk75)**

Interestingly, in addition to the strong diet matrix induced source strength drop of pk75 in publication I, an overall increase in HFD-fed mice was observed. As there is no effect in the mono-genetic MC4R-Ki mice, this is likely not a general obesity but HFD model specific phenomenon. Feeding a high fat diet is known to increase volatile esters in the feces headspace of obese humans suffering from non-alcoholic fatty liver disease (Raman et al., 2013). The emission of esters is thought to be derived from the high fat diet induced shift in gut microbiota. Another potential origin is the synthesis within the mouse organism, as synthesis of methyl acetate from e.g. human bronchial epithelial primary cells is known (Filipiak et al., 2010). In HFD fed mice, acetic acid is increased by trend in breath and described to be elevated in plasma when standardizing for fiber content as discussed

earlier. Thus, either intestinal or systemic changes induced by a high dietary fat content lead to elevated volatile esters like methyl acetate.

4.2.2.4. $^{18}\text{O}^{16}\text{O}$ oxygen (pk34A)

In the experimental setup higher consumption of oxygen leads to stronger decrease during a measurement phase, which is detectable in the oxygen isotope ^{18}O - ^{16}O (as in $^{16}\text{O}_2$ the detector is saturated). A stronger consumption of oxygen measured as more negative source strength in oxygen isotope is observed in HFD fed mice. As shown in publication II, HFD fed mice do consist of more metabolic active tissue, as both lean and fat mass are elevated. Thus, higher demand for energy supplied by primarily oxidative processes in the mitochondria lead to higher oxygen consumption (Tschöp et al., 2012).

4.2.2.5. $\text{H}_3\text{O}^+(\text{H}_2\text{O})_2$ water cluster and fragments of aldehydes (pk55B)

Pk55B is likely to consist of both $\text{H}_3\text{O}^+(\text{H}_2\text{O})_2$ water cluster and fragments of aldehydes e.g. butanal, hexanal, octanal or nonanal (Buhr et al., 2002). In HFD fed mice, elevated source strength of pk55B was found; in addition fasted mice were increased as well. It is not clear why humidity and water clustering should be increased in obese mice. However, a chronic increase in subclinical inflammation is known in obesity (Wellen and Hotamisligil, 2003). Breath (and breath condensate) in inflammatory oxidative stress with associated endogenous lipid peroxidation is known to contain elevated concentrations of aldehydes (Amann et al., 2014c).

4.2.2.6. Acetone and propanol (pk59, pk43B)

Source strengths of acetone and propanol were found to be elevated on HFD (acetone) and after fasting (both). The major energy source in states of food restriction or prolonged HFD uptake is shifted from glucose metabolism to lipid oxidation. Especially hepatic lipid oxidation fuels ketogenesis which provides ketone bodies for energy supply. Acetone is the only ketone body measurable from breath and is correlated to blood concentrations of the other two ketone bodies β -

hydroxybutyrate and aceto-acetate in humans (Musa-Veloso et al., 2006; Qiao et al., 2014). Thus, elevations of acetone can be used to monitor states of enhanced ketogenesis and lipid oxidation.

Acetone has a history of studies associating it with diabetic states of ketoacidosis (Leopold et al., 2014). Interestingly, the acetone - correlated ketone bodies β -hydroxybutyrate and aceto-acetate were associated to elevated fasting and 2h plasma glucose levels (Mahendran et al., 2013). Furthermore aceto-acetate could be used to predict increased GTT AUC and 5-year diabetes incidence. However, due to inter-individual variation, acetone itself is not suitable as a single signal biomarker for diabetes phenotyping (Leopold et al., 2014). In this context it has to be noted that in the gaussian graphical network model of publication II, a fasting responsive subnetwork connected to acetone was found. Two fragments of propanol, pk41 and pk43B, show a high partial correlation to acetone. This likely reflects the conversion from iso-propanol to acetone, which is catalyzed by alcohol dehydrogenase and elevated in states of ketosis (Lewis et al., 1984; Petersen et al., 2012). Breath iso-propanol in clinical settings are observed to be highly correlated to environmental air concentrations, which are originated likely from disinfectant use (Ghimenti et al., 2013). This reflects a relevant “wash-in” of iso-propanol, which can, as shown for ingested iso-propanol, influence exhaled acetone concentrations in the range of minutes (Ruzsanyi et al., 2014). Thus, propanol levels derived from various sources contribute to acetone breath levels and explain part of the variance observed. It might be possible to enhance the biomarker potential of acetone by measuring humans in a propanol reduced environment and taking exhaled propanol levels into consideration.

4.2.3. Volatiles specifically altered by MC4R nonsense knock-in

4.2.3.1. Cluster of unknowns 151, 153, 137B and 81B

In publication II, a subnetwork of VOCs was affected especially in MC4R-ki mice, including peaks 151, 153, 137B and 81B. Although impossible to interpret on individual level, a similar combination of peaks was observed in a PTR-MS study of monoterpenes like α - and β -pinene, 3-carene, limonene and camphor, which produced fragment ions of masses 67, 81 and 95 as well as a protonated molecular ion of mass 137 or 153 (Tani et al., 2003). Regarding the literature of mono-terpenes and metabolic diseases, a link to liver cirrhosis in humans was suggested (Morisco et al., 2013). Herein, the authors hypothesize a changed dietary pattern in cirrhotic subjects to cause altered mono-terpene levels. In addition, the modelling of plasma TG and FFA levels in a human lipid infusion study included β -limonene and β -pinene (Minh et al., 2012). In contrast, MC4R-ki mice show no changes in plasma TG/FFA (GMC primary screen, data not shown) indicating that this link does not explain

elevated monoterpenes. Therefore in MC4R-ki mice either increased dietary monoterpene intake due to hyperphagia, more storage in adipose tissue of this lipophilic substance class or an alteration monoterpene metabolism can provide explanations the elevated levels.

Regarding the hypothesized molecules, it has to be noted that the peak mass does not match the theoretical masses exactly. In PTR-MS measurements, typically an internal calibration is applied, using $\text{H}_3^{18}\text{O}^+$ (21.02), NO^+ (30.00) and protonated acetone (59.05). As the highest calibration mass is far lower than the peaks of interest, an additional calibration using external VOCs should be applied in future studies.

4.2.4. Unassigned candidate VOCs altered in obesity and strategies for identification

Furthermore, peaks 50, 81B and 117B were identified to be changed in response to obesity but could not be tentatively assigned to candidate VOCs. In HFD fed mice, the same was true for peaks 249B, 253A, 57A, 36B and 74B, while in MC4R-ki mice, peaks 44A, 123B, 135B, 36A, 122B, 137B, 113, 94B, 80B, 87B, 211 and 95B were altered but molecular structure is still unknown. Although these peaks are not yet identified, they do provide chances for further potential biomarkers for obesity and associated disease risks.

The obtained sets of candidate VOCs can be used to gain further insight into the origins and biology of VOCs in individual experiments. However, for understanding of physiological changes associated with the peaks a reasonable certainty of the volatile molecule needs to be achieved. This is a common problem in singly PTR-MS based studies, as this technology trades off specificity and detailed chemical information for extraordinary sensitivity and the possibility of measuring dynamics (Cappellin et al., 2013). Identification could be addressed using Fast-GC-PTR-TOF-MS or, the “gold standard”, GC-MS with pre-concentration (Romano et al., 2014). However, these methods do have decreased sensitivity in comparison to direct PTR-MS measurements and would probably demand changes in sampling procedure. Approaches like the nose-mask in rat provide undiluted breath sampling but do have limitations as well like animal stress and the necessity of train to avoid this (Aprea et al., 2012). In addition, data-driven approaches can be applied to address isotope ratios, formation of water clusters and underlying metabolic pathways to gain further insight on the potential molecules. Here, heatmaps combined with hierarchical clustering as well as gaussian graphical models as a novel tool for breath analysis were successfully utilized to clarify molecular origin of

certain candidate peaks for obesity. In extending the adaption of established tools found in other “omics”-technologies, further insight into the large PTR-TOF-MS datasets can be gained.

4.3. On the methodology of breath gas analysis in rodents

4.3.1. On the evolution of the unrestrained mouse VOC measurement and the corresponding data analysis

Based on the initial methodology, the analysis of VOCs in unrestrained and non-anesthetized mice was developed further to address needs of both secondary screening of mice within the systemic phenotyping in the GMC as well as individual VOC-driven projects. The implementation of the PTR-TOF-MS allowed a reduction from 15 to 5 min measurement time per profile while monitoring all volatiles simultaneously with massively higher precision in terms of mass resolution. However, this implementation did increase the amount of data per mouse by several orders of magnitude, therefore demanding a revision of the available data procession and data analysis tools. A server based data analysis platform was implemented to cope with the massively increased amount of data (Chang et al., 2015). Herein, source strength calculation can be performed in a semi-automated way still allowing for manual intervention and visual quality control during the analysis process. In addition, data exploration and quality control tools like a correlation heatmap and a visualizer of mass spectrometer machine parameters were implemented. Especially for the screening approach, the possibility to visualize data on cohort level in the server platform can be used to gain fast insight into experimental data.

The analysis of VOCs from mouse headspace typically results in $n \ll p$ datasets, thus a lot more parameters than animals are measured. The evaluation of such datasets from screening experiments with wide experimental hypothesis can be challenging with conventional statistical methods as e.g. interference testing. The more parameters are tested for differences at a certain significance level, the more likely a false positive finding gets (cumulation of alpha error). Methods have been developed to correct for multiple testing by e.g. Bonferroni’s control of family-wise error rate or Benjamini and Hochberg’s control of the false discovery rate which vary in power and beta error rate (Benjamini and Hochberg, 1995). Another approach to model such datasets is the use of machine learning techniques. Although computationally more expensive than classical statistics, machine learning algorithms as e.g. random forest based techniques can be used both for predicting responses on the basis of existing data as well as selecting the most important parameters different between experimental groups (Breiman, 2001). For the analysis of mouse VOC data, the selection of relevant VOCs is performed

by using RF-based methods in combination with cross validation and permutation analysis to obtain robust sets of candidate VOCs. These methods, which are state of the art in the field breath research (Smolinska et al., 2014), were established using R-based scripts to allow for rapid and custom evaluation of data.

4.3.2. Advantages and limitations of the established setup in comparison to published literature and implications on human breath analysis

Model organisms such as the laboratory mouse are widely used in biomedical research to study systemic responses in controlled and defined experimental situations. Concerning the mouse as a model organism for VOC analysis, some model characteristics have to be noted. The tidal volume of a mouse is only 0.2 - 0.3 ml while breathing frequency varies between 80 and 230 per minute (Reinhard et al., 2014). Thus, single breath exhalation VOC analysis as performed in humans is technically challenging if not impossible. In the initial methodology developed by Szymczak et. al, the emission of VOCs is calculated from saturation curves measured by a high sensitivity PTR-Quad-MS and expressed as a source strength which is an elegant way of assessing this information (Szymczak et al., 2014). Over the recent years, the lack of VOCs analysis in model organisms was recognized by the field and several other groups developed solutions to measure rodent VOCs. Table 2 gives an overview over the developed approaches and presents several key characteristics of the methods.

Table 2: Established methods for VOC measurements in rodent models

Method description	Saturation profiles of unrestrained mouse in chamber	Mouse in tube, Head-space using constant flow	Tracheal intubation	Tracheal intubation & ventilation & bag collection	Nose mask sampling	Tracheotomy & ventilation
Publications	(Kistler et al., 2016, 2014; Szymczak et al., 2014)	(Li et al., 2015)	(Neuhaus et al., 2011; Vautz et al., 2010)	(Zhu et al., 2013a, 2013b)	(Aprèa et al., 2012)	(Albrecht et al., 2015; Fink et al., 2015, 2014; Wolf et al., 2014)
Model	Mouse, C57Bl/6J and N	Mouse, C57Bl/6J	Mouse, BALB/c	Mouse, C57Bl/6J	Rat, Wistar	Rat, Sprague Dawley
Detection method	PTR-TOF-MS	Secondary electron spray ionization (SESI) –MS	Ion mobility spectrometry (IMS) & solid phase micro extraction (SPME) –GC MS	Secondary electron spray ionization (SESI) - MS	PTR-TOF-MS	Ion mobility spectrometry (IMS)
Food restriction before measurement	Yes (overnight) & No	No	No	No	No	Yes (12h)
Restrained	No	Partly, small container	No	No	Yes	No
Anesthesia	No	No	Yes (ketamine, xylazine)	Yes (pentobarbital, pancuronium bromide)	No	Yes (sevoflurane, sodium pentobarbital)
External ventilation	No	No	No	Yes	No	Yes
Invasiveness	None	None	Moderate	Moderate - High	None	High
Time per single data point / animal	3s conc.; 5 - 15 min source strength/ 21 – ≥ 60 min	10 s / 90 min	5-10 min / 5-10 min	40 – 60 min / 40 – 60 min/	20 s / 40 s	20 min / up to 24 h
Maximum repetitions/ Performed time range	arbitrarily repeatable / Several weeks	arbitrarily repeatable / 90 min	arbitrarily repeatable, limited due to anesthesia / 3 days	arbitrarily repeatable, limited due to anesthesia / several days	arbitrarily repeatable / 2 days	Terminal experiment / 24h
VOC sources / Dilution of VOCs	Whole body/ yes	Whole body/ yes	Breath/ no	Breath/ no	Breath/ no	Breath/ no
Animal stress potential	low	Low to moderate	Moderate	Moderate	Moderate to high (if no training)	Moderate
Contamination possible / controllable	Yes / yes	Yes / possible, not mentioned	No / -	No / -	No / -	No / -
Supply of synthetic air for breathing	Yes	Yes	No	No	No	Yes
Training necessary	No, acclimation performed	No	No	No	Yes	No
Research question(s)	Diet, Obesity	Drug pharmacodynamics	Allergy	Lung infections	Steatohepatitis	Inflammation, Sepsis, Fasting, Glucose infusion

The available methods are designed for a wide range of experimental questions and therefore differ in their particular implementations. The methodology applied in this thesis, which is based on the emitted VOCs from an un-restrained and non-anaesthetized mouse in a respiratory chamber, is established with the needs of large-scale phenotyping in mind. When a mouse is systemically phenotyped by using multiple experimental procedures within a limited time span, potential additional tests need to consider the impact on the following pipeline in terms of duration and burden. The implementation as a non-invasive method allows screening with low animal stress, reasonable throughput and the possibility to repeat measurements as often as desired. As a freely moving mouse in a box does emit VOCs from fur, urine and feces if present, one has to consider ways to control for such contaminations. We have established marker substances for urine and feces, which by using fast online VOC measurement are immediately indicating such contaminations. Other factors like emission of gases through skin are typically rather low in endothermal organisms (Feder and Burggren, 1985), although production on the skin and fur can occur. In addition, changes in respiratory activity can alter measured VOC source strengths; to what extent needs to be addressed in detail. The mouse in a box principle does not lead to simple concentrations but saturation curves. This data structure is more complex to handle (here by calculating source strength) but does give a robust estimate of VOC emission over the timespan of several minutes.

Interestingly, methods can be categorized by certain common design principles as summarized in table 2. A major decision is whether the complete rodent's headspace or just the breath is sampled. In the headspace situation, there is some dilution of the VOCs, potential fur contribution and the necessity to control for contamination. The direct sampling of breath in rodents is, unlike in the human situation, essentially combined with either restraintment or anesthesia. Although pure breath seems comparable to the human situation, animal stress and anesthesia do change e.g. on blood glucose and lactate levels, indicating a profound effect on rodent metabolism (Schwarzkopf et al., 2013). It thus depends on the specific research question whether this is acceptable or not. State of the art in most studies is the use of online measurement techniques which increases sensitivity, and, depending on the method applied, allows high time resolution. Bag sampling can be applied as well, but dynamics are lost and storage time needs to be as short as possible (Beauchamp et al., 2008; Steeghs et al., 2007).

Unfortunately, the use of multiple detection methods and VOC standards to get robust peak identification is not standard in the field yet. Thus, findings of altered VOC patterns and tentatively assigned candidates often cannot be related to the underlying biological mechanisms. This is,

especially in rodent studies where mechanistic experiments in single organs or tissues could be performed, a major barrier on the way of VOC analysis from laboratory to patient use.

Rodent studies have the potential to overcome several limitations of human studies. Environments can be controlled, genetic background of the study population is narrowed down tremendously and defined diets can be applied. However, the question whether and how long diet should be restricted (a rather common strategy to eliminate variation in human studies) is not answered yet, although relevant dietary contributions to VOC emissions are shown in publication I (Kistler et al., 2014). It might be favorable to restrict diet access before measurement, but depends on the very experiment if this restriction should be extended so that animals are already in a fasted state, which in itself leads to a relevant shift in VOC patterns as seen in publication II (Kistler et al., 2016).

Another often discussed problem in human measurements is the contribution of mouth microbiota to measured VOCs, thus the emerging question whether to sample from mouth or nose (Beauchamp and Pleil, 2013; Chen et al., 2014). In rodents, this problem is avoided in most approaches by directly sampling from the trachea, thus circumventing the upper respiratory system completely. Independently of sampling, rodents are obligately breathing through the nose and oral breathing occurs only in extreme situations like hypoxia (Drazen et al., 1999). Because of this characteristic, also non-invasive rodent methods do not suffer from mouth microbiota contribution as much as human measurements.

The so-called “wash-in/ wash-out” effect in human studies describes the phenomenon that VOCs present in inhaled air or incorporated via diet will be transferred to and be present in the organism. Consequently, such external VOCs will then be exhaled over time as soon as environmental concentrations are reduced, thus being washed-out (Maurer et al., 2014). Such “wash-in/wash-out” effects are dependent on the chemical properties of the volatiles and do contribute to the variability in human studies (Španěl et al., 2013). In rodents as model organisms, this wash-in from the cage environments is present but very comparable between animals due to controlled conditions. Before and during the measurement, it is probably good practice to use synthetic, thus VOC-reduced, air supply for the animals. Laboratory air is, similar to the clinical situation, often rich in VOCs used for disinfection as e.g. 2-propanol, which can be converted to acetone after ingestion and probably also after inhalation (Ruzsanyi et al., 2014). Notably, not all approaches do use synthetic air although this would be an easy way to further reduce VOC variability in rodent measurements.

5. Conclusion and Outlook

In this thesis, the effects of diet-matrix and both diet induced and mono-genetic obese mouse models on VOC signature were studied. In result, both diet-matrix and obesity show profound effects on the volatilome emitted and are as such very relevant to be standardized in further rodent and human measurements. In addition to that, especially the findings in both obesity models show potential for non-invasive biomarkers associated to obesity and/or its changed risk profile. In human breath analysis, recently efforts have been made to catalogue and summarize the volatilome from breath and various other sources (Amann et al., 2014a). Building on this, the field of rodent VOC analysis can add multiple contributions to deepen the understanding of VOC biology. First of all, the volatilome of the rat and the mouse also needs to be described in its completeness, without much of the human environmental contributions. This in comparison with its human counterpart could lead to a mammalian core volatilome, omitting external VOC sources. In addition to this catalogue, in rodents the origin and biological relevance of the encountered VOCs can and needs to be studied in detail, mapping VOCs to metabolic functions. The implicit knowledge of biological mechanisms will further aid to evolve laboratory-based breath analysis into a non-invasive diagnostic alternative applied in the clinic setting. In addition, screening studies in larger human cohorts are rare to date and will be needed to facilitate the breakthrough of breath analysis in every day clinics. And, hand in hand with the advanced knowledge on volatile metabolism, challenge tests using e.g. stable isotope labelled VOCs precursor molecules can be used as non-invasive clinical read-outs in the future. Those, as well, can and need to go back to rodent models to be established. Thus, as it has been stated, the field of breath analysis is currently in the process of “validating volatile promises” (Boots et al., 2015) and the use of rodent models is about to be a relevant part of that.

6. References

- Abate, N., Chandalia, M., Cabo-Chan, A.V., Moe, O.W., Sakhaee, K., 2004. The metabolic syndrome and uric acid nephrolithiasis: Novel features of renal manifestation of insulin resistance. *Kidney Int.* 65, 386–392. doi:10.1111/j.1523-1755.2004.00386.x
- Adams Jr., J.D., Klaidman, L.K., 1993. Acrolein-induced oxygen radical formation. *Free Radic. Biol. Med.* 15, 187–193. doi:10.1016/0891-5849(93)90058-3
- Adeva, M.M., Souto, G., Blanco, N., Donapetry, C., 2012. Ammonium metabolism in humans. *Metabolism.* 61, 1495–1511. doi:10.1016/j.metabol.2012.07.007
- Ahima, R.S., Prabakaran, D., Mantzoros, C., Qu, D., Lowell, B., Maratos-Flier, E., Flier, J.S., 1996. Role of leptin in the neuroendocrine response to fasting. *Nature* 382, 250–252. doi:10.1038/382250a0
- Ajibola, O.A., Smith, D., Španěl, P., Ferns, G.A.A., 2013. Effects of dietary nutrients on volatile breath metabolites. *J. Nutr. Sci.* 2. doi:10.1017/jns.2013.26
- Akter, S., Eguchi, M., Kuwahara, K., Kochi, T., Ito, R., Kurotani, K., Tsuruoka, H., Nanri, A., Kabe, I., Mizoue, T., 2015. High dietary acid load is associated with insulin resistance: The Furukawa Nutrition and Health Study. *Clin. Nutr.* doi:10.1016/j.clnu.2015.03.008
- Albrecht, F.W., Hüppe, T., Fink, T., Maurer, F., Wolf, A., Wolf, B., Volk, T., Baumbach, J.I., Kreuer, S., 2015. Influence of the respirator on volatile organic compounds: an animal study in rats over 24 hours. *J. Breath Res.* 9, 016007. doi:10.1088/1752-7155/9/1/016007
- Alkhoury, N., Cikach, F., Eng, K., Moses, J., Patel, N., Yan, C., Hanouneh, I., Grove, D., Lopez, R., Dweik, R., 2014. Analysis of breath volatile organic compounds as a noninvasive tool to diagnose nonalcoholic fatty liver disease in children: *Eur. J. Gastroenterol. Hepatol.* 26, 82–87. doi:10.1097/MEG.0b013e3283650669
- Alkhoury, N., Eng, K., Cikach, F., Patel, N., Yan, C., Brindle, A., Rome, E., Hanouneh, I., Grove, D., Lopez, R., Hazen, S.L., Dweik, R.A., 2015. Breathprints of childhood obesity: changes in volatile organic compounds in obese children compared with lean controls. *Pediatr. Obes.* 10, 23–29. doi:10.1111/j.2047-6310.2014.221.x
- Amann, A., Costello, B. de L., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., Ratcliffe, N., Risby, T., 2014a. The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* 8, 034001. doi:10.1088/1752-7155/8/3/034001
- Amann, A., Miekisch, W., Schubert, J., Buszewski, B., Ligor, T., Jezierski, T., Pleil, J., Risby, T., 2014b. Analysis of Exhaled Breath for Disease Detection. *Annu. Rev. Anal. Chem.* 7, 455–482. doi:10.1146/annurev-anchem-071213-020043
- Amann, A., Miekisch, W., Schubert, J., Buszewski, B., Ligor, T., Jezierski, T., Pleil, J., Risby, T., 2014c. Analysis of Exhaled Breath for Disease Detection. *Annu. Rev. Anal. Chem.* 7, 455–482. doi:10.1146/annurev-anchem-071213-020043
- Amirshahrokhi, K., Bohlooli, S., 2013. Effect of Methylsulfonylmethane on Paraquat-Induced Acute Lung and Liver Injury in Mice. *Inflammation* 36, 1111–1121. doi:10.1007/s10753-013-9645-8
- Aprea, E., Morisco, F., Biasioli, F., Vitaglione, P., Cappellin, L., Soukoulis, C., Lembo, V., Gasperi, F., D'Argenio, G., Fogliano, V., Caporaso, N., 2012. Analysis of breath by proton transfer reaction time of flight mass spectrometry in rats with steatohepatitis induced by high-fat diet. *J. Mass Spectrom.* JMS 47, 1098–1103. doi:10.1002/jms.3009
- Axelrod, J., Daly, J., 1965. Pituitary gland: enzymic formation of methanol from S-adenosylmethionine. *Science* 150, 892–893.
- Bajtarevic, A., Ager, C., Pienz, M., Klieber, M., Schwarz, K., Ligor, M., Ligor, T., Filipiak, W., Denz, H., Fiegl, M., Hilbe, W., Weiss, W., Lukas, P., Jamnig, H., Hackl, M., Haidenberger, A.,

- Buszewski, B., Miekisch, W., Schubert, J., Amann, A., 2009. Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC Cancer* 9, 348. doi:10.1186/1471-2407-9-348
- Barker, M., Hengst, M., Schmid, J., Buers, H.-J., Mittermaier, B., Klemp, D., Koppmann, R., 2006. Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis. *Eur. Respir. J.* 27, 929–936. doi:10.1183/09031936.06.00085105
- Beauchamp, J., 2011. Inhaled today, not gone tomorrow: pharmacokinetics and environmental exposure of volatiles in exhaled breath. *J. Breath Res.* 5, 037103. doi:10.1088/1752-7155/5/3/037103
- Beauchamp, J.D., Pleil, J.D., 2013. Simply breath-taking? Developing a strategy for consistent breath sampling. *J. Breath Res.* 7, 042001. doi:10.1088/1752-7155/7/4/042001
- Beauchamp, J., Herbig, J., Gutmann, R., Hansel, A., 2008. On the use of Tedlar® bags for breath-gas sampling and analysis. *J. Breath Res.* 2, 046001.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- Biasioli, F., Yeretian, C., Märk, T.D., Dewulf, J., Van Langenhove, H., 2011. Direct-injection mass spectrometry adds the time dimension to (B)VOC analysis. *TrAC Trends Anal. Chem., Biogenic Volatile Organic Compounds S.I.* 30, 1003–1017. doi:10.1016/j.trac.2011.04.005
- Blanco Vela, C.I., Bosques Padilla, F.J., 2011. Determination of ammonia concentrations in cirrhosis patients—still confusing after all these years? *Ann. Hepatol.* 10 Suppl 2, S60–65.
- Bolze, F., Rink, N., Brumm, H., Kühn, R., Mocek, S., Schwarz, A.-E., Kless, C., Biebermann, H., Wurst, W., Rozman, J., Klingenspor, M., 2011. Characterization of the melanocortin-4-receptor nonsense mutation W16X in vitro and in vivo. *Pharmacogenomics J.* 13. doi:10.1038/tpj.2011.43
- Boots, A.W., Berkel, J.J.B.N. van, Dallinga, J.W., Smolinska, A., Wouters, E.F., Schooten, F.J. van, 2012. The versatile use of exhaled volatile organic compounds in human health and disease. *J. Breath Res.* 6, 027108. doi:10.1088/1752-7155/6/2/027108
- Boots, A.W., Bos, L.D., van der Schee, M.P., van Schooten, F.-J., Sterk, P.J., 2015. Exhaled Molecular Fingerprinting in Diagnosis and Monitoring: Validating Volatile Promises. *Trends Mol. Med.* 21, 633–644. doi:10.1016/j.molmed.2015.08.001
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32. doi:10.1023/A:1010933404324
- Buhr, K., van Ruth, S., Delahunty, C., 2002. Analysis of volatile flavour compounds by Proton Transfer Reaction-Mass Spectrometry: fragmentation patterns and discrimination between isobaric and isomeric compounds. *Int. J. Mass Spectrom.* 221, 1–7. doi:10.1016/S1387-3806(02)00896-5
- Buśko, M., Jeleń, H., Góral, T., Chmielewski, J., Stuper, K., Szwajkowska-Michalek, L., Tyrakowska, B., Perkowski, J., 2010. Volatile metabolites in various cereal grains. *Food Addit. Contam. Part Chem. Anal. Control Expo. Risk Assess.* 27, 1574–1581. doi:10.1080/19440049.2010.506600
- Butler, A.A., Kozak, L.P., 2010. A Recurring Problem With the Analysis of Energy Expenditure in Genetic Models Expressing Lean and Obese Phenotypes. *Diabetes* 59, 323–329. doi:10.2337/db09-1471
- Butler, A.A., Marks, D.L., Fan, W., Kuhn, C.M., Bartolome, M., Cone, R.D., 2001. Melanocortin-4 receptor is required for acute homeostatic responses to increased dietary fat. *Nat. Neurosci.* 4, 605–611. doi:10.1038/88423
- Buuren, S. van, Groothuis-Oudshoorn, K., 2011. mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* 45, 1–67.
- Cappellin, L., Loreto, F., Aprea, E., Romano, A., del Pulgar, J., Gasperi, F., Biasioli, F., 2013. PTR-MS in Italy: A Multipurpose Sensor with Applications in Environmental, Agri-Food and Health Science. *Sensors* 13, 11923–11955. doi:10.3390/s130911923
- Carvalho, B.M., Guadagnini, D., Tsukumo, D.M.L., Schenka, A.A., Latuf-Filho, P., Vassallo, J., Dias, J.C., Kubota, L.T., Carvalheira, J.B.C., Saad, M.J.A., 2012. Modulation of gut microbiota by

- antibiotics improves insulin signalling in high-fat fed mice. *Diabetologia* 55, 2823–2834. doi:10.1007/s00125-012-2648-4
- Chang, W., Cheng, J., Allaire, J.J., Xie, Y., McPherson, J., 2015. shiny: Web Application Framework for R.
- Chen, S., Zieve, L., Mahadevan, V., 1970. Mercaptans and dimethyl sulfide in the breath of patients with cirrhosis of the liver. Effect of feeding methionine. *J. Lab. Clin. Med.* 75, 628–635.
- Chen, W., Metsälä, M., Vaittinen, O., Halonen, L., 2014. The origin of mouth-exhaled ammonia. *J. Breath Res.* 8, 036003. doi:10.1088/1752-7155/8/3/036003
- Cikach, F.S., Dweik, R.A., 2012. Cardiovascular Biomarkers In Exhaled Breath. *Prog. Cardiovasc. Dis.* 55, 34–43. doi:10.1016/j.pcad.2012.05.005
- Cordain, L., Eaton, S.B., Sebastian, A., Mann, N., Lindeberg, S., Watkins, B.A., O’Keefe, J.H., Brand-Miller, J., 2005. Origins and evolution of the Western diet: health implications for the 21st century. *Am. J. Clin. Nutr.* 81, 341–354.
- Costello, B. de L., Amann, A., Al-Kateb, H., Flynn, C., Filipiak, W., Khalid, T., Osborne, D., Ratcliffe, N.M., 2014. A review of the volatiles from the healthy human body. *J. Breath Res.* 8, 014001. doi:10.1088/1752-7155/8/1/014001
- Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., Almeida, M., Quinquis, B., Levenez, F., Galleron, N., Gougis, S., Rizkalla, S., Batto, J.-M., Renault, P., Consortium, A.M., Doré, J., Zucker, J.-D., Clément, K., Ehrlich, S.D., ANR MicroObes consortium Members, 2013. Dietary intervention impact on gut microbial gene richness. *Nature* 500, 585–588. doi:10.1038/nature12480
- Daniel, H., Moghaddas Gholami, A., Berry, D., Desmarchelier, C., Hahne, H., Loh, G., Mondot, S., Lepage, P., Rothballer, M., Walker, A., Böhm, C., Wenning, M., Wagner, M., Blaut, M., Schmitt-Kopplin, P., Kuster, B., Haller, D., Clavel, T., 2014. High-fat diet alters gut microbiota physiology in mice. *ISME J.* 8, 295–308. doi:10.1038/ismej.2013.155
- David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., Biddinger, S.B., Dutton, R.J., Turnbaugh, P.J., 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. doi:10.1038/nature12820
- Da Yu Lin, S.-Z.Z., Block, E., Katz, L.C., 2005. Encoding social signals in the mouse main olfactory bulb. *Nature* 434, 470–477.
- Després, J.-P., Lemieux, I., 2006. Abdominal obesity and metabolic syndrome. *Nature* 444, 881–887. doi:10.1038/nature05488
- Donovan, A., 1996. Antoine Lavoisier: Science, Administration and Revolution. Cambridge University Press.
- Dorokhov, Y.L., Komarova, T.V., Petrunia, I.V., Kosorukov, V.S., Zinovkin, R.A., Shindyapina, A.V., Frolova, O.Y., Gleba, Y.Y., 2012. Methanol May Function as a Cross-Kingdom Signal. *PLoS ONE* 7. doi:10.1371/journal.pone.0036122
- Dorokhov, Y.L., Shindyapina, A.V., Sheshukova, E.V., Komarova, T.V., 2015. Metabolic Methanol: Molecular Pathways and Physiological Roles. *Physiol. Rev.* 95, 603–644. doi:10.1152/physrev.00034.2014
- Dragonieri, S., Porcelli, F., Longobardi, F., Carratù, P., Aliani, M., Ventura, V.A., Tutino, M., Quaranta, V.N., Resta, O., Gennaro, G. de, 2015. An electronic nose in the discrimination of obese patients with and without obstructive sleep apnoea. *J. Breath Res.* 9, 026005. doi:10.1088/1752-7155/9/2/026005
- Drazen, J.M., Finn, P.W., De Sanctis, G.T., 1999. Mouse models of airway responsiveness: physiological basis of observed outcomes and analysis of selected examples using these outcome indicators. *Annu. Rev. Physiol.* 61, 593–625.
- Duan, X., Block, E., Li, Z., Connelly, T., Zhang, J., Huang, Z., Su, X., Pan, Y., Wu, L., Chi, Q., Thomas, S., Zhang, S., Ma, M., Matsunami, H., Chen, G.-Q., Zhuang, H., 2012. Crucial role of copper in detection of metal-coordinating odorants. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3492–3497. doi:10.1073/pnas.1111297109

- Eckel-Mahan, K.L., Patel, V.R., de Mateo, S., Orozco-Solis, R., Ceglia, N.J., Sahar, S., Dilag-Penilla, S.A., Dyar, K.A., Baldi, P., Sassone-Corsi, P., 2013. Reprogramming of the Circadian Clock by Nutritional Challenge. *Cell* 155, 1464–1478. doi:10.1016/j.cell.2013.11.034
- Eisenmann, A., Amann, A., Said, M., Datta, B., Ledochowski, M., 2008. Implementation and interpretation of hydrogen breath tests. *J. Breath Res.* 2, 046002. doi:10.1088/1752-7155/2/4/046002
- Eriksen, S.P., Kulkarni, A.B., 1963. Methanol in Normal Human Breath. *Science* 141, 639–640. doi:10.1126/science.141.3581.639
- European Parliament, Council of the European Union, 2004. Directive 2004/42/CE of the European Parliament and of the Council of 21 April 2004 on the limitation of emissions of volatile organic compounds due to the use of organic solvents in certain paints and varnishes and vehicle refinishing products and amending Directive 1999/13/EC.
- Fagherazzi, G., Vilier, A., Bonnet, F., Lajous, M., Balkau, B., Boutron-Ruault, M.-C., Clavel-Chapelon, F., 2014. Dietary acid load and risk of type 2 diabetes: the E3N-EPIC cohort study. *Diabetologia* 57, 313–320. doi:10.1007/s00125-013-3100-0
- Faulkner, L., Dowling, A., Stuart, R., Nillni, E., Hill, J., 2015. Reduced melanocortin production causes sexual dysfunction in male mice with POMC neuronal insulin and leptin insensitivity. *Endocrinology* en.2014–1788. doi:10.1210/en.2014-1788
- Feder, M.E., Burggren, W.W., 1985. Cutaneous Gas Exchange in Vertebrates: Design, Patterns, Control and Implications. *Biol. Rev.* 60, 1–45. doi:10.1111/j.1469-185X.1985.tb00416.x
- Fink, T., Albrecht, F.W., Maurer, F., Kleber, A., Hüppe, T., Schnauber, K., Wolf, B., Baumbach, J.I., Volk, T., Kreuer, S., 2015. Exhalation pattern changes during fasting and low dose glucose treatment in rats. *Anal. Bioanal. Chem.* 1–11. doi:10.1007/s00216-015-8602-9
- Fink, T., Wolf, A., Maurer, F., Albrecht, F.W., Heim, N., Wolf, B., Hauschild, A.C., Bödeker, B., Baumbach, J.I., Volk, T., Sessler, D.I., Kreuer, S., 2014. Volatile Organic Compounds during Inflammation and Sepsis in Rats: A Potential Breath Test Using Ion-mobility Spectrometry. *Anesthesiology*. doi:10.1097/ALN.0000000000000420
- Fuchs, H., Gailus-Durner, V., Adler, T., Pimentel, J.A.A., Becker, L., Bolle, I., Brielmeier, M., Calzada-Wack, J., Dalke, C., Ehrhardt, N., Fasnacht, N., Ferwagner, B., Frischmann, U., Hans, W., Hölter, S.M., Hölzlwimmer, G., Horsch, M., Javaheri, A., Kallnik, M., Kling, E., Lengger, C., Maier, H., Mossbrugger, I., Mörth, C., Naton, B., Nöth, U., Pasche, B., Prehn, C., Przemeck, G., Puk, O., Racz, I., Rathkolb, B., Rozman, J., Schäble, K., Schreiner, R., Schrewe, A., Sina, C., Steinkamp, R., Thiele, F., Willershäuser, M., Zeh, R., Adamski, J., Busch, D.H., Beckers, J., Behrendt, H., Daniel, H., Esposito, I., Favor, J., Graw, J., Heldmaier, G., Höfler, H., Ivandic, B., Katus, H., Klingenspor, M., Klopstock, T., Lengeling, A., Mempel, M., Müller, W., Neschen, S., Ollert, M., Quintanilla-Martinez, L., Rosenstiel, P., Schmidt, J., Schreiber, S., Schughart, K., Schulz, H., Wolf, E., Wurst, W., Zimmer, A., Hrabé de Angelis, M., 2009. The German Mouse Clinic: a platform for systemic phenotype analysis of mouse models. *Curr. Pharm. Biotechnol.* 10, 236–243.
- Gentleman, R.C., Carey, V.J., Bates, D.M., others, 2004. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* 5, R80.
- Ghimenti, S., Tabucchi, S., Lomonaco, T., Di Francesco, F., Fuoco, R., Onor, M., Lenzi, S., Trivella, M.G., 2013. Monitoring breath during oral glucose tolerance tests. *J. Breath Res.* 7, 017115.
- Giannini, E.G., Fasoli, A., Borro, P., Botta, F., Malfatti, F., Fumagalli, A., Testa, E., Polegato, S., Cotellessa, T., Milazzo, S., Risso, D., Testa, R., 2005. ¹³C-galactose breath test and ¹³C-aminopyrine breath test for the study of liver function in chronic liver disease. *Clin. Gastroenterol. Hepatol.* 3, 279–285. doi:10.1016/S1542-3565(04)00720-7
- Gillooly, J.F., Brown, J.H., West, G.B., Savage, V.M., Charnov, E.L., 2001. Effects of Size and Temperature on Metabolic Rate. *Science* 293, 2248–2251. doi:10.1126/science.1061967
- Gregory, P.J., Sperry, M., Wilson, A.F., 2008. Dietary supplements for osteoarthritis. *Am. Fam. Physician* 77, 177–184.

- Greiter, M.B., Keck, L., Siegmund, T., Hoeschen, C., Oeh, U., Paretzke, H.G., 2010. Differences in exhaled gas profiles between patients with type 2 diabetes and healthy controls. *Diabetes Technol. Ther.* 12, 455–463. doi:10.1089/dia.2009.0181
- Guh, D.P., Zhang, W., Bansback, N., Amarsi, Z., Birmingham, C.L., Anis, A.H., 2009. The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health* 9, 88. doi:10.1186/1471-2458-9-88
- Hakim, M., Broza, Y.Y., Barash, O., Peled, N., Phillips, M., Amann, A., Haick, H., 2012. Volatile organic compounds of lung cancer and possible biochemical pathways. *Chem. Rev.* 112, 5949–5966. doi:10.1021/cr300174a
- Halbritter, S., Fedrigo, M., Höllriegl, V., Szymczak, W., Maier, J.M., Ziegler, A.-G., Hummel, M., 2012. Human Breath Gas Analysis in the Screening of Gestational Diabetes Mellitus. *Diabetes Technol. Ther.* doi:10.1089/dia.2012.0076
- Hall, K.D., Heymsfield, S.B., Kemnitz, J.W., Klein, S., Schoeller, D.A., Speakman, J.R., 2012. Energy balance and its components: implications for body weight regulation. *Am. J. Clin. Nutr.* 95, 989–994. doi:10.3945/ajcn.112.036350
- Herbig, J., Amann, A., 2009. Proton transfer reaction-mass spectrometry applications in medical research. *J. Breath Res.* 3, 020201. doi:10.1088/1752-7163/3/2/020201
- Herbig, J., Müller, M., Schallhart, S., Titzmann, T., Graus, M., Hansel, A., 2009. On-line breath analysis with PTR-TOF. *J. Breath Res.* 3, 027004. doi:10.1088/1752-7155/3/2/027004
- Jensen, M.N., Ritskes-Hoitinga, M., 2007. How isoflavone levels in common rodent diets can interfere with the value of animal models and with experimental results. *Lab. Anim.* 41, 1–18. doi:10.1258/00236770779399428
- Kaiyala, K.J., Morton, G.J., Leroux, B.G., Ogimoto, K., Wisse, B., Schwartz, M.W., 2010. Identification of Body Fat Mass as a Major Determinant of Metabolic Rate in Mice. *Diabetes* 59, 1657–1666. doi:10.2337/db09-1582
- Kaji, H., Hisamura, M., Saito, N., Sakai, H., Aikawa, T., Kondo, T., Ide, H., Murao, M., 1979. Clinical application of breath analysis for dimethyl sulfide following ingestion of DL-methionine. *Clin. Chim. Acta Int. J. Clin. Chem.* 93, 377–380.
- Kamel, R., Morsy, E.M.E., 2013. Hepatoprotective effect of methylsulfonylmethane against carbon tetrachloride-induced acute liver injury in rats. *Arch. Pharm. Res.* 36, 1140–1148. doi:10.1007/s12272-013-0110-x
- Karinje, K.U., Ogata, M., 1990. Methanol metabolism in acatalasemic mice. *Physiol. Chem. Phys. Med. NMR* 22, 193–198.
- Karpiévitch, Y.V., Hill, E.G., Leclerc, A.P., Dabney, A.R., Almeida, J.S., 2009. An Introspective Comparison of Random Forest-Based Classifiers for the Analysis of Cluster-Correlated Data by Way of RF++. *PLoS ONE* 4. doi:10.1371/journal.pone.0007087
- Katan, M.B., Ludwig, D.S., 2010. Extra calories cause weight gain—but how much? *Jama* 303, 65–66.
- King, J., Mochalski, P., Kupferthaler, A., Unterkofler, K., Koc, H., Filipiak, W., Teschl, S., Hinterhuber, H., Amann, A., 2010. Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. *Physiol. Meas.* 31, 1169. doi:10.1088/0967-3334/31/9/008
- Kistler, M., Muntean, A., Szymczak, W., Rink, N., Fuchs, H., Valerie Gailus-Durner, Wurst, W., Hoeschen, C., Klingenspor, M., Angelis, M.H. de, Jan Rozman, 2016. Diet-induced and mono-genetic obesity alter volatile organic compound signature in mice. *J. Breath Res.* 10, 016009. doi:10.1088/1752-7155/10/1/016009
- Kistler, M., Szymczak, W., Fedrigo, M., Fiamoncini, J., Höllriegl, V., Hoeschen, C., Klingenspor, M., Hrabě de Angelis, M., Rozman, J., 2014. Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice. *J. Breath Res.* 8, 016004. doi:10.1088/1752-7155/8/1/016004

- Klein, P.D., Malaty, H.M., Martin, R.F., Graham, K.S., Genta, R.M., Graham, D.Y., 1996. Noninvasive detection of *Helicobacter pylori* infection in clinical practice: the ¹³C urea breath test. *Am. J. Gastroenterol.* 91, 690–694.
- Kohl, I., Beauchamp, J., Cakar-Beck, F., Herbig, J., Dunkl, J., Tietje, O., Tiefenthaler, M., Boesmueller, C., Wisthaler, A., Breitenlechner, M., Langebner, S., Zabernigg, A., Reinstaller, F., Winkler, K., Gutmann, R., Hansel, A., 2013. First observation of a potential non-invasive breath gas biomarker for kidney function. *J. Breath Res.* 7, 017110. doi:10.1088/1752-7155/7/1/017110
- Kohsaka, A., Laposky, A.D., Ramsey, K.M., Estrada, C., Joshu, C., Kobayashi, Y., Turek, F.W., Bass, J., 2007. High-Fat Diet Disrupts Behavioral and Molecular Circadian Rhythms in Mice. *Cell Metab.* 6, 414–421. doi:10.1016/j.cmet.2007.09.006
- Krumsiek, J., Suhre, K., Evans, A.M., Mitchell, M.W., Mohny, R.P., Milburn, M.V., Wagele, B., Romisch-Margl, W., Illig, T., Adamski, J., Gieger, C., Theis, F.J., Kastenmuller, G., 2012. Mining the Unknown: A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. *PLoS Genet.* 8. doi:10.1371/journal.pgen.1003005
- Krumsiek, J., Suhre, K., Illig, T., Adamski, J., Theis, F.J., 2011. Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Syst. Biol.* 5, 21. doi:10.1186/1752-0509-5-21
- Lalazar, G., Adar, T., Ilan, Y., 2009. Point-of-care continuous ¹³C-methacetin breath test improves decision making in acute liver disease: Results of a pilot clinical trial. *World J. Gastroenterol.* WJG 15, 966–972. doi:10.3748/wjg.15.966
- Layden, B.T., Yalamanchi, S.K., Wolever, T.M., Dunaif, A., Lowe, W.L., 2012. Negative association of acetate with visceral adipose tissue and insulin levels. *Diabetes Metab. Syndr. Obes. Targets Ther.* 5, 49–55. doi:10.2147/DMSO.S29244
- Lay Jr., J.O., Liyanage, R., Borgmann, S., Wilkins, C.L., 2006. Problems with the “omics.” *TrAC Trends Anal. Chem., Use and abuse of chemometrics* 25, 1046–1056. doi:10.1016/j.trac.2006.10.007
- Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.-M., Kennedy, S., Leonard, P., Li, J., Burgdorf, K., Grarup, N., Jørgensen, T., Brandslund, I., Nielsen, H.B., Juncker, A.S., Bertalan, M., Levenez, F., Pons, N., Rasmussen, S., Sunagawa, S., Tap, J., Tims, S., Zoetendal, E.G., Brunak, S., Clément, K., Doré, J., Kleerebezem, M., Kristiansen, K., Renault, P., Sicheritz-Ponten, T., de Vos, W.M., Zucker, J.-D., Raes, J., Hansen, T., MetaHIT Consortium, Bork, P., Wang, J., Ehrlich, S.D., Pedersen, O., MetaHIT consortium additional Members, 2013. Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546. doi:10.1038/nature12506
- Lee, J., Ngo, J., Blake, D., Meinardi, S., Pontello, A.M., Newcomb, R., Galassetti, P.R., 2009. Improved predictive models for plasma glucose estimation from multi-linear regression analysis of exhaled volatile organic compounds. *J. Appl. Physiol.* 107, 155–160. doi:10.1152/jappphysiol.91657.2008
- Leopold, J.H., van Hooijdonk, R.T., Sterk, P.J., Abu-Hanna, A., Schultz, M.J., Bos, L.D., 2014. Glucose prediction by analysis of exhaled metabolites – a systematic review. *BMC Anesthesiol.* 14, 46. doi:10.1186/1471-2253-14-46
- Lewis, G.D., Laufman, A.K., McAnalley, B.H., Garriott, J.C., 1984. Metabolism of acetone to isopropyl alcohol in rats and humans. *J. Forensic Sci.* 29, 541–549.
- Li, J., Peng, Y., Liu, Y., Li, W., Jin, Y., Tang, Z., Duan, Y., 2014. Investigation of potential breath biomarkers for the early diagnosis of breast cancer using gas chromatography–mass spectrometry. *Clin. Chim. Acta* 436, 59–67. doi:10.1016/j.cca.2014.04.030
- Lindinger, W., Hansel, A., Jordan, A., 1998. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Process.* 173, 191–241.

- Lindinger, W., Taucher, J., Jordan, A., Hansel, A., Vogel, W., 1997. Endogenous Production of Methanol after the Consumption of Fruit. *Alcohol. Clin. Exp. Res.* 21, 939–943. doi:10.1111/j.1530-0277.1997.tb03862.x
- Lin, S., Thomas, T.C., Storlien, L.H., Huang, X.F., others, 2000. Development of high fat diet-induced obesity and leptin resistance in C 57 Bl/6 J mice. *Int. J. Obes.* 24, 639–646.
- Li, X., Martinez-Lozano Sinues, P., Dallmann, R., Bregy, L., Hollmén, M., Proulx, S., Brown, S.A., Detmar, M., Kohler, M., Zenobi, R., 2015. Drug Pharmacokinetics Determined by Real-Time Analysis of Mouse Breath. *Angew. Chem. Int. Ed.* 54, 7815–7818. doi:10.1002/anie.201503312
- Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., Croteau-Chonka, D.C., Esko, T., Fall, T., Ferreira, T., Gustafsson, S., Kutalik, Z., Luan, J., Mägi, R., Randall, J.C., Winkler, T.W., Wood, A.R., Workalemahu, T., Faul, J.D., Smith, J.A., Hua Zhao, J., Zhao, W., Chen, J., Fehrmann, R., Hedman, Å.K., Karjalainen, J., Schmidt, E.M., Absher, D., Amin, N., Anderson, D., Beekman, M., Bolton, J.L., Bragg-Gresham, J.L., Buyske, S., Demirkan, A., Deng, G., Ehret, G.B., Feenstra, B., Feitosa, M.F., Fischer, K., Goel, A., Gong, J., Jackson, A.U., Kanoni, S., Kleber, M.E., Kristiansson, K., Lim, U., Lotay, V., Mangino, M., Mateo Leach, I., Medina-Gomez, C., Medland, S.E., Nalls, M.A., Palmer, C.D., Pasko, D., Pechlivanis, S., Peters, M.J., Prokopenko, I., Shungin, D., Stančáková, A., Strawbridge, R.J., Ju Sung, Y., Tanaka, T., Teumer, A., Trompet, S., van der Laan, S.W., van Setten, J., Van Vliet-Ostaptchouk, J.V., Wang, Z., Yengo, L., Zhang, W., Isaacs, A., Albrecht, E., Ärnlöv, J., Arscott, G.M., Attwood, A.P., Bandinelli, S., Barrett, A., Bas, I.N., Bellis, C., Bennett, A.J., Berne, C., Blagieva, R., Blüher, M., Böhringer, S., Bonnycastle, L.L., Böttcher, Y., Boyd, H.A., Bruinenberg, M., Caspersen, I.H., Ida Chen, Y.-D., Clarke, R., Warwick Daw, E., de Craen, A.J.M., Delgado, G., Dimitriou, M., Doney, A.S.F., Eklund, N., Estrada, K., Eury, E., Folkersen, L., Fraser, R.M., Garcia, M.E., Geller, F., Giedraitis, V., Gigante, B., Go, A.S., Golay, A., Goodall, A.H., Gordon, S.D., Gorski, M., Grabe, H.-J., Grallert, H., Grammer, T.B., Gräßler, J., Grönberg, H., Groves, C.J., Gusto, G., Haessler, J., Hall, P., Haller, T., Hallmans, G., Hartman, C.A., Hassinen, M., Hayward, C., Heard-Costa, N.L., Helmer, Q., Hengstenberg, C., Holmen, O., Hottenga, J.-J., James, A.L., Jeff, J.M., Johansson, Å., Jolley, J., Juliusdottir, T., Kinnunen, L., Koenig, W., Koskenvuo, M., Kratzer, W., Laitinen, J., Lamina, C., Leander, K., Lee, N.R., Lichtner, P., Lind, L., Lindström, J., Sin Lo, K., Lobbens, S., Lorbeer, R., Lu, Y., Mach, F., Magnusson, P.K.E., Mahajan, A., McArdle, W.L., McLachlan, S., Menni, C., Merger, S., Mihailov, E., Milani, L., Moayyeri, A., Monda, K.L., Morcken, M.A., Mulas, A., Müller, G., Müller-Nurasyid, M., Musk, A.W., Nagaraja, R., Nöthen, M.M., Nolte, I.M., Pilz, S., Rayner, N.W., Renstrom, F., Rettig, R., Ried, J.S., Ripke, S., Robertson, N.R., Rose, L.M., Sanna, S., Scharnagl, H., Scholtens, S., Schumacher, F.R., Scott, W.R., Seufferlein, T., Shi, J., Vernon Smith, A., Smolonska, J., Stanton, A.V., Steinhorsdottir, V., Stirrups, K., Stringham, H.M., Sundström, J., Swertz, M.A., Swift, A.J., Syvänen, A.-C., Tan, S.-T., Tayo, B.O., Thorand, B., Thorleifsson, G., Tyrer, J.P., Uh, H.-W., Vandenput, L., Verhulst, F.C., Vermeulen, S.H., Verweij, N., Vonk, J.M., Waite, L.L., Warren, H.R., Waterworth, D., Weedon, M.N., Wilkens, L.R., Willenborg, C., Wilsgaard, T., Wojczynski, M.K., Wong, A., Wright, A.F., Zhang, Q., Brennan, E.P., Choi, M., Dastani, Z., Drong, A.W., Eriksson, P., Franco-Cereceda, A., Gådin, J.R., Gharavi, A.G., Goddard, M.E., Handsaker, R.E., Huang, J., Karpe, F., Kathiresan, S., Keildson, S., Kiryluk, K., Kubo, M., Lee, J.-Y., Liang, L., Lifton, R.P., Ma, B., McCarroll, S.A., McKnight, A.J., Min, J.L., Moffatt, M.F., Montgomery, G.W., Murabito, J.M., Nicholson, G., Nyholt, D.R., Okada, Y., Perry, J.R.B., Dorajoo, R., Reinmaa, E., Salem, R.M., Sandholm, N., Scott, R.A., Stolk, L., Takahashi, A., Tanaka, T., van't Hooft, F.M., Vinkhuyzen, A.A.E., Westra, H.-J., Zheng, W., Zondervan, K.T., Heath, A.C., Arveiler, D., Bakker, S.J.L., Beilby, J., Bergman, R.N., Blangero, J., Bovet, P., Campbell, H., Caulfield, M.J., Cesana, G., Chakravarti, A., Chasman, D.I., Chines, P.S., Collins, F.S., Crawford, D.C., Adrienne Cupples, L., Cusi, D., Danesh, J., de Faire, U., den Ruijter, H.M.,

- Dominiczak, A.F., Erbel, R., Erdmann, J., Eriksson, J.G., Farrall, M., Felix, S.B., Ferrannini, E., Ferrières, J., Ford, I., Forouhi, N.G., Forrester, T., Franco, O.H., Gansevoort, R.T., Gejman, P.V., Gieger, C., Gottesman, O., Gudnason, V., Gyllensten, U., Hall, A.S., Harris, T.B., Hattersley, A.T., Hicks, A.A., Hindorff, L.A., Hingorani, A.D., Hofman, A., Homuth, G., Kees Hovingh, G., Humphries, S.E., Hunt, S.C., Hyppönen, E., Illig, T., Jacobs, K.B., Jarvelin, M.-R., Jöckel, K.-H., Johansen, B., Jousilahti, P., Wouter Jukema, J., Jula, A.M., Kaprio, J., Kastelein, J.J.P., Keinänen-Kiukaanniemi, S.M., Kiemeny, L.A., Knekt, P., Kooner, J.S., Kooperberg, C., Kovacs, P., Kraja, A.T., Kumari, M., Kuusisto, J., Lakka, T.A., Langenberg, C., Le Marchand, L., Lehtimäki, T., Lyssenko, V., Männistö, S., Marette, A., Matisse, T.C., McKenzie, C.A., McKnight, B., Moll, F.L., Morris, A.D., Morris, A.P., Murray, J.C., Nelis, M., Ohlsson, C., Oldehinkel, A.J., Ong, K.K., Madden, P.A.F., Pasterkamp, G., Peden, J.F., Peters, A., Postma, D.S., Pramstaller, P.P., Price, J.F., Qi, L., Raitakari, O.T., Rankinen, T., Rao, D.C., Rice, T.K., Ridker, P.M., Rioux, J.D., Ritchie, M.D., Rudan, I., Salomaa, V., Samani, N.J., Saramies, J., Sarzynski, M.A., Schunkert, H., Schwarz, P.E.H., Sever, P., Shuldiner, A.R., Sinisalo, J., Stolk, R.P., Strauch, K., Tönjes, A., Trégouët, D.-A., Tremblay, A., Tremoli, E., Virtamo, J., Vohl, M.-C., Völker, U., Waeber, G., Willemsen, G., Witteman, J.C., Carola Zillikens, M., Adair, L.S., Amouyel, P., Asselbergs, F.W., Assimes, T.L., Bochud, M., Boehm, B.O., Boerwinkle, E., Bornstein, S.R., Bottinger, E.P., Bouchard, C., Cauchi, S., Chambers, J.C., Chanock, S.J., Cooper, R.S., de Bakker, P.I.W., Dedoussis, G., Ferrucci, L., Franks, P.W., Froguel, P., Groop, L.C., Haiman, C.A., Hamsten, A., Hui, J., Hunter, D.J., Hveem, K., Kaplan, R.C., Kivimäki, M., Kuh, D., Laakso, M., Liu, Y., Martin, N.G., März, W., Melbye, M., Metspalu, A., Moebus, S., Munroe, P.B., Njølstad, I., Oostra, B.A., Palmer, C.N.A., Pedersen, N.L., Perola, M., Pérusse, L., Peters, U., Power, C., Quertermous, T., Rauramaa, R., Rivadeneira, F., Saaristo, T.E., Saleheen, D., Sattar, N., Schadt, E.E., Schlessinger, D., Eline Slagboom, P., Snieder, H., Spector, T.D., Thorsteinsdottir, U., Stumvoll, M., Tuomilehto, J., Uitterlinden, A.G., Uusitupa, M., van der Harst, P., Walker, M., Wallaschofski, H., Wareham, N.J., Watkins, H., Weir, D.R., Wichmann, H.-E., Wilson, J.F., Zanen, P., Borecki, I.B., Deloukas, P., Fox, C.S., Heid, I.M., O'Connell, J.R., Strachan, D.P., Stefansson, K., van Duijn, C.M., Abecasis, G.R., Franke, L., Frayling, T.M., McCarthy, M.I., Visscher, P.M., Scherag, A., Willer, C.J., Boehnke, M., Mohlke, K.L., Lindgren, C.M., Beckmann, J.S., Barroso, I., North, K.E., Ingelsson, E., Hirschhorn, J.N., Loos, R.J.F., Speliotes, E.K., 2015. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–206. doi:10.1038/nature14177
- Mahendran, Y., Vangipurapu, J., Cederberg, H., Stančáková, A., Pihlajamäki, J., Soininen, P., Kangas, A.J., Paananen, J., Civelek, M., Saleem, N.K., Pajukanta, P., Lusi, A.J., Bonnycastle, L.L., Morken, M.A., Collins, F.S., Mohlke, K.L., Boehnke, M., Ala-Korpela, M., Kuusisto, J., Laakso, M., 2013. Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. *Diabetes*. doi:10.2337/db12-1363
- Marti, A., Corbalán, M.S., Forga, L., Martínez, J.A., Hinney, A., Hebebrand, J., 2003. A novel nonsense mutation in the melanocortin-4 receptor associated with obesity in a Spanish population. *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.* 27, 385–388. doi:10.1038/sj.ijo.0802244
- Mathur, R., Amichai, M., Chua, K.S., Mirocha, J., Barlow, G.M., Pimentel, M., 2013. Methane and Hydrogen Positivity on Breath Test Is Associated With Greater Body Mass Index and Body Fat. *J. Clin. Endocrinol. Metab.* 98, E698–E702. doi:10.1210/jc.2012-3144
- Matsuzawa-Nagata, N., Takamura, T., Ando, H., Nakamura, S., Kurita, S., Misu, H., Ota, T., Yokoyama, M., Honda, M., Miyamoto, K., Kaneko, S., 2008. Increased oxidative stress precedes the onset of high-fat diet-induced insulin resistance and obesity. *Metabolism* 57, 1071–1077. doi:10.1016/j.metabol.2008.03.010
- Maurer, F., Wolf, A., Fink, T., Rittershofer, B., Heim, N., Volk, T., Baumbach, J.I., Kreuer, S., 2014. Wash-out of ambient air contaminations for breath measurements. *J. Breath Res.* 8, 027107. doi:10.1088/1752-7155/8/2/027107

- McCurdy, M.R., Sharafkhaneh, A., Abdel-Monem, H., Rojo, J., Tittel, F.K., 2011. Exhaled nitric oxide parameters and functional capacity in chronic obstructive pulmonary disease. *J. Breath Res.* 5, 016003. doi:10.1088/1752-7155/5/1/016003
- McNelis, J.C., Lee, Y.S., Mayoral, R., Kant, R. van der, Johnson, A.M.F., Wollam, J., Olefsky, J.M., 2015. GPR43 potentiates beta cell function in obesity. *Diabetes* db141938. doi:10.2337/db14-1938
- Meinardi, S., Jin, K.-B., Barletta, B., Blake, D.R., Vaziri, N.D., 2013. Exhaled breath and fecal volatile organic biomarkers of chronic kidney disease. *Biochim. Biophys. Acta* 1830, 2531–2537. doi:10.1016/j.bbagen.2012.12.006
- Mendis, S., World Health Organization, 2014. Global status report on noncommunicable diseases 2014.
- Miekisch, W., Schubert, J.K., 2006. From highly sophisticated analytical techniques to life-saving diagnostics: Technical developments in breath analysis. *TrAC Trends Anal. Chem., On-site Instrumentation and Analysis* 25, 665–673. doi:10.1016/j.trac.2006.05.006
- Minh, T.D.C., Oliver, S.R., Flores, R.L., Ngo, J., Meinardi, S., Carlson, M.K., Midyett, J., Rowland, F.S., Blake, D.R., Galassetti, P.R., 2012. Noninvasive Measurement of Plasma Triglycerides and Free Fatty Acids from Exhaled Breath. *J. Diabetes Sci. Technol.* 6, 86–101.
- Minh, T.D.C., Oliver, S.R., Ngo, J., Flores, R., Midyett, J., Meinardi, S., Carlson, M.K., Rowland, F.S., Blake, D.R., Galassetti, P.R., 2011. Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects. *Am. J. Physiol. Endocrinol. Metab.* 300, E1166–1175. doi:10.1152/ajpendo.00634.2010
- Mochalski, P., Al-Zoairy, R., Niederwanger, A., Unterkofler, K., Amann, A., 2014. Quantitative analysis of volatile organic compounds released and consumed by rat L6 skeletal muscle cells *in vitro*. *J. Breath Res.* 8, 046003. doi:10.1088/1752-7155/8/4/046003
- Moghe, A., Ghare, S., Lamoreau, B., Mohammad, M., Barve, S., McClain, C., Joshi-Barve, S., 2015. Molecular Mechanisms of Acrolein Toxicity: Relevance to Human Disease. *Toxicol. Sci.* 143, 242–255. doi:10.1093/toxsci/kfu233
- Moraal, M., Leenaars, P.P. a. M., Arnts, H., Smeets, K., Savenije, B.S., Curfs, J.H. a. J., Ritskes-Hoitinga, M., 2012. The influence of food restriction versus ad libitum feeding of chow and purified diets on variation in body weight, growth and physiology of female Wistar rats. *Lab. Anim.* 46, 101–107. doi:10.1258/la.2011.011011
- Morin, A.M., Liss, M., 1973. Evidence for a methylated protein intermediate in pituitary methanol formation. *Biochem. Biophys. Res. Commun.* 52, 373–378. doi:10.1016/0006-291X(73)90721-3
- Morisco, F., Aprea, E., Lembo, V., Fogliano, V., Vitaglione, P., Mazzone, G., Cappellin, L., Gasperi, F., Masone, S., De Palma, G.D., Marmo, R., Caporaso, N., Biasioli, F., 2013. Rapid “Breath-Print” of Liver Cirrhosis by Proton Transfer Reaction Time-of-Flight Mass Spectrometry. A Pilot Study. *PloS One* 8, e59658. doi:10.1371/journal.pone.0059658
- Müller, P.D.J., 1898. Ueber die Ausscheidungsstätten des Acetons und die Bestimmung desselben in der Athemluft und den Hautausdünstungen des Menschen. *Arch. Für Exp. Pathol. Pharmacol.* 40, 351–362. doi:10.1007/BF01825265
- Munson, M.S.B., Field, F.H., 1966. Chemical Ionization Mass Spectrometry. I. General Introduction. *J. Am. Chem. Soc.* 88, 2621–2630. doi:10.1021/ja00964a001
- Musa-Veloso, K., Likhodii, S.S., Cunnane, S.C., 2002. Breath acetone is a reliable indicator of ketosis in adults consuming ketogenic meals. *Am. J. Clin. Nutr.* 76, 65–70.
- Musa-Veloso, K., Likhodii, S.S., Rarama, E., Benoit, S., Liu, Y.-M.C., Chartrand, D., Curtis, R., Carmant, L., Lortie, A., Comeau, F.J.E., Cunnane, S.C., 2006. Breath acetone predicts plasma ketone bodies in children with epilepsy on a ketogenic diet. *Nutr. Burbank Los Angel. Cty. Calif* 22, 1–8. doi:10.1016/j.nut.2005.04.008
- Myers, M.G., Münzberg, H., Leininger, G.M., Leshan, R.L., 2009. The geometry of leptin action in the brain: more complicated than a simple ARC. *Cell Metab.* 9, 117–123. doi:10.1016/j.cmet.2008.12.001

- Neuhaus, S., Seifert, L., Vautz, W., Nolte, J., Bufe, A., Peters, M., 2011. Comparison of metabolites in exhaled breath and bronchoalveolar lavage fluid samples in a mouse model of asthma. *J. Appl. Physiol.* 111, 1088–1095. doi:10.1152/jappphysiol.00476.2011
- Nicklas, W., Baneux, P., Boot, R., Decelle, T., Deeny, A.A., Fumanelli, M., Illgen-Wilcke, B., FELASA (Federation of European Laboratory Animal Science Associations Working Group on Health Monitoring of Rodent and Rabbit Colonies), 2002. Recommendations for the health monitoring of rodent and rabbit colonies in breeding and experimental units. *Lab. Anim.* 36, 20–42.
- Novak, B.J., Blake, D.R., Meinardi, S., Rowland, F.S., Pontello, A., Cooper, D.M., Galassetti, P.R., 2007. Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proc. Natl. Acad. Sci.* 104, 15613–15618. doi:10.1073/pnas.0706533104
- Opgen-Rhein, R., Strimmer, K., 2006. Inferring gene dependency networks from genomic longitudinal data: a functional data approach. *RevStat* 4, 53–65.
- Panaro, B.L., Tough, I.R., Engelstoft, M.S., Matthews, R.T., Digby, G.J., Møller, C.L., Svendsen, B., Gribble, F., Reimann, F., Holst, J.J., Holst, B., Schwartz, T.W., Cox, H.M., Cone, R.D., 2014. The Melanocortin-4 Receptor Is Expressed in Enteroendocrine L Cells and Regulates the Release of Peptide YY and Glucagon-like Peptide 1 In Vivo. *Cell Metab.* 20, 1018–1029. doi:10.1016/j.cmet.2014.10.004
- Paredi, P., Kharitonov, S.A., Barnes, P.J., 2000. Elevation of exhaled ethane concentration in asthma. *Am. J. Respir. Crit. Care Med.* 162, 1450–1454. doi:10.1164/ajrccm.162.4.2003064
- Pauling, L., Robinson, A.B., Teranishi, R., Cary, P., 1971. Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography. *Proc. Natl. Acad. Sci.* 68, 2374–2376.
- Pearson, T.W., Dawson, H.J., Lackey, H.B., 1981. Natural occurring levels of dimethyl sulfoxide in selected fruits, vegetables, grains, and beverages. *J. Agric. Food Chem.* 29, 1089–1091.
- Pellizzari, E.D., Wallace, L.A., Gordon, S.M., 1992. Elimination kinetics of volatile organics in humans using breath measurements. *J Expo Anal Env. Epidemiol* 2, 341–356.
- Peng, G., Hakim, M., Broza, Y.Y., Billan, S., Abdah-Bortnyak, R., Kuten, A., Tisch, U., Haick, H., 2010. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. *Br J Cancer* 103, 542–551.
- Peng, X., Li, S., Luo, J., Wu, X., Liu, L., 2013. Effects of dietary fibers and their mixtures on short chain fatty acids and microbiota in mice guts. *Food Funct.* 4, 932–938. doi:10.1039/C3FO60052A
- Petersen, T.H., Williams, T., Nuwayhid, N., Harruff, R., 2012. Postmortem Detection of Isopropanol in Ketoacidosis: POSTMORTEM DETECTION OF ISOPROPANOL IN KETOACIDOSIS. *J. Forensic Sci.* 57, 674–678. doi:10.1111/j.1556-4029.2011.02045.x
- Pinheiro, J., Bates, D., DebRoy, S., Deepayan, S., R-core, 2015. nlme: Linear and Nonlinear Mixed Effects Models. <http://cran.us.r-project.org/web/packages/nlme/index.html>.
- Ploner, A., 2014. Heatplus: Heatmaps with row and/or column covariates and colored clusters. <http://bioconductor.org/packages/Heatplus/>.
- Qiao, Y., Gao, Z., Liu, Y., Cheng, Y., Yu, M., Zhao, L., Duan, Y., Liu, Y., 2014. Breath Ketone Testing: A New Biomarker for Diagnosis and Therapeutic Monitoring of Diabetic Ketosis. *BioMed Res. Int.* 2014. doi:10.1155/2014/869186
- Qin, T., Liu, H., Song, Q., Song, G., Wang, H., Pan, Y., Xiong, F., Gu, K., Sun, G., Chen, Z., 2010. The screening of volatile markers for hepatocellular carcinoma. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 19, 2247–2253. doi:10.1158/1055-9965.EPI-10-0302
- Qu, H., Li, J., Chen, W., Li, Y., Jiang, Q., Jiang, H., Huo, J., Zhao, Z., Liu, B., Zhang, Q., 2014. Differential expression of the melanocortin-4 receptor in male and female C57BL/6J mice. *Mol. Biol. Rep.* 41, 3245–3256. doi:10.1007/s11033-014-3187-5
- Raman, M., Ahmed, I., Gillevet, P.M., Probert, C.S., Ratcliffe, N.M., Smith, S., Greenwood, R., Sikaroodi, M., Lam, V., Crotty, P., Bailey, J., Myers, R.P., Rioux, K.P., 2013. Fecal Microbiome and Volatile Organic Compound Metabolome in Obese Humans With

- Nonalcoholic Fatty Liver Disease. *Clin. Gastroenterol. Hepatol.* 11, 868–875.e3. doi:10.1016/j.cgh.2013.02.015
- R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Romano, A., Fischer, L., Herbig, J., Campbell-Sills, H., Coulon, J., Lucas, P., Cappellin, L., Biasioli, F., 2014. Wine analysis by FastGC proton-transfer reaction-time-of-flight-mass spectrometry. *Int. J. Mass Spectrom.* 369, 81–86. doi:10.1016/j.ijms.2014.06.006
- Rudolf, J.L., Bauerly, K.A., Tchapanian, E., Rucker, R.B., Mitchell, A.E., 2008. The influence of diet composition on phase I and II biotransformation enzyme induction. *Arch. Toxicol.* 82, 893–901. doi:10.1007/s00204-008-0310-1
- Rumessen, J.J., 1992. Hydrogen and methane breath tests for evaluation of resistant carbohydrates. *Eur. J. Clin. Nutr.* 46 Suppl 2, S77–90.
- Ruzsanyi, V., Lederer, W., Seger, C., Calenic, B., Liedl, K.R., Amann, A., 2014. Non-¹³C₂ targeted breath tests: a feasibility study. *J. Breath Res.* 8, 046005. doi:10.1088/1752-7155/8/4/046005
- Sakakibara, I., Fujino, T., Ishii, M., Tanaka, T., Shimosawa, T., Miura, S., Zhang, W., Tokutake, Y., Yamamoto, J., Awano, M., Iwasaki, S., Motoike, T., Okamura, M., Inagaki, T., Kita, K., Ezaki, O., Naito, M., Kuwaki, T., Chohnan, S., Yamamoto, T.T., Hammer, R.E., Kodama, T., Yanagisawa, M., Sakai, J., 2009. Fasting-Induced Hypothermia and Reduced Energy Production in Mice Lacking Acetyl-CoA Synthetase 2. *Cell Metab.* 9, 191–202. doi:10.1016/j.cmet.2008.12.008
- Schaefer, J., Opgen-Rhein, R., Strimmer, K., 2015. GeneNet: Modeling and Inferring Gene Networks. <http://CRAN.R-project.org/package=GeneNet>.
- Schäfer, J., Strimmer, K., 2005. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* 4.
- Schwarzkopf, T.M., Horn, T., Lang, D., Klein, J., 2013. Blood gases and energy metabolites in mouse blood before and after cerebral ischemia: the effects of anesthetics. *Exp. Biol. Med.* Maywood NJ 238, 84–89. doi:10.1258/ebm.2012.012261
- Shearer, J., Duggan, G., Weljie, A., Hittel, D.S., Wasserman, D.H., Vogel, H.J., 2008. Metabolomic profiling of dietary-induced insulin resistance in the high fat–fed C57BL/6J mouse. *Diabetes Obes. Metab.* 10, 950–958. doi:10.1111/j.1463-1326.2007.00837.x
- Simpson, J.L., Wark, P.A., 2008. The role of exhaled nitric oxide and exhaled breath condensates in evaluating airway inflammation in asthma. *Expert Opin. Med. Diagn.* 2, 607–620. doi:10.1517/17530059.2.6.607
- Siragusa, R.J., Cerda, J.J., Baig, M.M., Burgin, C.W., Robbins, F.L., 1988. Methanol production from the degradation of pectin by human colonic bacteria. *Am. J. Clin. Nutr.* 47, 848–851.
- Smolinska, A., Hauschild, A.-C., Fijten, R.R.R., Dallinga, J.W., Baumbach, J., Schooten, F.J. van, 2014. Current breathomics—a review on data pre-processing techniques and machine learning in metabolomics breath analysis. *J. Breath Res.* 8, 027105. doi:10.1088/1752-7155/8/2/027105
- Solga, S.F., Mudalel, M., Spacek, L.A., Lewicki, R., Tittel, F.K., Loccioni, C., Russo, A., Ragnoni, A., Risby, T.H., 2014. Changes in the concentration of breath ammonia in response to exercise: a preliminary investigation. *J. Breath Res.* 8, 037103. doi:10.1088/1752-7155/8/3/037103
- Solga, S.F., Mudalel, M., Spacek, L.A., Lewicki, R., Tittel, F., Loccioni, C., Russo, A., Risby, T.H., 2013. Factors influencing breath ammonia determination. *J. Breath Res.* 7, 037101. doi:10.1088/1752-7155/7/3/037101
- Sonnenburg, E.D., Smits, S.A., Tikhonov, M., Higginbottom, S.K., Wingreen, N.S., Sonnenburg, J.L., 2016. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 529, 212–215. doi:10.1038/nature16504
- Španěl, P., Dryahina, K., Rejšková, A., Chippendale, T.W.E., Smith, D., 2011. Breath acetone concentration; biological variability and the influence of diet. *Physiol. Meas.* 32, N23. doi:10.1088/0967-3334/32/8/N01

- Španěl, P., Dryahina, K., Smith, D., 2013. A quantitative study of the influence of inhaled compounds on their concentrations in exhaled breath. *J. Breath Res.* 7, 017106. doi:10.1088/1752-7155/7/1/017106
- Steeghs, M.M., Cristescu, S.M., Harren, F.J., 2007. The suitability of Tedlar bags for breath sampling in medical diagnostic research. *Physiol. Meas.* 28, 73.
- Szymczak, W., Rozman, J., Höllriegel, V., Kistler, M., Keller, S., Peters, D., Kneipp, M., Schulz, H., Hoeschen, C., Klingenspor, M., de Angelis, M.H., 2014. Online breath gas analysis in unrestrained mice by hs-PTR-MS. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* 25, 129–140. doi:10.1007/s00335-013-9493-8
- Tang, C., Ahmed, K., Gille, A., Lu, S., Gröne, H.-J., Tunaru, S., Offermanns, S., 2015. Loss of FFA2 and FFA3 increases insulin secretion and improves glucose tolerance in type 2 diabetes. *Nat. Med.* doi:10.1038/nm.3779
- Tani, A., Hayward, S., Hewitt, C.N., 2003. Measurement of monoterpenes and related compounds by proton transfer reaction-mass spectrometry (PTR-MS). *Int. J. Mass Spectrom., In Memoriam of Werner Lindinger* 223–224, 561–578. doi:10.1016/S1387-3806(02)00880-1
- Tassopoulos, C.N., Barnett, D., Russell Fraser, T., 1969. BREATH-ACETONE AND BLOOD-SUGAR MEASUREMENTS IN DIABETES. *The Lancet* 293, 1282–1286. doi:10.1016/S0140-6736(69)92222-3
- Temelkova-Kurktschiev, T., Stefanov, T., 2012. Lifestyle and Genetics in Obesity and type 2 Diabetes. *Exp. Clin. Endocrinol. Amp Diabetes* 120, 1–6. doi:10.1055/s-0031-1285832
- Thekedar, B., Szymczak, W., Höllriegel, V., Hoeschen, C., Oeh, U., 2009. Investigations on the variability of breath gas sampling using PTR-MS. *J. Breath Res.* 3, 027007. doi:10.1088/1752-7155/3/2/027007
- Tiengo, A., Valerio, A., Molinari, M., Meneghel, A., Lapolla, A., 1981. Effect of ethanol, acetaldehyde, and acetate on insulin and glucagon secretion in the perfused rat pancreas. *Diabetes* 30, 705–709.
- Tschöp, M.H., Speakman, J.R., Arch, J.R.S., Auwerx, J., Brüning, J.C., Chan, L., Eckel, R.H., Farese Jr, R.V., Galgani, J.E., Hambly, C., Herman, M.A., Horvath, T.L., Kahn, B.B., Kozma, S.C., Maratos-Flier, E., Müller, T.D., Münzberg, H., Pfluger, P.T., Plum, L., Reitman, M.L., Rahmouni, K., Shulman, G.I., Thomas, G., Kahn, C.R., Ravussin, E., 2012. A guide to analysis of mouse energy metabolism. *Nat. Methods* 9, 57–63. doi:10.1038/nmeth.1806
- Turner, C., Španěl, P., Smith, D., 2006. A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol. Meas.* 27, 637–648. doi:10.1088/0967-3334/27/7/007
- Uchida, K., Kanematsu, M., Morimitsu, Y., Osawa, T., Noguchi, N., Niki, E., 1998. Acrolein Is a Product of Lipid Peroxidation Reaction Formation of free acrolein and its conjugate with lysine residues in oxidized low density lipoproteins. *J. Biol. Chem.* 273, 16058–16066. doi:10.1074/jbc.273.26.16058
- Urrea, V., Calle, M.L., 2012. AUCCRF: Variable Selection with Random Forest and the Area Under the Curve. <http://cran.us.r-project.org/web/packages/AUCCRF/index.html>.
- Van den Velde, S., Nevens, F., Van Hee, P., van Steenberghe, D., Quirynen, M., 2008. GC-MS analysis of breath odor compounds in liver patients. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* 875, 344–348. doi:10.1016/j.jchromb.2008.08.031
- Van der Ploeg, L.H.T., Martin, W.J., Howard, A.D., Nargund, R.P., Austin, C.P., Guan, X., Drisko, J., Cashen, D., Sebat, I., Patchett, A.A., Figueroa, D.J., DiLella, A.G., Connolly, B.M., Weinberg, D.H., Tan, C.P., Palyha, O.C., Pong, S.-S., MacNeil, T., Rosenblum, C., Vongs, A., Tang, R., Yu, H., Sailer, A.W., Fong, T.M., Huang, C., Tota, M.R., Chang, R.S., Stearns, R., Tamvakopoulos, C., Christ, G., Drazen, D.L., Spar, B.D., Nelson, R.J., MacIntyre, D.E., 2002. A role for the melanocortin 4 receptor in sexual function. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11381–11386. doi:10.1073/pnas.172378699

- Vautz, W., Nolte, J., Bufer, A., Baumbach, J.I., Peters, M., 2010. Analyses of mouse breath with ion mobility spectrometry: a feasibility study. *J. Appl. Physiol.* 108, 697–704. doi:10.1152/jappphysiol.00658.2009
- Vitali, B., Ndagijimana, M., Cruciani, F., Carnevali, P., Candela, M., Guerzoni, M.E., Brigidi, P., 2010. Impact of a synbiotic food on the gut microbial ecology and metabolic profiles. *BMC Microbiol.* 10, 4. doi:10.1186/1471-2180-10-4
- Wellen, K.E., Hotamisligil, G.S., 2003. Obesity-induced inflammatory changes in adipose tissue. *J. Clin. Invest.* 112, 1785–1788. doi:10.1172/JCI200320514
- Wickham, H., 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wolever, T., Josse, R.G., Leiter, L.A., Chiasson, J.-L., 1997. Time of day and glucose tolerance status affect serum short-chain fatty concentrations in humans. *Metabolism* 46, 805–811.
- Wolf, A., Baumbach, J.I., Kleber, A., Maurer, F., Maddula, S., Favrod, P., Jang, M., Fink, T., Volk, T., Kreuer, S., 2014. Multi-capillary column-ion mobility spectrometer (MCC-IMS) breath analysis in ventilated rats: a model with the feasibility of long-term measurements. *J. Breath Res.* 8, 016006. doi:10.1088/1752-7155/8/1/016006
- Wong, C.K., Botta, A., Pither, J., Dai, C., Gibson, W.T., Ghosh, S., 2015. A high-fat diet rich in corn oil reduces spontaneous locomotor activity and induces insulin resistance in mice. *J. Nutr. Biochem.* 26, 319–326. doi:10.1016/j.jnutbio.2014.11.004
- Yamashita, H., Kaneyuki, T., Tagawa, K., 2001. Production of acetate in the liver and its utilization in peripheral tissues. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* 1532, 79–87. doi:10.1016/S1388-1981(01)00117-2
- Zarrinpar, A., Chaix, A., Yooseph, S., Panda, S., 2014. Diet and Feeding Pattern Affect the Diurnal Dynamics of the Gut Microbiome. *Cell Metab.* 20, 1006–1017. doi:10.1016/j.cmet.2014.11.008
- Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., Friedman, J.M., 1994. Positional cloning of the mouse obese gene and its human homologue. *Nature* 372, 425–432. doi:10.1038/372425a0
- Zhu, J., Bean, H.D., Jiménez-Díaz, J., Hill, J.E., 2013a. Secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting of multiple bacterial lung pathogens, a mouse model study. *J. Appl. Physiol.* 114, 1544–1549. doi:10.1152/jappphysiol.00099.2013
- Zhu, J., Jiménez-Díaz, J., Bean, H.D., Daphary, N.A., Aliyeva, M.I., Lundblad, L.K.A., Hill, J.E., 2013b. Robust detection of *P. aeruginosa* and *S. aureus* acute lung infections by secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting: from initial infection to clearance. *J. Breath Res.* 7, 037106. doi:10.1088/1752-7155/7/3/037106

7. Acknowledgments

I would like to thank numerous people for their support and guidance throughout my PhD Thesis.

Prof. Martin Hrabě de Angelis for giving me the opportunity to work on this project and his feedback, which helped a lot to focus my work.

Prof. Martin Klingenspor for his valuable supervision and his support during this thesis.

Dr. Jan Rozman for his open ear during the everyday business, the fruitful discussions and for granting me freedom in exploring the volatile world.

Dr. Wilfried Szymczak for sharing his experience and his keen sense for technical questions, from which I learned a lot.

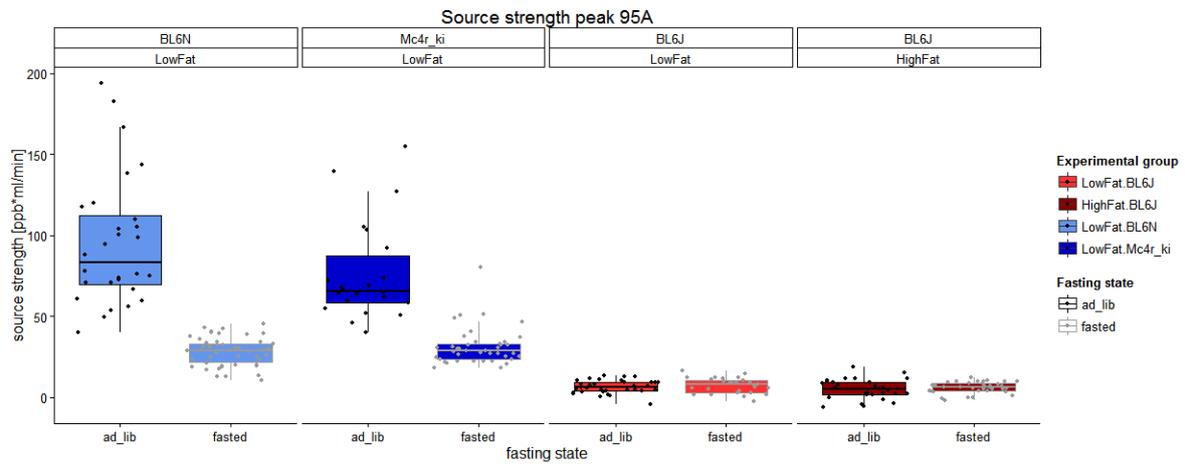
Ann-Elisabeth Schwarz, Stefan Keller, Garlich Fischbeck, Julia Günther, Andreea Muntean, Tobias Hopfner, Lilly Vargas Panesso, Nirav Chhabra and all other people from my microcosm for open ears to for discuss science and just random stuff and for the nice working atmosphere.

Dr. Helmut Fuchs, Dr. Valerie Gailus-Durner and the whole GMC-Team for the kind support whenever needed.

My friends and my family for supporting me in busy times and create meaning and motivation for what I do.

8. Appendix

8.1. Supplementary figure



Supplementary figure 1: DMSO₂ in chow and purified diet fed mice

8.2. Publication I

Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice

M. Kistler^{1*}, W. Szymczak², M. Fedrigo², J. Fiamoncini⁵, V. Höllriegel², C. Hoeschen², M. Klingenspor³, M. Hrabě de Angelis^{1,4,6} and J. Rozman^{1,3,6*}

¹ Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), Institute of Experimental Genetics, German Mouse Clinic, 85764 Neuherberg, Germany

² Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), Research Unit Medical Radiation Physics and Diagnostics, 85764 Neuherberg, Germany

³ ZIEL Department of Molecular Nutritional Medicine, Else Kröner-Fresenius Center, Technische Universität München, 85350 Freising, Germany

⁴ Chair for Experimental Genetics, Life and Food Science Center Weihenstephan, Technische Universität Munich, 85354 Freising-Weihenstephan, Germany

⁵ Department of Physiology and Biophysics, Institute of Biomedical Sciences, University of São Paulo, São Paulo – SP, Brazil

⁶ Member of German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany

*corresponding authors: Martin Kistler (martin.kistler@helmholtz-muenchen.de) and Jan Rozman (jan.rozman@helmholtz-muenchen.de)

Contributions: Conceived and designed research: MKi, WS, CH, MK, MHA; JR. Contributed reagents/materials/analysis tools: MKi, WS, JF, VH. Performed experiments: MKi, WS, JF. Analyzed data: MKi, WS, MF. Wrote manuscript: MKi, WS, MF, JR. Read and approved manuscript: all authors.

Keywords: Volatile organic compounds, proton transfer reaction mass spectrometry (PTR-MS), diet-induced obesity, high-fat diet, mouse breath, rodent diet

8.2.1. Abstract

Breath gas analysis in humans proved successful in identifying disease states and assessing metabolic functions in a non-invasive way. While many studies report diagnostic capability using volatile organic compounds (VOC) in breath, the inter-individual variability even in healthy human cohorts is rather large and not completely understood in its biochemical origin. Laboratory mice are the predominant animal model system for human disorders and are analyzed under highly standardized and controlled conditions. We established a novel setup to monitor VOCs as biomarkers for disease in the breath gas of non-anesthetized, non-restrained mice using a proton transfer reaction mass spectrometer with time of flight detection (PTR-TOF-MS). In this study we implemented breath gas analysis in a dietary intervention study in C57BL/6J mice with the aim to assess the variability in VOC signatures due to a change in diet matrix. Mice were fed a standard laboratory chow and then exposed to four semi-purified low- or high-fat diets for four weeks. Random forest (RF++) was used to identify VOCs that specifically respond to the diet matrix change. Interestingly, we found that the change from a chow diet to semi-purified diets resulted in a considerable drop of several VOC levels. Our results suggest that diet matrix impacts VOC signatures and the underlying metabolic functions and may be one source of variability in exhaled volatiles.

8.2.2. Introduction

The analysis of volatile organic compounds (VOCs) in human breath as a non-invasive tool to diagnose and monitor various diseases recently attracted increasing attention. Technological advancements combined with mass spectrometric (MS) measurement as the selected ion flow tube - MS [1], ion mobility-MS [2], proton transfer-MS [3,4] and pre-concentration methods in combination with gas chromatography - mass spectrometry (GC-MS) [5] led to a high sensitivity in the analysis of VOCs. Analysis of VOCs even in low concentrations in the range of parts per billion allows the access to a large quantity of molecules that may provide additional information beyond blood, urine or tissue based metabolomics techniques or classical gas exchange analysis of carbon dioxide and oxygen concentrations for indirect calorimetry. Several studies using VOCs exhaled with breath as a diagnostic procedure were conducted [6], for example in the fields of cancer [7], liver cirrhosis [8], pulmonary diseases [9,10] and systemic infections [11]. Despite the success in identifying diagnostic targets, volatiles detected in breath vary notably even between age-matched healthy humans [12]. This inter-individual variability in breath VOCs is not fully understood yet and is thought to consist of several components. Volatiles measured in breath may originate from enzymatic activity in tissues or non-enzymatic reactions depending on the chemical state of the organism (e.g. pH or redox state) [13]. Furthermore, the complex endogenous variability may be difficult to be distinguished from environmental wash-in contributions [14]. A further contribution from mouth or gut microbial metabolism and food ingredients is also known [15,16] and in some cases even dominates the endogenous emission [17]. This variability is difficult to control for in a human studies. In contrast to human cohorts, inbred strains of laboratory mice (*Mus musculus*) are housed under standardized conditions and can be fed highly defined diets. Under such controlled conditions, specific effects of the diet on the emission of VOCs can be studied avoiding confounding factors attributed to other environmental factors. In the field of energy metabolism and associated diseases, diets with an elevated relative fat content are frequently used to model “western-style” energy rich human diet-induced obesity in mice. Typically, in those dietary studies not only macronutrient composition of the diet is varied but the matrix of the diet also changes from raw material-based laboratory chow diets to highly specific semi-purified diets containing refined components. In this study, we aim to particularly explore the impact of diet matrix on VOC signatures of laboratory mice superimposing effects of a switch of macro-nutrient composition.

8.2.3. Methods

8.2.3.1. Animals

Two cohorts of male C57BL/6J mice (n=40, age 12 weeks, 10 per diet group) were obtained either directly from Charles River Laboratories (Sulzfeld, Germany, Lard groups) or from in-house breeding from C57BL/6J founder animals (tallow/soy based diet groups). Mice were housed in isolated ventilated cages at a temperature of 24 +/- 1°C. A 12:12h light/dark cycle and air humidity of 50 to 60% was applied. Animals were located in the German Mouse Clinic [18] and kept under specific pathogen-free conditions in type II polycarbonate cages in individually ventilated caging systems (IVC). Mice were maintained in groups of up to four animals (tallow/soy oil based diets) or were singly housed after arrival (Lard/soy oil based diets). Wood shavings were used as bedding (Altromin GmbH, Germany); if singly housed, tissue paper and a mouse house were supplied. Mice were weighed every week to the nearest 0.1 g and body composition was monitored by non-invasive qNMR scans (Bruker Minispec LF50 body composition analyzer, Ettlingen, Germany). All experiments were performed following animal welfare regulations with permission from the district government of Upper Bavaria (Regierung von Oberbayern).

8.2.3.2. Experimental diets

Mice had ad libitum access to a pelleted standard laboratory chow diet (no. 1314, Altromin, Lage, Germany) and drinking water from weaning on. After a baseline breath gas analysis at the age of 12 weeks, diets were switched to semi-purified diets containing low amount of soy oil versus a high amount of soy oil and tallow fat (LF, HF) or low versus high amount of soy oil and lard (LL, HL). The low fat soy oil-based diet was pelleted and semi-purified (LF, E 15000-04, Ssniff, Soest Germany). The corresponding high fat diet was a pelleted diet with 60% of the energy coming from fat (HF, E 15741-347, Ssniff, Soest, Germany). A third group of mice (LL) was fed a low fat diet with 90% lard and 10% soybean oil in the dietary fat. The fourth group of mice (HL) was fed a high-fat diet with the same ratio of lard to soybean oil in the lipid fraction (see Table 1 for an overview of the used diets). Lard was added to pre-mixtures directly in the animal facility. All diets and the respective separate ingredients were sterilized with 25 kGy/min γ -radiation (Isotron Deutschland - Allershausen, Germany) before import into the facility. Diets were refreshed on weekly (pelleted) or twice a week basis (self-mixed).

Table 3. Diet composition related to percentages of metabolizable energy content (estimated from the conversion of weight percentages using Atwater factors).

Diet	Carbohydrates	Protein	Fat	Energy content [kJ/g]	Matrix
Laboratory chow	60 en%	27 en%	13 en%	12.5	Grain-based
Low soy oil diet (LF)	66 en%	23 en%	11 en%	15.0	Semi-purified
High soy oil/tallow fat diet (HF)	21 en%	19 en%	60 en%	21.4	Semi-purified
Low soy oil/lard diet (LL)	74 en%	16 en%	10 en%	15.6	Semi-purified
High soy oil/lard diet (HL)	25 en %	15 en%	60 en%	22.3	Semi-purified

8.2.3.3. Proton-transfer time-of-flight mass spectrometry

The PTR-TOF2000 MS is a high sensitivity PTR-MS (e.g. benzene 100 cps/ppbv) with a medium mass resolution of $\Delta m/m \leq 2000$. The PTR-MS principle and the implementation of a TOF-mass analyser is well documented [4,19]. The sensitivity and the sub-ppbV limits of detection in combination with short integration times of the full mass range comply with the standards in on-line breath analysis in mice.

The PTR-MS was operated with a drift temperature of 80 °C, drift voltage of 600 V and drift pressure 2.3 mbar. A mass range from m/z 0 to 349.5 was recorded with a repetition rate of 77 kHz and the sum spectra integrated over 3s stored (TOF-DAQ, ToFwerk AG, Switzerland).

The stored sum spectra were analyzed using TOFViewer (Version 3.0.1, Ionicon analytic GmbH, Innsbruck, Austria). Spectra were mass calibrated internally using $m/z=29.9971$ (NO) and 59.0491 ($C_3H_7O^+$; protonated acetone). Peak borders of 164 peaks in the TOF spectrum were set manually. The deconvolution of overlapping, multiple peaks was performed using a gauss-based fit in TOFViewer. Peak concentrations in ppbV were calculated using the semi-quantitative estimation

formula [4] with a constant k-rate of 2. The performance of the PTR-TOF2000 was routinely controlled by transmission measurement. With a gas calibration unit (GCU, Ionicon Analytic GmbH, Innsbruck, Austria) a mixture of substances (VOC gas standard, Ionicon Analytic GmbH, Innsbruck, Austria) with a stepwise changing concentration was fed into the PTR-MS inlet. From the resulting linear calibration curve the instrumental sensitivity for each standard compound was determined and the transmission factor calculated.

8.2.3.4. Setup for real-time breath gas analysis in unrestrained mice

The accumulation of exhaled VOCs of an unrestrained mouse was monitored in the headspace inside a respirometry chamber as previously described [20]. In brief, a polypropylene box (volume 600 mL) was connected to a proton transfer time-of-flight mass spectrometer (PTR-TOF2000-MS, Ionicon Analytic GmbH, Innsbruck, Austria) and face-to-face to a gas supply of synthetic air (20% oxygen, 80% nitrogen, concentration of hydrocarbons ≤ 0.1 ppm, Linde AG, Germany). Whereas the inlet capillary (PEEK) and the drift tube of the PTR-MS was heated to 80°C. The respirometry chamber was in equilibrium with ambient temperature. The repeated flushing of the chamber and continuous diluting of the headspace in presence of a mouse kept the temperature within the chamber roughly at ambient temperature (20 - 25°C) comparable to home cage conditions. A membrane filter (PTFE, pore size 2.0 μm , PALL Corporation, Ann Arbor, U.S.A.) cleans the air stream into the PTR-MS from particular contaminations. After flushing with synthetic air blank samples were drawn from the empty box to control for leakage and contaminations (5 min, flow 60 mL min^{-1} controlled by PTR). During a second flushing with synthetic air, the mouse was placed into the respirometry chamber and the accumulation of exhaled VOCs was monitored. The cycle of flushing (2 min, flow 3 L h^{-1}) and accumulation of VOCs (5 min, flow 60 mL min^{-1} controlled by PTR) was repeated three times. During the accumulation the continuous loss of sampled air was replenished from a Teflon bag reservoir (capacity of 10 L, Welch Fluorocarbon Inc., Dover, U.S.A.) connected to the chamber

8.2.3.5. Monitoring of and controlling for contaminations

This setup facilitated measuring unrestrained, non-anaesthetized mice without inducing obvious stress-derived alterations in volatile organics emissions of the animal [21]. However, as the concentration of VOCs was monitored in a headspace, other sources, e.g. not endogenous, of VOCs may confound the accumulation profile. If necessary, mice were cleaned before starting the

measurement to remove fur contaminations (e.g. remaining wood shavings, dust). In blank measurements, e.g. an occasional increase in the acetone ($m/z=59.05$) and propanol ($m/z=41.06$) concentration pointed to leakage and high concentration in the laboratory air. The accumulation profile of humidity measured at ($H_2^{18}O$) $_2$ H $^+$, ($m/z=39.05$)), a distinct breath-driven profile due to the exhaled humidity, was affected by changes in locomotor behavior as mice showed a typical behavioral pattern from activity to settling down. The induced variation in the source strength was about a factor of two to mice at rest and to active mice [20]. Furthermore small droplets of urine in the respirometry chamber not only distorted the accumulation profile of humidity but polluted the head space. Feces and urine contaminations in the VOC profile were identified by a sudden increase in the concentration of methanethiol ($m/z=49.02$), trimethylamine ($m/z=60.07$) and pk127B ($m/z=127.02$). In case of crossing a threshold concentration of about 5 ppbV, feces were removed and in the case of an urine contamination the respirometry chamber was replaced, the mouse was gently cleaned using soft tissue paper (Kimtech Science, Kimberly-Clark), and the measurement was restarted.

8.2.3.6. Data analysis

The peak integral values exported from the TOF-Viewer were plotted as function of time (track) for all 164 masses to visualize the blank profile and the subsequent accumulation profiles (figure 1). The start and endpoint of the accumulation profiles were determined with the help of a subset of high-count mass signals (the water cluster, acetone and ethanol). This was done by looking for the local maxima and minima of the smoothed selected signals in order to identify increasing signals during accumulation. The profiles were fitted (nonlinear least-square fit) with the solution function of a simple compartment model [20]. In this approach, we described the concentration of a VOC C_i at time through a constant source strength S_i which is counterbalanced by the steady flow rate F out of the respirometry chamber (figure 1). The solution function included two free parameters, C_o (basal concentration after the ventilation) and S_i (source strength given in ppb*mL*min $^{-1}$). The fit include two noise models, which were the standard Gaussian uniform noise model (used as a first, rough baseline) and a uniformly linearly scaled Poisson noise model. This second model is based on the Poisson noise approach for electronic counting measurements, linearly scaled to roughly compensate the effect of count rate normalization to the primary ion count rate. Closeness of mean and variance estimates for these two models was later used as an indicator of goodness of fit.

If necessary the tracks were split into consecutive, non-overlapping subsets according to the statistical properties of the differential signals of the selected masses within the tracks. This was done in order to

identify trend-breaking events during the accumulation period, like a sudden change of breathing frequency in mice or the evaporation of water and VOCs from urine. Such events were recursively identified whenever the maximum p-value of a Kolmogorov-Smirnov test on the statistical distributions of the smoothed differential signals before and after each time a trend-breaking event was greater than a threshold p value (0.05). This parameter controlled the sensitivity to non-smooth accumulation profiles.

Profile visualization, profile selection and fitting were scripted in Matlab (VersionR2012a, The Mathworks Inc., MA, USA). Detected profiles were proof read for correct fitting and confounded profiles were removed ($S_i(\text{pk127B}) > 300 \text{ ppb} \cdot \text{mL} \cdot \text{min}^{-1}$; $S_i(\text{pk49}) > 100 \text{ ppb} \cdot \text{mL} \cdot \text{min}^{-1}$). The source strength, given in ppb mL min^{-1} was calculated for both blank and VOCs accumulation profiles. VOC source strengths measured during blanks were subtracted from VOC source strengths during mouse accumulation profiles.

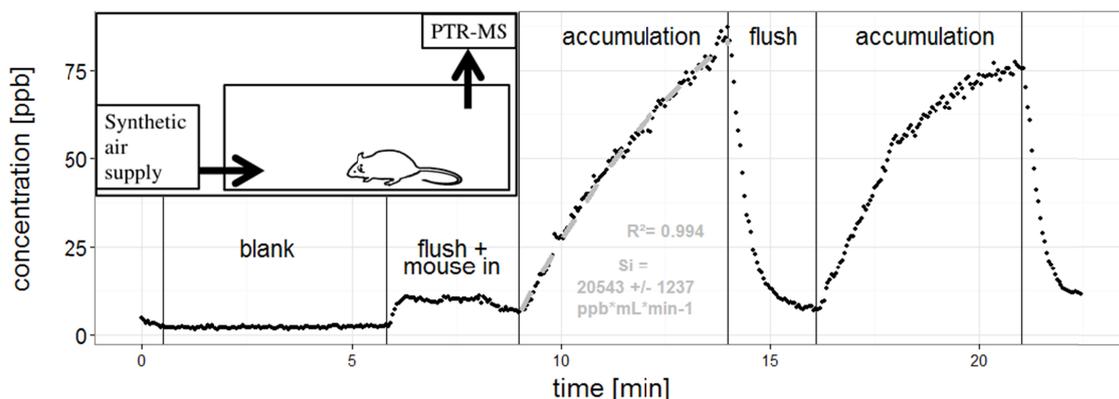


Figure 1. Real-time monitored methanol ($m/z=33.05$) signal in a respirometry chamber: blank profile of the empty chamber, mouse transfer to the chamber and two accumulation profiles separated by flushing with synthetic air. The first accumulation profile was fitted with the solution function of the compartment model to determine the source strength (shown in dashed grey). On the top left: a schematic view of the respirometry chamber (details in [20]).

8.2.3.7. Statistics

Due to the high number of variables per sample, Random Forest was used for feature identification as was previously published for breath data [22]. The original approach for replicated measures is to use averaged data per case for Random Forest based classifications only. However, this is discussed to be

accompanied with loss of information and poorer classification [23]. To account for this, RF++ was used which implements a subject-based bootstrapping in the original Random Forest algorithm (5000 trees, 13 splitting variables per tree; RF++, <http://sourceforge.net/projects/rfpp/>). For the factor diet-matrix (chow vs. semi-purified), blank-corrected source strengths data from baseline measurement were compared to week 1 data. The model was evaluated using the internal out-of-bag error. For the top 3 independent factors identified by the Random Forest, we calculated a linear mixed effects model accounting for longitudinal data (repeated measures ANOVA) to confirm the effects on the chosen volatile compounds. We started with a full model consisting of the factors diet matrix (chow vs. semi-purified), fat content (low vs high), fat quality (soy/tallow vs. lard) and the interaction of fat quality and quantity; mouse ID was used as random effect (nlme package, method: maximum likelihood, R [24]). Akaike information criterion (AIC) was then used to stepwise reduce the model (stepAIC, R). Homogeneity of variances was checked for plotting residuals versus fitted values. Normal distribution was monitored using Q-Q plotting; data for peaks 33B, 75 and 95A was log-transformed to achieve normality. As this is a study with rather exploratory character, lme model p values were not adjusted for multiple testing [25]. However, in case of a diet quality and quality interaction, we used pairwise t-testing to identify differences between single groups (pairwise.ttest, R) and adjusted the post-hoc testing using Bonferroni correction.

8.2.4. Results

8.2.4.1. Body mass gain in response to feeding the semi-purified diets

Within the first week after the diet switch mice on low calorie diets gained 1.6 g (LF) respectively 1.7 g (LL) in body mass whereas the high calorie diets induced higher body mass gain with 3.3 g (HF) respectively 2.2 g (HL). After four weeks the absolute difference in body mass was even more pronounced (LF: 3.2 g, HF: 7.5 g, LL: 3.8 g, HL: 6.4 g). About 40% of this mass gain could be attributed to an increase in body fat reserves as detected by qNMR (LF: 1.2 g, HF: 3.1 g, LL: 1.6 g, HL: 2.6 g).

8.2.4.2. Identification of discriminating volatile organic compounds for diet matrix

For the change of diet matrix from chow to semi-purified diet, an out-of-bag error of 0.00 was calculated in the RF++ and the confusion matrix showed showed no false classifications ($n_{\text{(chow)}}=27$, $n_{\text{(semi-purified)}}=26$), indicating that the groups were strongly classifiable. The importance of volatiles relevant for classification is shown as the mean decrease in margin for the diet matrix switch (figure 2). Here, three independent VOCs and related isotopes as well as water clusters were identified.

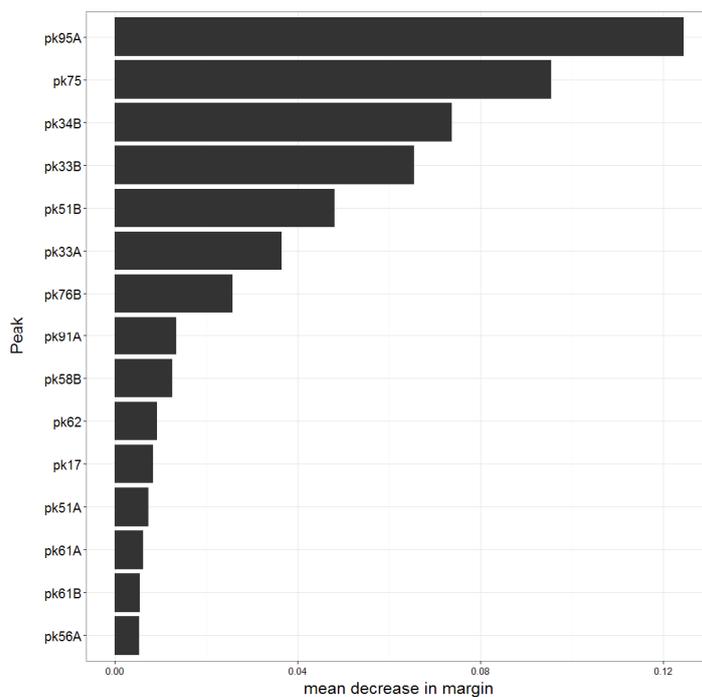


Figure 2. Variable importance (as mean decrease in margin) calculated by RF++ to identify effects of the diet matrix switch on VOCs comparing weeks zero and one.

8.2.4.3. Reduced source strengths of volatiles after change from laboratory chow to semi-purified diets

Pk33B - Methanol

The methanol signal is found in the spectrum at pk33B ($m/z=33.05$, figure 3(a)). In comparison to baseline, methanol levels drop significantly in response to a change to semi-purified diets ($p_{\text{matrix}} < 0.0001$, table 2). Interestingly, increased methanol levels in the LF group compared to other semi-purified diets were detected ($p_{\text{qual:quant}} < 0.0001$, table 2; pairwise ttest: figure 3(a)). Furthermore, the intense methanol signal on baseline chow feeding dominates as a C-13 isotope at pk34B ($m/z=34.06$) as well as a water cluster at pk51B ($m/z=51.03$) and is therefore found in the RF++ variable importance (figure 2(b), boxplot data not shown).

Pk75 - Methyl acetate/ Propionate

Pk75 ($m/z=75.02$, figure 3(b)), which we assigned to methyl acetate or propionate, shows a massive reduction in source strength from chow to semi-purified diets ($p_{\text{matrix}} < 0.0001$). Additionally, the linear mixed effects model also shows a slight decrease on low fat mice, in the soy oil/tallow-based diets and over time ($p_{\text{quant}}=0.0189$, $p_{\text{fatqual}}=0.0397$, $p_{\text{time}} < 0.0001$). Furthermore, pk76B ($m/z=76.01$) identified in the RF++ approach corresponds to C-13 isotope of pk75 (figure 2(b), boxplot data not shown).

Pk95 - Dimethyl disulphone

In rodent breath three candidates with a nominal mass of 95 (DMSO_2 , phenol, C_7H_{10}) are present and can be detected with a high-resolution PTR-ToF-MS [26]. Pk95A ($m/z=95.97$, figure 3(c)) is assigned to dimethyl disulphone (DMSO_2 , $m/z=94.998479$). Pk95A showed a decrease from baseline to semi-purified diet breath measurements ($p_{\text{matrix}} < 0.0001$). In addition, a diet quality to quantity interaction was detected, however none of the four groups differed from each other at the single time points ($p_{\text{qual:quant}}=0.0182$, table 2).

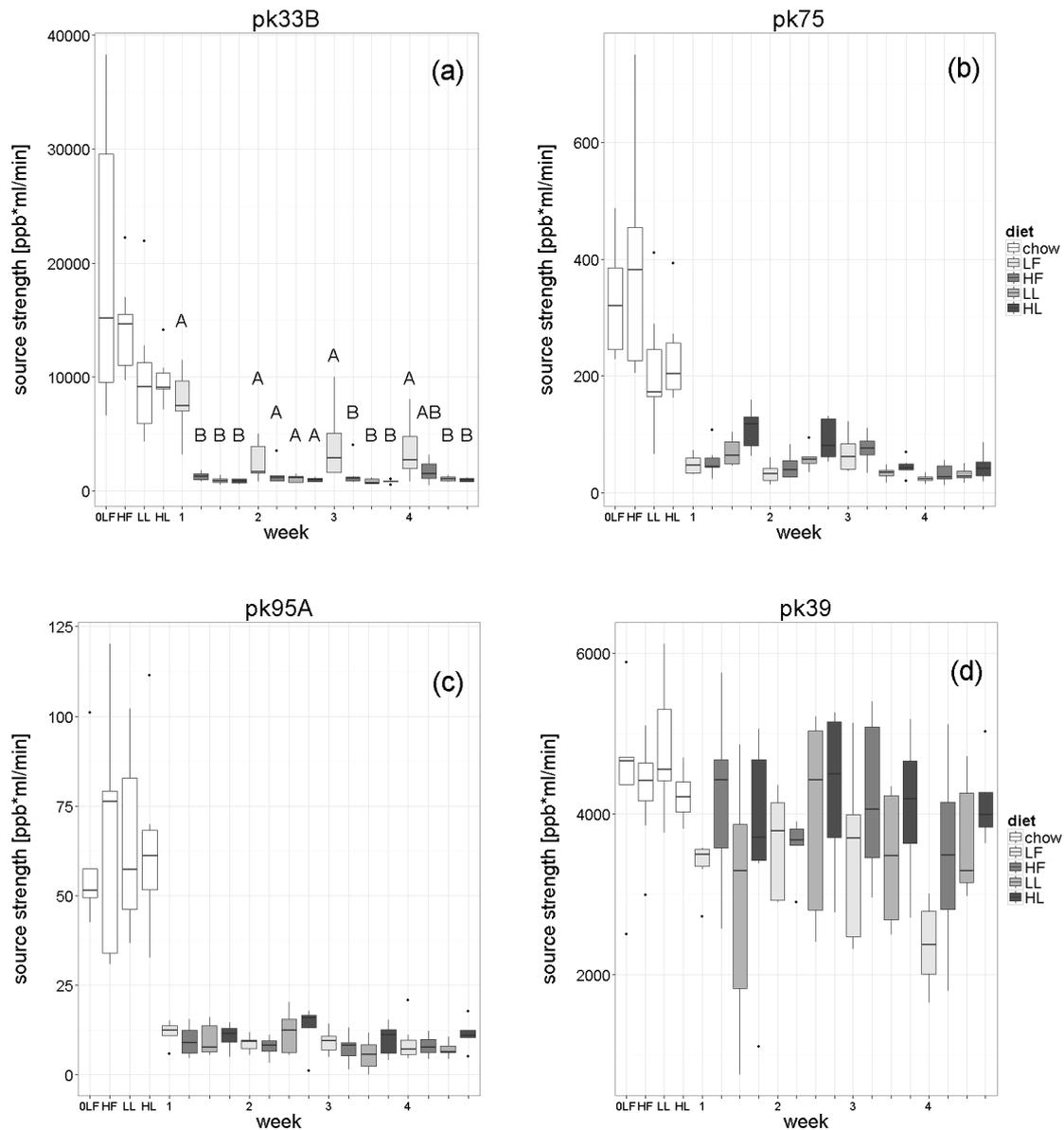


Figure 3. Effects of a switch from chow (week 0) to purified diets (weeks 1-4) on emitted breath VOCs (a-c) are shown (5-10 mice per group and date). If an interaction in diet quantity and quality was detected, different capital letters were used to show significant differences between diet groups tested with pairwise student's t-test at a time point (significance level: $p < 0.05$). For pk95A, no differences at a time point were found in pairwise t-testing. Water isotope cluster pk39 (d) is shown as a measure for humidity to estimate activity and breath frequency.

Table 4. A nested linear mixed effects model was applied on repeated measures of unique VOC signals identified in the RF++-approach. A stepwise reduction to identify the best model according to AIC was performed, not included parameters are marked (n.i.). Significance level: $p < 0.05$.

Name	Mouse weight	Fat quantity	Fat quality	Diet matrix	time	Quantity : quality interaction
Pk33B - Methanol	0.0876	0.4672	0.5167	<0.0001	0.0664	<0.0001
Pk75 – Methyl acetate/ Propionate	n.i.	0.0189	0.0397	<0.0001	<0.0001	n.i.
Pk95A – DMSO ₂	n.i.	0.0400	0.1333	<0.0001	0.0888	0.0182
Pk39 – H ₂ O*H ₃ O	n.i	0.0017	n.i	<0.0001	n.i	n.i

8.2.5. Discussion

We analyzed breath volatiles in C57BL/6J mice that experienced the shift from a standard laboratory chow to four different defined regularly used experimental diets which model high fat containing western style human diets. Chow diets contain considerable amounts of homogenized grain material that constitute the major matrix of the diet and depending on the season and harvesting time as well as the location of origin those diets can vary [27,28]. In contrast, semi-purified experimental rodent diets contain defined pure ingredients that are mixed according to the requirements of individual researchers and study aims. In our study, we particularly analyzed the immediate effects within one week after the diet shift. Those VOCs that showed the strongest contribution for the separation of VOC signatures between the experimental groups by Random Forrest were followed up over the consecutive four weeks.

Interestingly, the matrix switch from chow to semi-purified diets resulted in a massive drop in a set of volatiles. Methanol is one of the most evident reduced VOCs and its presence in (human) breath has been known for decades [29]. A significant contribution to methanol levels in breath may be related to the consumption of fruit-derived pectin [30,31]. Soy is a major ingredient of laboratory chow containing pectic polysaccharides that may be the main origin of methanol in the breath of chow-fed mice [32]. Interestingly, the source strength of methanol was found to be increased in mice fed the

low-fat soy-oil based diet (LF). In the literature a reduction of methanol in breath is reported for obese humans and rats whereas an increase was described in human individuals suffering from non-alcoholic fatty liver disease [8,26,33]. In obese individuals it was discussed whether the decline is due to decreased intake of soluble fibres or altered microbial cultivation and endogenous enzymatic activities. Small chain alcohols like methanol can also be produced by intestinal microbiota from water-soluble and non-resorbed carbohydrates[34]. The absorption of luminal glucose is increased in diabetic individuals and might be already present in an obese state, therefore less methanol might be produced[35]. In this mouse study, methanol source strengths did not depend on body mass, but on diet matrix and feeding the LF diet. Hence, these findings indicate that the production of methanol from dietary substrates by gut microbiota may be dominant over shifts due to metabolic alterations on high fat diet or obesity.

Another volatile with blunted source strengths on purified diets is pk75, which can be assigned to either methyl acetate or propionate. Both substances are short chain fatty acids (SCFA) or their derivatives. SCFA are also known to be predominantly produced by intestinal microbiota from dietary fibre, although there is some basal endogenous production of acetate [36,37]. The biochemical endogenous origin of methyl acetate is unknown so far. In vivo emission of pk75 could be shown in human primary bronchial cells [38]. As methanol source strengths similar pattern as methyl acetate, it can be hypothesized that part of the methyl acetate signal may be explained by an esterification of methanol and acetate.

Methyl acetate and in particular isoprene are known to be influenced by exercise or exercise-induced changes in cardiac output or blood pressure in humans [39]. This might hint to increased locomotor activity which could not be systematically monitored. However, the mice explored the respiratory box during the first measurement but habituated to the conditions later on and decreased activity levels. This change in behavior may provide an alternative explanation for the temporal pattern observed in this VOC. However, humidity (monitored in H₂O cluster isotope m/z39, Figure) and isoprene (data not shown) as additional markers for physical activity do not present an equivalent reduction (~19% reduction in humidity vs. ~82% reduction in pk75) after chow feeding, suggesting this drop being mainly an effect of the diet matrix switch and not decrease of activity levels due to habituation.

The third signal identified was found at pk95A, which we assigned to dimethyl sulphone (alias: methylsulfonylmethane). In the basal measurement on a chow diet, high levels of pk95A were present. Dimethyl sulphone (DMSO₂) breath levels are reported to be slightly reduced in obese rats and, in combination with dimethyl sulphide, might be used to detect non-alcoholic fatty liver disease (NAFLD) [26]. However, mice on normal laboratory chow with low fat and calorie content are

unlikely to be affected by NAFLD. Hence, the pk95A signal might originate rather from plant raw material e.g. grain included in the chow diet [40]. As this potential biomarker for NAFLD differs between chow and purified diets, it has to be emphasized that beyond macronutrient composition and disease state the diet matrix has a notable impact on VOC profiles.

The emphasis of the method for volatile measurement in unrestrained and non-stressed mice excluded alternative more invasive techniques as breath masks or anesthesia. Therefore, contributions to volatile source strengths from non-breath sources as urine or feces, from the skin, dust from home cage wood shavings stuck in the fur, might confound the monitoring. For the major contaminations urine and feces, marker substances were identified and only undisturbed profiles were further considered. Changes in locomotor behavior, e.g. from active exploration to settling down in the chamber, modulate to the derived source strength of exhaled VOCs [20].

The analysis of the spectra was performed using the software TOF-Viewer (Version 3.0.1) and the concentration was calculated using the semi-quantitative estimation formula [4]. To improve the quantification would include additional steps in spectra analysis as noise reduction, baseline removal and (automated) peak deconvolution and extraction as described by other groups [41]. The peak deconvolution with gaussian peaks used by the Tof-Viewer does not reflect the skewed shape of the peaks. This might over- or underestimate the peak area of the deconvoluted peaks.

8.2.6. Conclusion

To our knowledge, this is the first assessment of the effect of dietary matrix on breath gas analysis in non-anesthetized and non-restrained mice. Here we demonstrated that the switch from a plant-material based chow diet to several semi-purified diets independent of the main experimental factor fat content already results in considerably affected VOC signatures. Despite controlling for macronutrient composition, the respective diet matrix is highly relevant for breath gas analysis in rodents. This finding may be relevant because it is hypothesized to provide an explanation for the high inter-individual variability in rodent studies as well as in human clinical studies.

Acknowledgments:

We would like to thank Ann-Elisabeth Schwarz and Brigitte Herrmann as well as the GMC animal caretaker team for expert technical assistance. This work was funded by the German Federal Ministry of Education and Research to the German Center for Diabetes Research (DZD e.V.) and to the GMC (Infrafrontier grant 01KX1012).

References

- [1] Smith D, Turner C and Španěl P 2007 Volatile metabolites in the exhaled breath of healthy volunteers: their levels and distributions *J. Breath Res.* **1** 014004
- [2] Vautz W, Nolte J, Bufe A, Baumbach J I and Peters M 2010 Analyses of mouse breath with ion mobility spectrometry: a feasibility study *J. Appl. Physiol.* **108** 697–704
- [3] Hansel A, Jordan A, Warneke C, Holzinger R and Lindinger W 1998 Improved detection limit of the proton-transfer reaction mass spectrometer: on-line monitoring of volatile organic compounds at mixing ratios of a few pptv *Rapid Commun. Mass Spectrom.* **12** 871–5
- [4] Lindinger W, Hansel A and Jordan A 1998 On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research *Int. J. Mass Spectrom. Ion Process.* **173** 191–241
- [5] Miekisch W, Schubert J K and Noeldge-Schomburg G F . 2004 Diagnostic potential of breath analysis—focus on volatile organic compounds *Clin. Chim. Acta* **347** 25–39
- [6] Whittle C L, Fakharzadeh S, Eades J and Preti G 2007 Human Breath Odors and Their Use in Diagnosis *Ann. N. Y. Acad. Sci.* **1098** 252–66
- [7] Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, Ligor T, Filipiak W, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Haidenberger A, Buszewski B, Miekisch W, Schubert J and Amann A 2009 Noninvasive detection of lung cancer by analysis of exhaled breath *BMC Cancer* **9** 348
- [8] Morisco F, Aprea E, Lembo V, Fogliano V, Vitaglione P, Mazzone G, Cappellin L, Gasperi F, Masone S, De Palma G D, Marmo R, Caporaso N and Biasioli F 2013 Rapid “Breath-Print” of Liver Cirrhosis by Proton Transfer Reaction Time-of-Flight Mass Spectrometry. A Pilot Study *PLoS One* **8** e59658
- [9] Ibrahim B, Basanta M, Cadden P, Singh D, Douce D, Woodcock A and Fowler S J 2011 Non-invasive phenotyping using exhaled volatile organic compounds in asthma *Thorax* **66** 804–9
- [10] Van Berkel J J B N, Dallinga J W, Möller G M, Godschalk R W L, Moonen E J, Wouters E F M and Van Schooten F J 2010 A profile of volatile organic compounds in breath discriminates COPD patients from controls *Respir. Med.* **104** 557–63
- [11] Barbour A G, Hirsch C M, Ghalyanchi Langeroudi A, Meinardi S, Lewis E R G, Estabragh A S and Blake D R 2013 Elevated Carbon Monoxide in the Exhaled Breath of Mice during a Systemic Bacterial Infection *PLoS ONE* **8** e69802
- [12] Martinez-Lozano Sinues P, Kohler M and Zenobi R 2013 Human Breath Analysis May Support the Existence of Individual Metabolic Phenotypes *PLoS ONE* **8** e59909
- [13] Phillips M, Cataneo R N, Cheema T and Greenberg J 2004 Increased breath biomarkers of oxidative stress in diabetes mellitus *Clin. Chim. Acta* **344** 189–94
- [14] Španěl P, Dryahina K and Smith D 2013 A quantitative study of the influence of inhaled compounds on their concentrations in exhaled breath *J. Breath Res.* **7** 017106
- [15] Koeth R A, Wang Z, Levison B S, Buffa J A, Org E, Sheehy B T, Britt E B, Fu X, Wu Y, Li L, Smith J D, DiDonato J A, Chen J, Li H, Wu G D, Lewis J D, Warrier M, Brown J M, Krauss R M, Tang W H W, Bushman F D, Lusis A J and Hazen S L 2013 Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis *Nat. Med.*
- [16] Baranska A, Tigchelaar E, Smolinska A, Dallinga J W, Moonen E J C, Dekens J A M, Wijmenga C, Zhernakova A and van Schooten F J 2013 Profile of volatile organic compounds in exhaled breath changes as a result of gluten-free diet *J. Breath Res.* **7** 037104
- [17] Španěl P, Turner C, Wang T, Bloor R and Smith D 2006 Generation of volatile compounds on mouth exposure to urea and sucrose: implications for exhaled breath analysis *Physiol. Meas.* **27** N7

- [18] Fuchs H, Gailus-Durner V, Adler T, Pimentel J A A, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, Dalke C, Ehrhardt N, Fasnacht N, Ferwagner B, Frischmann U, Hans W, Hölter S M, Hölzlwimmer G, Horsch M, Javaheri A, Kallnik M, Kling E, Lengger C, Maier H, Mossbrugger I, Mörth C, Naton B, Nöth U, Pasche B, Prehn C, Przemeczek G, Puk O, Racz I, Rathkolb B, Rozman J, Schäble K, Schreiner R, Schrewe A, Sina C, Steinkamp R, Thiele F, Willershäuser M, Zeh R, Adamski J, Busch D H, Beckers J, Behrendt H, Daniel H, Esposito I, Favor J, Graw J, Heldmaier G, Höfler H, Ivandic B, Katus H, Klingenspor M, Klopstock T, Lengeling A, Mempel M, Müller W, Neschen S, Ollert M, Quintanilla-Martinez L, Rosenstiel P, Schmidt J, Schreiber S, Schughart K, Schulz H, Wolf E, Wurst W, Zimmer A and Hrabé de Angelis M 2009 The German Mouse Clinic: a platform for systemic phenotype analysis of mouse models *Curr. Pharm. Biotechnol.* **10** 236–43
- [19] Petersson F, Sulzer P, Mayhew C A, Watts P, Jordan A, Märk L and Märk T D 2009 Real-time trace detection and identification of chemical warfare agent simulants using recent advances in proton transfer reaction time-of-flight mass spectrometry *Rapid Commun. Mass Spectrom.* **23** 3875–80
- [20] Szymczak W, Rozman J, Höllriegel V, Kistler M, Keller S, Peters D, Kneipp M, Schulz H, Hoeschen C, Klingenspor M and Angelis M H de 2013 Online breath gas analysis in unrestrained mice by hs-PTR-MS *Mamm. Genome* 1–12
- [21] Lima P, Calil C and Marcondes F 2012 Influence of gender and stress on the volatile sulfur compounds and stress biomarkers production *Oral Dis.*
- [22] Hauschild A-C, Baumbach J I and Baumbach J 2012 Integrated statistical learning of metabolic ion mobility spectrometry profiles for pulmonary disease identification *Genet. Mol. Res. GMR* **11**
- [23] Karpievitch Y V, Hill E G, Leclerc A P, Dabney A R and Almeida J S 2009 An Introspective Comparison of Random Forest-Based Classifiers for the Analysis of Cluster-Correlated Data by Way of RF++ *PLoS One* **4**
- [24] R Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing (Vienna, Austria)
- [25] Bender R and Lange S 2001 Adjusting for multiple testing—when and how? *J. Clin. Epidemiol.* **54** 343–9
- [26] Aprea E, Morisco F, Biasioli F, Vitaglione P, Cappellin L, Soukoulis C, Lembo V, Gasperi F, D’Argenio G, Fogliano V and Caporaso N 2012 Analysis of breath by proton transfer reaction time of flight mass spectrometry in rats with steatohepatitis induced by high-fat diet *J. Mass Spectrom.* **47** 1098–103
- [27] Jensen M N and Ritskes-Hoitinga M 2007 How isoflavone levels in common rodent diets can interfere with the value of animal models and with experimental results *Lab. Anim.* **41** 1–18
- [28] Moraal M, Leenaars P P a. M, Arnts H, Smeets K, Savenije B S, Curfs J H a. J and Ritskes-Hoitinga M 2012 The influence of food restriction versus ad libitum feeding of chow and purified diets on variation in body weight, growth and physiology of female Wistar rats *Lab. Anim.* **46** 101–7
- [29] Eriksen S P and Kulkarni A B 1963 Methanol in Normal Human Breath *Science* **141** 639–40
- [30] Siragusa R J, Cerda J J, Baig M M, Burgin C W and Robbins F L 1988 Methanol production from the degradation of pectin by human colonic bacteria. *Am. J. Clin. Nutr.* **47** 848–51
- [31] Lindinger W, Taucher J, Jordan A, Hansel A and Vogel W 1997 Endogenous Production of Methanol after the Consumption of Fruit *Alcohol. Clin. Exp. Res.* **21** 939–43
- [32] Choct M, Dersjant-Li Y, McLeish J and Peisker M 2010 Soy oligosaccharides and soluble non-starch polysaccharides: a review of digestion, nutritive and anti-nutritive effects in pigs and poultry *Asian-Australas. J. Anim. Sci.* **23** 1386–98
- [33] Turner C, Spanel P and Smith D 2006 A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS *Physiol. Meas.* **27** 637–48

-
- [34] Hryniuk A and Ross B M 2010 A preliminary investigation of exhaled breath from patients with celiac disease using selected ion flow tube mass spectrometry *J. Gastrointest. Liver Dis. JGLD* **19** 15–20
- [35] Young R L, Chia B, Isaacs N J, Ma J, Khoo J, Wu T, Horowitz M and Rayner C K 2013 Disordered control of intestinal sweet taste receptor expression and glucose absorption in type 2 diabetes *Diabetes* **62** 3532–41
- [36] Macfarlane S and Macfarlane G T 2003 Regulation of short-chain fatty acid production *Proc. Nutr. Soc.* **62** 67–72
- [37] Buckley B M and Williamson D H 1977 Origins of blood acetate in the rat *Biochem. J.* **166** 539–45
- [38] Filipiak W, Sponring A, Filipiak A, Ager C, Schubert J, Miekisch W, Amann A and Troppmair J 2010 TD-GC-MS Analysis of Volatile Metabolites of Human Lung Cancer and Normal Cells In vitro *Cancer Epidemiol. Biomarkers Prev.* **19** 182–95
- [39] King J, Mochalski P, Kupferthaler A, Unterkofler K, Koc H, Filipiak W, Teschl S, Hinterhuber H and Amann A 2010 Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study *Physiol. Meas.* **31** 1169–84
- [40] Pearson T W, Dawson H J and Lackey H B 1981 Natural occurring levels of dimethyl sulfoxide in selected fruits, vegetables, grains, and beverages *J. Agric. Food Chem.* **29** 1089–91
- [41] Cappellin L, Biasioli F, Granitto P M, Schuhfried E, Soukoulis C, Costa F, Märk T D and Gasperi F 2011 On data analysis in PTR-TOF-MS: From raw spectra to data mining *Sensors Actuators B Chem.* **155** 183–90

8.3. Publication II

Diet-induced and mono-genetic obesity alter volatile organic compound signature in mice

Martin Kistler^{1,2,5}, Andreea Muntean³, Wilfried Szymczak³, Nadine Rink⁴, Helmut Fuchs^{1,2}, Valerie Gailus-Durner^{1,2}, Wolfgang Wurst^{6,7}, Christoph Hoeschen³, Martin Klingenspor⁴, Martin Hrabě de Angelis^{1,2,5,8} and Jan Rozman^{1,2,5,*}

¹Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany

²German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany

³Research Unit Medical Radiation Physics and Diagnostics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany

⁴ZIEL Department of Molecular Nutritional Medicine, Else Kröner-Fresenius Center, Technische Universität München, 85350 Freising, Germany

⁵German Center for Diabetes Research (DZD)

⁶Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany

⁷Developmental Genetics, Centre of Life and Food Sciences Weihenstephan, Technische Universität München, Freising, Germany

⁸Chair for Experimental Genetics, Life and Food Science Center Weihenstephan, Technische Universität München, D-85354 Freising-Weihenstephan, Germany

*Corresponding author. phone: +49 89 3187 3807; E-mail: jan.rozman@helmholtz-muenchen.de (J. Rozman), Postal: Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Munich, Germany

Abbreviations: VOCs, volatile organic compounds; HFD, high fat diet; LFD low fat diet; AUC, area under the curve; RF, random forest; ROC, receiver operating characteristic; FDR, false discovery rate; PTR-MS, proton transfer reaction mass spectrometry; TOF, time of flight; MC4R, melanocortin 4 receptor; MC4R-ki, MC4R W16X knock in; qNMR, quantitative nuclear magnetic resonance peak; gt, genotype; adlib, ad libitum; MTMT, (methylthio)methanethiol; DMS, dimethyl sulfide; OGTT oral glucose tolerance test.

Abstract

The prevalence of obesity is still rising in many countries, resulting in an increased risk of associated metabolic diseases. In this study we aimed to describe the volatile organic compound (VOC) patterns symptomatic for obesity. We analyzed high fat diet (HFD) induced obese and mono-genetic obese mice (global knock-in mutation in melanocortin-4 receptor MC4R-ki). The source strengths of 208 VOCs were analyzed in ad libitum fed mice and after overnight food restriction. Volatiles relevant for a random forest-based separation of obese mice were detected (26 in MC4R-ki, 22 in HFD mice). Eight volatiles were found to be important in both obesity models. Interestingly, by creating a partial correlation network of the volatile metabolites, the chemical and metabolic origins of several volatiles were identified. HFD-induced obese mice showed an elevation in the ketone body acetone and acrolein, a marker of lipid peroxidation, and several unidentified volatiles. In MC4R-ki mice, several yet-unidentified VOCs were found to be altered. Remarkably, the pheromone (methylthio)methanethiol was found to be reduced, linking metabolic dysfunction and reproduction. The signature of volatile metabolites can be instrumental in identifying and monitoring metabolic disease states, as shown in the screening of the two obese mouse models in this study. Our findings show the potential of breath gas analysis to non-invasively assess metabolic alterations for personalized diagnosis.

8.3.1. Introduction

Obesity has progressed to a worldwide epidemic linked to a number of co-morbidities, such as diabetes, cardiovascular disease, dyslipidemia and certain types of cancers (Guh et al 2009). Easily accessible biomarkers are central to the assessment of the individual risks of patients suffering from such pathologies and developing personalized medicine approaches for prevention and treatment. The dysregulation of metabolic pathways and the associated changes in body fluid metabolite concentrations are increasingly studied and also used for risk prediction (Mahendran et al 2013, Elliott et al 2015, Wahl et al 2015). A variety of normal and disease-associated metabolic reactions produce small volatile organic compounds (VOCs), which can be detected in body fluids but also non-invasively in exhaled breath. Over the past decade, advances in the methodology have made it possible to determine VOCs online in a concentration range of ppm to ppt and led to studies linking VOC signatures to various pathologies (Boots et al 2012). Regarding diseases associated with energy metabolism, several human studies were conducted in an attempt to monitor glucose levels (Lee et al 2009, Minh et al 2011), identify gestational, type 1, or type 2 diabetes (Halbritter et al 2012, Novak et al 2007, Greiter et al 2010), and characterize non-alcoholic fatty liver disease and liver cirrhosis (Morisco et al 2013, Alkhoury et al 2014). Prerequisites for broader clinical application are (1) clear identification of the molecules that are exhaled as VOCs and (2) a better understanding of the considerable inter- and intra-individual variation in VOCs found in breath even in healthy humans (Phillips et al 1999, Basanta et al 2012, Martinez-Lozano Sinues et al 2014). Environmental “wash-in”, diet and associated microbial changes as well as circadian rhythm might increase this variation. Animal models, and especially rodent models, for human diseases are a tremendously valuable tool used to deepen the understanding of molecular mechanisms and decipher the various sources of volatiles in a controlled environment (Rosenthal and Brown 2007). In this study, we were particularly interested in differences in the VOC signatures of normal weight mice and mice with manifested obesity induced either by feeding a high fat diet (HFD) or induced by targeted loss of function mutation of the melanocortin 4 receptor (Bolze et al 2011). VOC signatures could reflect the degree of obesity but could also be due to changes in diet composition, which is a confounding interaction that cannot be avoided in diet induced obesity studies (Baranska et al 2013, Kistler et al 2014). In contrast to diet-induced modifications of VOC signatures, to our knowledge no investigation on the effects of genetically-induced obesity on the volatilome, defined as the total amount of all VOCs emitted, has yet been conducted. Therefore, the aim of this study was to characterize alterations in exhaled VOCs both in a diet-induced and a mono-genetic obese mouse model and to evaluate whether a symptomatic pattern of VOCs related to obesity can be determined. In addition, individual changes in the volatilome of two specific obese models with distinct metabolic deregulations are of interest. We

employed statistical analysis methods that identified typical correlations between VOC emission rates, in the following called source strengths, that could be used to unravel the biochemical origin of the respective molecules.

8.3.2. Material and methods

8.3.2.1. Mice, animal housing and challenge experiments

Mice were housed in type IIL polycarbonate cages in individually ventilated cages (Tecniplast, Italy). A 12:12 h light/dark cycle at a temperature of 24 ± 1 °C and air humidity of 50–60% were maintained. Animals were housed in groups of 2-5 animals per cage in specific pathogen-free conditions at the German Mouse Clinic (GMC) (Fuchs et al 2009). Wood shavings were used for bedding (Altromin GmbH, Germany). For the generation of the diet-induced obesity model, 20 male C57BL/6J mice from in-house breeding were fed a pelleted laboratory chow from weaning onwards with ad libitum access to food and drinking water (no. 1314, Altromin, Lage, Germany). From the age of 12 weeks until the start of the VOC measurement (24 ± 2 weeks), their diet was changed to pelleted purified low fat and high fat diets (low fat: E 15000-04; high fat: E 15741-34; both: Ssniff, Soest, Germany). Assignment to diet groups was performed randomly using existing cage stocking to avoid single housing while ensuring balanced group numbers. A mono-genetic hyperphagic obesity model having a melanocortin-4-receptor nonsense allele W16X was used (Mc4r-ki mouse, as previously published (Bolze et al, 2011)). Fifteen homozygous MC4R-ki BL6/J mice as well as 15 controls were transferred to the GMC from the provider's laboratory at the age of 5 weeks and analyzed at the age of 24 ± 2 weeks. Mice had ad libitum access to drinking water and a pelleted laboratory chow from weaning onwards (<5 weeks: 'RM-Z autoklavierbar', Ssniff; >5 weeks: no. 1314, Altromin, Lage, Germany). All experiments were performed following animal welfare regulations with permission from the district government of Upper Bavaria (Regierung von Oberbayern).

For the analysis of VOCs from ad libitum fed mice, gas measurements took place between 1 pm and 6 pm. During this time, food consumption is low compared to night-time. Therefore, this period was chosen to reduce the contribution of food-derived volatiles to measured VOC patterns. Mice were measured in a random order and alternating between control and obese mice to remove potential systemic bias. For the fasted VOC measurements, mice were food deprived overnight beginning around 5-6 pm and were measured in the same order as in the ad libitum state between 8 to 12 am the following day. Mice were weighed before every VOC measurement to the nearest 0.1 g and body composition was monitored by non-invasive quantitative nuclear magnetic resonance scans in an ad

libitum fed state (Bruker Minispec LF50 body composition analyser, Ettlingen, Germany). The comparisons of body, lean and fat mass between groups were performed by using a linear regression model.

8.3.2.2. Proton-transfer reaction time-of-flight mass spectrometry and protocol for real-time breath gas analysis in unrestrained mice

A high-sensitivity proton transfer reaction mass spectrometer (PTR-MS, e.g. benzene 100 cps/pppV; PTR-MS, Ionicon Analytic GmbH, Innsbruck, Austria) with a resolution of $\Delta m/m \leq 2000$ was used. The principle of PTR-MS using H_3O^+ ions to softly ionize and detect VOCs was developed in the late 1990s (Lindinger et al 1998, Petersson et al 2009). A drift tube temperature of 80 °C, a drift tube voltage of 600 V and a drift pressure 2.3 mbar were applied. A mass range from m/z 0 to 349.5 was recorded (repetition rate of 77 kHz); the sum spectra with an integration time of 3 s were recorded (TOF-DAQ, Tofwerk AG, Switzerland). For the integration of peaks from the TOF-spectra, the PTR-MS Viewer software was used (Version 3.2.6, Ionicon analytic GmbH, Innsbruck, Austria). An internal calibration with the known peaks $H_3^{18}O^+$ (m/z 21.0221), NO^+ (m/z 29.9971) and protonated acetone (m/z 59.0491, $C_3H_6O.H^+$) was performed. 306 peaks were selected manually from the spectra. The deconvolution of overlapping peaks was performed by fitting a Gaussian distribution to the peaks in the PTR-MS Viewer. VOC concentrations were calculated using a constant k -rate of $2 * 10^{-9}$ [$cm^3 * s^{-1}$] in the semi-quantitative estimation formula (Lindinger et al, 1998). The system sensitivity was controlled regularly using a gas calibration unit (GCU, Ionicon Analytic GmbH, Innsbruck, Austria) with a mixture of substances (VOC gas standard, Ionicon Analytic GmbH, Innsbruck, Austria) regularly. From a set of compounds, a linear calibration curve obtained from multiple concentrations was used to calculate the individual transmission factors.

A setup and protocol for real-time measurement of breath gas in unrestrained mice using respiratory chambers was applied as described previously (Szymczak et al 2014; Kistler et al 2014). Deviating from the published set up, mice were acclimated to a training respiratory chamber for 7 min to reduce stress levels in the mice and consequently contamination by urine and feces.

8.3.2.3. Data analysis and statistics

Calculation of source strength and data pre-processing. A compartment model was used to describe the emission of a certain peak from recorded saturation curves as a source strength in $ppb * ml/min$ (non-linear regression, described in (Szymczak et al 2014)) and data pre-processing was performed as

described previously if not stated otherwise (Kistler et al 2014). Measurement starts and ends were defined manually using an in-house web-application based on R and shiny package (R Core Team 2015, Chang et al 2015). The information of group membership was not recorded in raw data files but added later on to ensure fully blinded analysis of saturation curves. As a further contamination control step, data was filtered for high concentrations or sudden increases of urinary and feces markers (pk127B, pk60 and pk49 > 1 ppb). For every peak, outliers (defined as greater than five standard deviations from mean) were removed. Peak data and single measurement data with more than 10% missing values were excluded. Peaks with source strengths not different or lower compared to corresponding blank source strengths were excluded as well (linear regression modeling with $p < 0.1$ to ensure enclosure of low signal candidate VOCs). An exception was made for known oxygen isotopes, as the negative source strength (=consumption) is expected. These filter steps resulted in final group sizes of 15 for fasted MC4R-wt mice, 14 for fasted MC4R-ki mice, 11 for ad libitum fed MC4R-wt mice and nine for every other experimental group. As a complete data-matrix is required to calculate random forest (RF) and Gaussian graphical models, missing data was imputed using chained equations (mice R package (van Buuren and Groothuis-Oudshoorn, 2011)), which accounted for 0.29% of data.

Feature selection and statistical testing of individual VOCs. By using TOF mass spectrometric detection of volatiles, the dataset consisted of a large number of peaks relative to animal numbers. We applied the area under the curve – RF (AUC-RF) algorithm as recently published (Calle *et al* 2011) to find a reduced set of candidate volatiles. In this algorithm, an initial RF is computed to obtain a ranking of predictors and an area-under-the receiver operating characteristic (ROC) curve. During the elimination process, less important variables are removed and AUCs of the resulting RFs are computed; an optimal set of predictors based on the AUC is finally reported. We used this algorithm for both HFD-fed and MC4R-ki datasets independently; setting strata to allow only one measurement per mouse and fasting status in every decision tree. A five-fold cross validation was applied 20 times to avoid over-fitting of each of the resulting RF model (using again a modified version of the algorithm allowing for stratification). Receiver operating characteristic (ROC) curves were created using R package pROC (Robin *et al*, 2011). The complete feature selection pipeline was repeated using randomly permuted class labels and plotted for comparison. Average and variability of areas under the ROC curves were visualized in boxplots for all 20 iterations of both pipelines. Peaks with a selection probability higher than 70% in the cross-validation AUC-RFs were used for further analysis.

For the analysis of genotype-induced and diet-induced effects on the VOC source strengths, two-sided mixed effects models were applied (Pinheiro *et al*, 2015). Both diet and genotype subsets of data were

log-transformed to approximate a normal distribution (tested visually by qq-plotting). The variance between groups was controlled using both boxplots of source strength as well as residuals and residual versus fitted data plots. For every peak of both subsets of data, effects of the corresponding intervention variable (diet or genotype, respectively), the fasting status as well as the interaction of both were tested using a mixed effects model accounting for repeated measures. If a significant interaction could be detected, individual group comparisons were performed by using the `multcomp` package (Hothorn *et al* 2008). As a larger number of tests leads to summation of type I – error, control of false discovery rate after Benjamini and Hochberg (1995) was applied and all p-values were adjusted according to a 10% false discovery rate.

Data visualization. Both sets of data were visualized in a clustered heatmap using the `Heatplus` (Ploner 2015) package from the Bioconductor project (Gentleman *et al* 2004). Mean ad libitum fed as well as mean fasted source strength data per mouse was used and shown individually. Boxplots were created using the R package `ggplot2` using all repeatedly measured source strength data (Wickham, 2009). In addition we applied a Gaussian graphical model to log-transformed source strengths of breath volatiles to visualize information about fragmentation, isotopic, water cluster and/or metabolic correlations. The complete dataset features more variables than number of mice, therefore we used a shrinkage approach to estimate a partial correlation matrix (Schäfer and Strimmer 2005). As the data has a longitudinal structure, we created a network accounting for that using dynamic (partial) correlation (Opgen-Rhein and Strimmer 2006). A network was extracted from the estimated partial correlation matrix using a local false discovery rate of 3% (GeneNet R package (Schaefer *et al*, 2015). A ‘dummy’ variable to correct for inter-experimental differences between HFD-fed and MC4R-ki mice was included in the network but not plotted. For every peak within a selected subset with significant fasting state, genotype or diet effect, the percentaged coefficients from mixed effects model are shown in the nodes as a pie-chart. The top 20% of connections is shown with bold lines, minor 20% with gray lines and negative partial correlations with dotted lines. Direct positive connections of significant nodes were highlighted and combined for overlapping subnetworks containing multiple significant nodes. Peaks included in AUC-RF model data-subsets but without significant connections were included in the graphical model for illustration purposes.

8.3.3. Results

8.3.3.1. Obesity state of HFD fed and mono-genetic mice

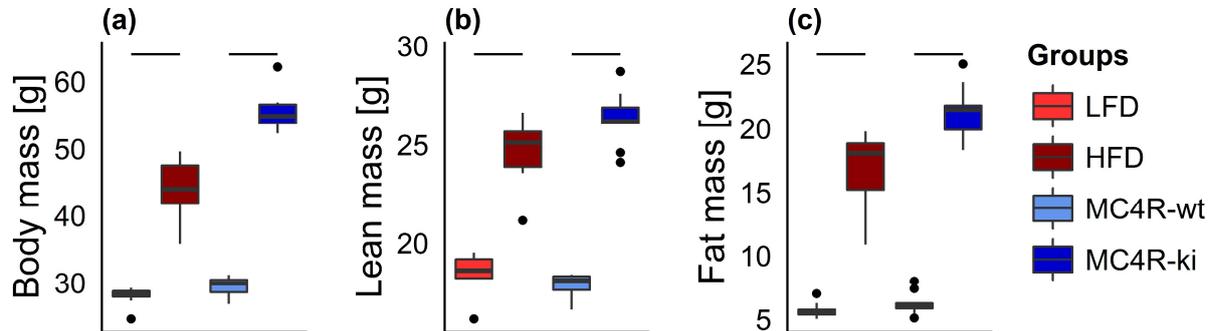


Figure 1: Body mass, lean mass and fat mass. Body mass (a), lean mass (b) and fat mass (c) for high-fat diet fed (HFD, dark red) and melanocortin-4-receptor *W16X* knock-in (*MC4R-ki*, dark blue) mice as well as corresponding controls (low-fat diet LFD, red; melanocortin-4-receptor wild type, *MC4R-wt*, blue) is shown in boxplots. Significant differences between controls and respective obesity mouse models are shown as black lines over individual boxes (linear regression model, $p < 0.05$). Group sizes: *MC4R-wt* ad lib ($n = 11$), other groups ($n = 9$).

Both HFD and the *MC4R-ki* mutation resulted in clear states of obesity as was evident from increased body mass as well as lean and fat mass. HFD-fed mice were heavier compared to littermate controls (44.21 ± 4.17 g versus 27.94 ± 1.47 g, $p = 4.01 \cdot 10^{-8}$, figure 1(a)). This gain in mass was partly due to an increase in lean mass (24.64 ± 1.7 g versus 18.47 ± 1.08 g, $p = 2.56 \cdot 10^{-7}$, figure 1(b)) as well as an increase in fat mass (16.83 ± 2.88 g versus 5.69 ± 0.65 g, $p = 2.16 \cdot 10^{-8}$, figure 1(c)). For *MC4R-ki* mice, also a considerable difference in body mass compared to littermate controls could be detected (50.99 ± 3.19 g versus 28.01 ± 1.28 g, $p = 2.00 \cdot 10^{-14}$, figure 1(a)). This difference was in part attributed to lean mass (26.28 ± 1.4 g versus 17.8 ± 0.64 g, $p = 5.67 \cdot 10^{-13}$, figure 1(b)), but largely due to elevated fat mass (21.3 ± 2.09 g versus 6.2 ± 0.83 g, $p = 1.77 \cdot 10^{-14}$, figure 1(c)). Overall, the impact of the *MC4R-ki* on body mass and body composition was more pronounced compared to the HFD model.

8.3.3.2. Selection of VOCs relevant for classification

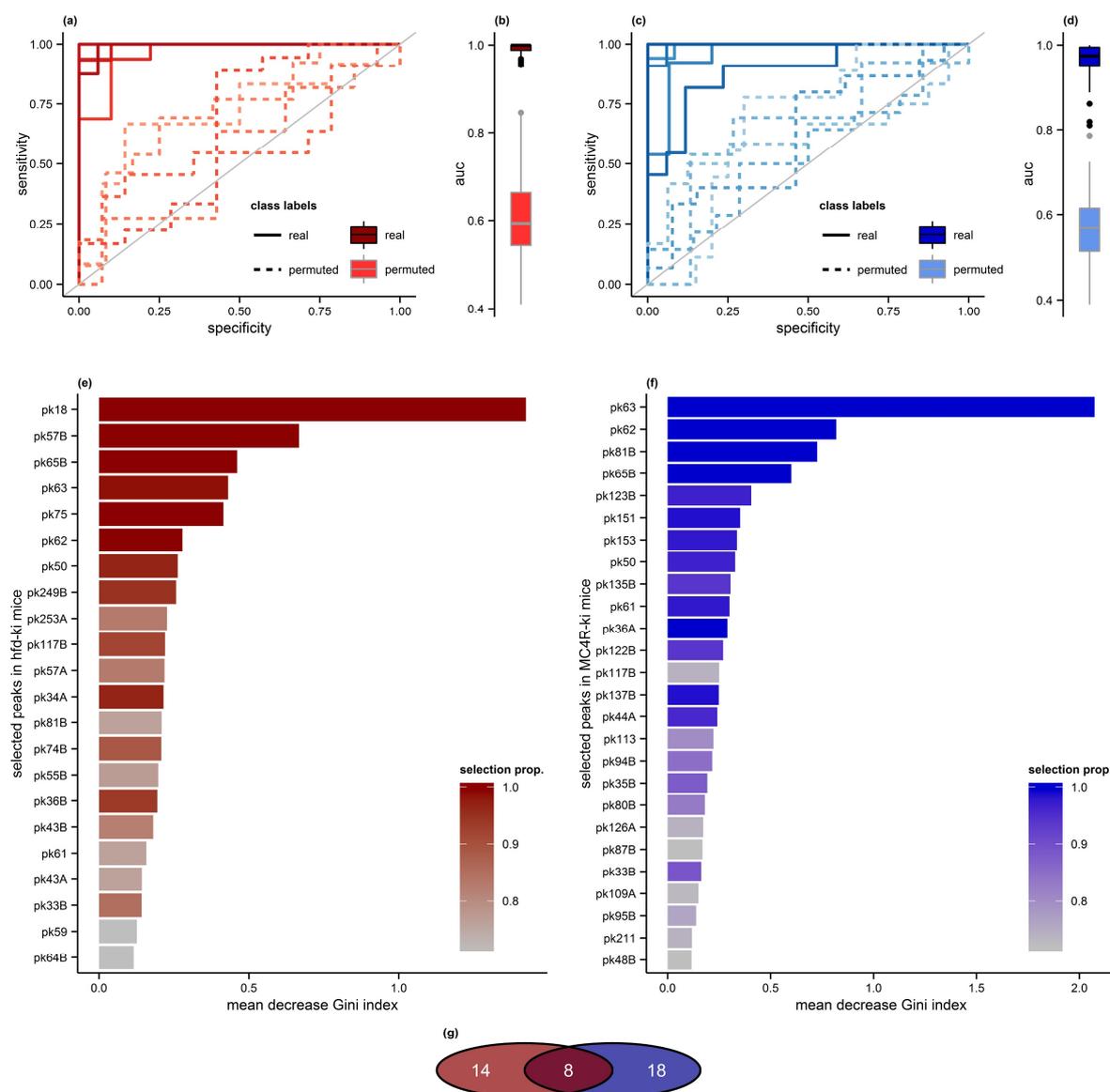


Figure 2: Machine learning strategy for selection of obesity relevant peaks. The performance of classification models was evaluated using ROC curves. For high-fat diet fed mice (HFD, (a)) and melanocortin-4-receptor *W16X* knock-in mice (MC4R-ki, (c)), ROC curves are shown for a single five-fold cross-validation of the recursive feature selection using AUC-RF algorithm. Results for real class labels are shown in dark solid lines. In comparison, light dashed lines represent ROC curves from using randomly permuted class labels in the classification procedure. The variation in area under ROC curves (AUCs) is shown in boxplots (HFD (b), MC4R-ki (d)). For this, 20 different 5-fold cross-validation sets were analyzed. Dark boxplot fills represent real class labels while light colors represent permuted labels. For further analysis, peaks with more than 70% selection probability in the repeated cross-validation procedure were selected. The variable importance of selected peaks is shown (HFD, (e); MC4R-ki, (f)). Color gradients indicate selection probability after 20

iterations of a five-fold cross validation procedure. Overlapping of selected peaks is shown as Venn diagram (g).

For both obesity models, feature selection was performed using the AUC-RF algorithm. In this algorithm, an optimal set of features is created by optimization of the area under the receiver operating characteristics (ROC) curve in a series of RF models. The ROC curves for a five-fold cross validation are shown for HFD (figure 2(a)) and MC4R (figure 2(c)). In addition, the ROC curves for the same procedure but with randomly permuted labels are plotted. For both datasets, a series of 20 different five-fold cross validations was computed both for real and permuted class labels to analyze robustness of the classification. Here, a drop of AUCs was observed in both permuted datasets (figures 2(b) and (d)). A cut-off of at least 70% selection probability was used to select 22 candidate peaks with the highest variable importance in HFD mice (figure 2(e)). For MC4R-ki mice, 26 candidate peaks fulfilled the cut-off criterion (figure 2(f)). Interestingly, within these peaks with the highest classification importance an overlap of eight peaks could be detected between the two mouse models (figure 2(g)). The eight peaks present in both groups were pk33B (methanol), pk50 (unassigned), pk61 (acetic acid), pk62 (MTMT), pk63 (CO₂, dimethyl sulphide (DMS)), pk65B (CO₂, DMS isotopes), pk81B (unassigned) and pk117B (unassigned).

8.3.3.3. Visualization of selected source strength data

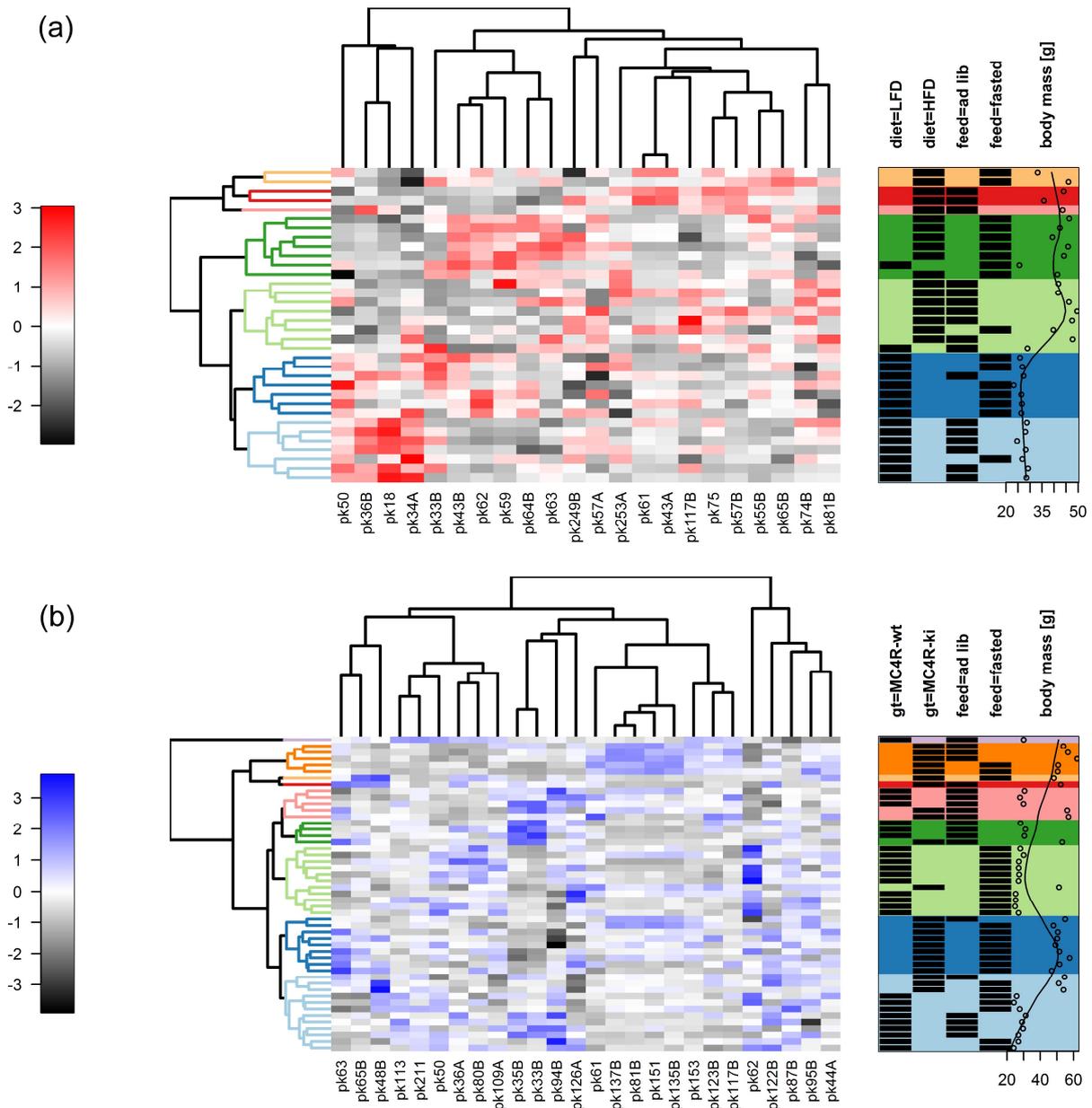


Figure 3: Heatmaps of selected VOCs. Heatmap of selected peaks in HFD fed (a) and MC4R-ki mice (b) are shown with hierarchical clustering of individual mice (mean data, rows, sub-clusters colored) and VOC peaks (columns, labels according to nominal mass). Data is scaled and centered. Color-coding legend shown on the left. Classification of individual mice is annotated on the right ((a): diet = LFD or HFD; (b): gt = MC4R-wt or MC4R-ki; both feed = ad libitum fed or fasted, body mass [g], subcluster-membership coloured). Group sizes: MC4R-ki fasted (n=15), MC4R-wt fasted (n=14), MC4R-wt ad libitum (n=11), other groups (n=9).

Heatmaps consisting of RF-selected peaks for both models were created to gain further insight into data structure by using unsupervised hierarchical clustering (figure 3). HFD-fed mice clustered in the top half of the heatmap with a remote subgroup within the dark blue sub cluster (figure 3(a)). Interestingly, despite the selection for obesity relevant peaks, a clustering according to fasting status was observed (fasted within light red, light green and dark blue, predominantly). Contrary to the findings in HFD-fed mice, the feeding status seemed to be the dominant clustering principle with MC4R wild type and knock-in mice showing fasted mice in the light green, dark blue and light blue sub clusters (figure 3(b)). Notably, fasted MC4R-ki mice clustered together mostly in the dark blue subcluster, whereas ad libitum fed knock-in mice showed a weaker clustering in the “warm colored” sub clusters.

8.3.3.4. Effects on VOC signature in diet-induced and monogenetic obesity

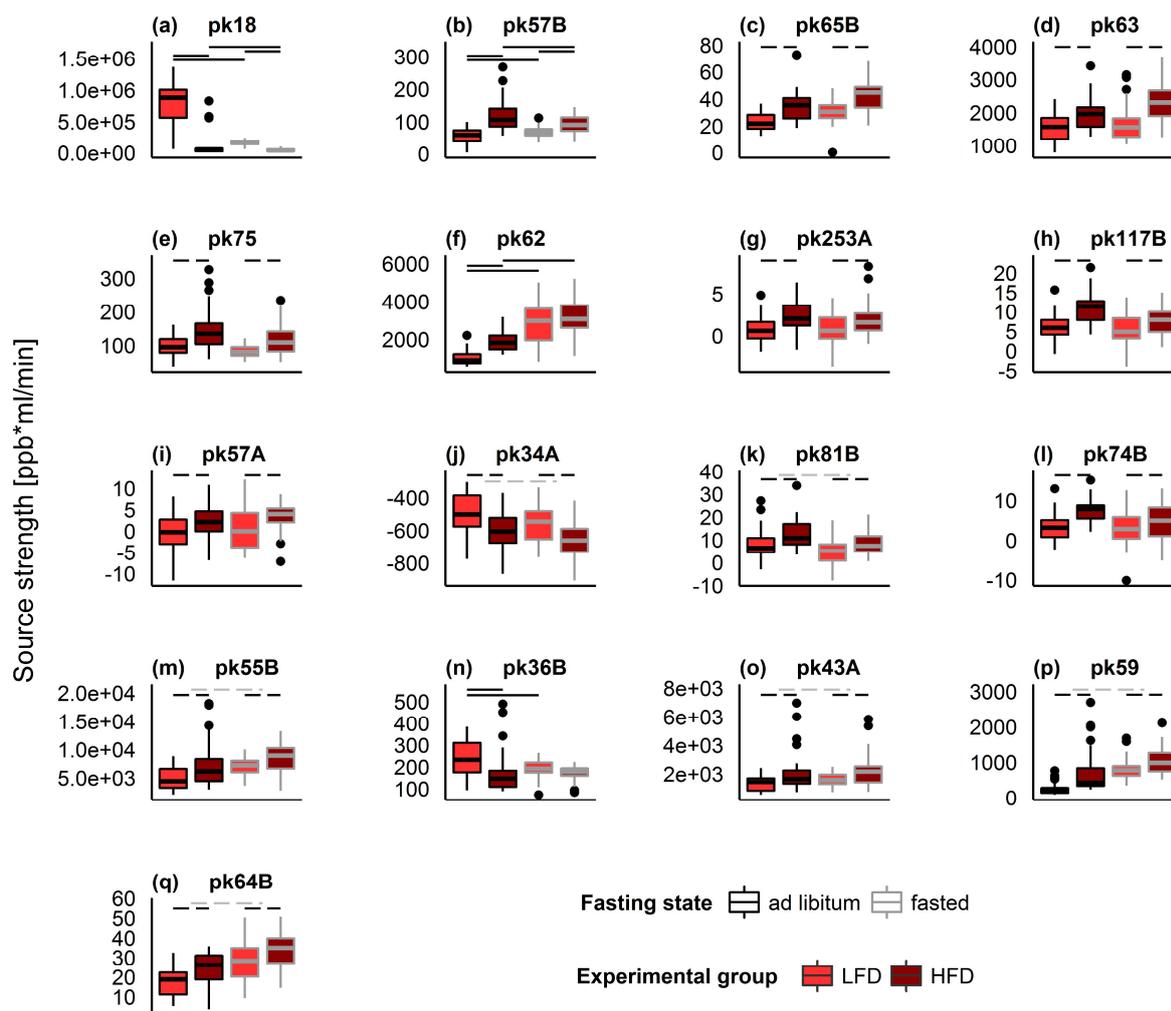


Figure 4: VOC source strengths with significant high fat diet effects. Source strengths for nominal mass-labeled peaks 18 (a), 57B (b), 65B (c), 63 (d), 75 (e), 62 (f), 253A (g), 117B (h), 57A (i), 34A (j), 81B (k), 74B (l), 55B (m), 36B (n), 43A (o), 59 (p) and 64B (q) are shown as boxplots (ordered after selection probability in cross-validated AUC-RF algorithm). Box fill corresponds to diet (red: low fat diet; dark red: high fat diet). Box border corresponds to fasting state (black: ad libitum fed; gray: fasted). Significant main effects in mixed effects model are shown as dotted lines (black: diet, gray: fasting state). In case of interaction, significant group differences are shown as black lines. P-values are adjusted for a false discovery rate of 10 %. Group sizes ($n = 9$).

In addition to obesity effects, we were interested in the effect of fasting on obesity candidate VOCs. Thus, we used statistical interference modeling to assess both diet and the food restriction effects. Linear mixed effects models were used to reflect the repetitive structure of the data (see supplementary Table 1 for details). For the HFD, a significant increase in source strength could be

found in 11 peaks (figure 4). Those peaks are 65B (unassigned, figure 4(c)), 63 ($\text{CO}_2 \cdot \text{H}_2\text{O}$ / DMS, figure 4(d)), 75 (methyl acetate, figure 4(e)), 253A (unassigned, figure 4(g)), 117B (unassigned, figure 4(h)), 57A (unassigned, figure 4(i)), 81B (unassigned, figure 4(k)), 74B (unassigned, figure 4(l)), 55B ($\text{H}_3\text{O}^+ \cdot \text{H}_2\text{O}$ / $\text{C}_4\text{H}_6\text{H}^+$, figure 4(m)), 43A ($\text{C}_2\text{H}_2\text{O} \cdot \text{H}^+$, figure 4(o)), 59 (acetone, figure 4(p)) and 64B ($^{13}\text{CO}_2 \cdot \text{H}_3\text{O}^+$ / $^{13}\text{CCH}_6\text{S} \cdot \text{H}^+$, figure 4(q)). Source strength in peak 34A ($^{16}\text{O}^{18}\text{O}$, figure 4(j)) was decreased in diet-induced obese mice.

In the HFD mouse model, several volatiles were affected by the fasting status of the mouse. Fasting induced higher emitted source strength in the seven peaks: 50 (unassigned, supplemental figure 1(a)), 55B ($\text{H}_3\text{O}^+ \cdot \text{H}_2\text{O}$ / $\text{C}_4\text{H}_6\text{H}^+$, figure 4(m)), 43B ($\text{C}_3\text{H}_6\text{H}^+$, supplemental figure 1(c)), 43A ($\text{C}_2\text{H}_2\text{O} \cdot \text{H}^+$, supplemental figure 1(o)), 33B (methanol, supplemental figure 1(e)), 59 (acetone, figure 4(p)) and 64B ($^{13}\text{CO}_2 \cdot \text{H}_3\text{O}^+$ / $^{13}\text{CCH}_6\text{S} \cdot \text{H}^+$, figure 4(q)). Three volatiles were reduced after overnight food restriction: 249B (unassigned, supplemental figure 1(b)), 34A ($^{17}\text{O}_2$, figure 4(j)) and 81B (unassigned, figure 4(k)).

In four peaks, an interaction of HFD feeding and food restriction was present. Ammonia (figure 4(a), pk18) was decreased in HFD mice (with a larger decrease in source strengths in the ad libitum fed state) and upon overnight fasting. Acrolein (Fig 4(b), pk57B (2-propenal, $\text{C}_3\text{H}_4\text{O} \cdot \text{H}^+$)) was elevated in obese mice in both states. Upon fasting source strength was increased in wild type mice but decreased in HFD fed animals. Peak 62 was assigned to a thiol-loss fragment of (methylthio)methanethiol (MTMT, $\text{CH}_3\text{SCH}_2 \cdot \text{H}^+$) as described in (Da Yu Lin et al, 2005) (figure 4(f)). MTMT was increased upon fasting in both groups. Furthermore, ad libitum fed HFD mice showed a higher source strength of MTMT compared to LFD fed mice, an effect which was no longer present when fasted mice were measured in the morning. pk36B was decreased in HFD ad libitum fed mice and fasted LFD fed mice (Fig 4(n), unassigned).

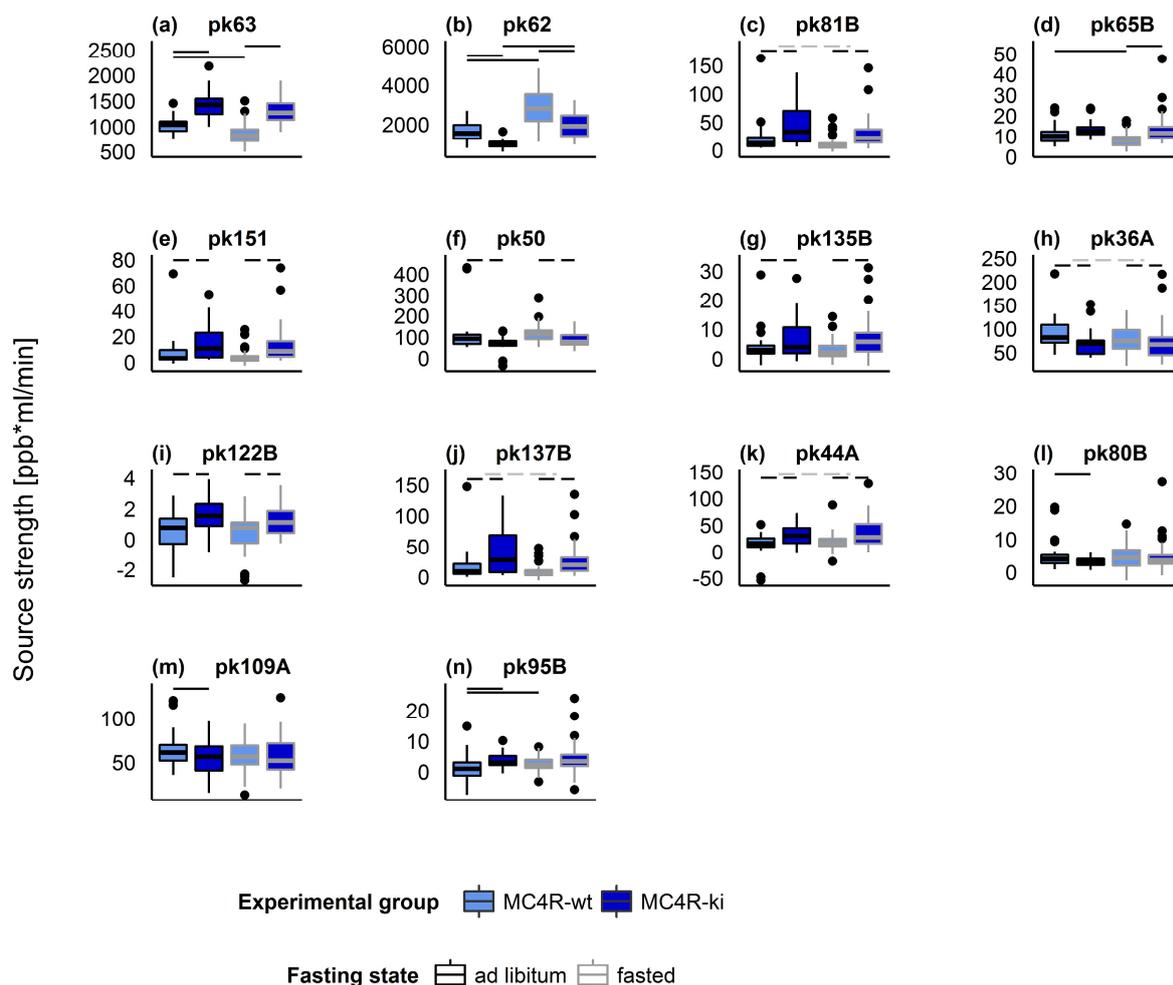


Figure 5: VOC source strengths with significant genotype effects. Source strengths for nominal mass-labeled peaks 63 (a), 62 (b), 81B (c), 65B (d), 151 (e), 50 (f), 135B (g), 36A (h), 122B (i), 137B (j), 44A (k), 80B (l), 109A (m) and 95B (n) shown as boxplots (ordered after selection probability in AUC-RF feature selection). Box fill corresponds to genotype (blue: melanocortin-4-receptor wild type; dark blue: melanocortin-4-receptor W16X knock-in). Box border corresponds to fasting state (black: ad libitum fed; gray: fasted). Significant main effects in mixed effects model are shown as dotted lines (black: genotype, gray: fasting state). In case of interaction, significant group differences are shown as black lines. P-values are adjusted for a false discovery rate of 10 %. Group sizes: MC4R-ki fasted ($n = 15$), MC4R-wt fasted ($n = 14$), MC4R-wt ad libitum ($n = 11$), MC4R-ki ad libitum ($n = 9$).

In MC4R-ki mice several VOC source strengths differed between genotypes as shown by multifactorial linear mixed effects modeling using genotype and fasting status (figure 5, detailed model results can be found in supplementary table 1). A significant increase was found in eight volatiles: peak 63 ($\text{CO}_2 \cdot \text{H}_2\text{O}$ / DMS, figure 5(a)), 81B (unassigned, figure 5(c)), 151 (unassigned,

figure 5(e)), 135B (unassigned, figure 5(g)), 122B (unassigned, figure 5(i)), 137B (unassigned, figure 5(j)), 44A ($C_2H_3O / ^{13}CCH_2O.H^+$, figure 5(k)) and 95B (unassigned, figure 5(n)). In five volatiles, a decrease in source strength was observed in MC4R-ki mice: peak 62 (MTMT, figure 5(b)), 50 (unassigned, figure 5(f)), 36A (unassigned, Fig 5(h)), 80B (unassigned, figure 5(l)) and 109A (unassigned, figure 5(m)).

In addition to genotype effects, overnight food restriction affected four peaks positively and eleven peaks negatively. An increase in fasted state was observed in peaks 62 (MTMT, figure 5(b)), 44A ($C_2H_3O / ^{13}CCH_2O.H^+$, figure 5(k)), 87B (unassigned, supplemental figure 2(i)) and 95B (unassigned, figure 5(n)). Decreased fasting source strengths were found for peaks 63 (CO_2*H_2O/ DMS , figure 5(a)), 81B (unassigned, figure 5(c)), 123B (unassigned, supplemental figure 2(a)), 61 (acetic acid, supplemental figure 2(c)), 36A (unassigned, figure 5(h)), 117B (unassigned, supplemental figure 2(d)), 137B (unassigned, figure 5(j)), 113 (unassigned, supplemental figure 2(e)), 94B (unassigned, supplemental figure 2(f)), 35B ($CH_3^{16}OH.H^+$, supplemental figure 2(g)) and 211 (unassigned, supplemental figure 2(k)).

An interaction of MC4R-ki genotype and food restriction was present in peak 65B (figure 5(d), unassigned). Here, source strength was increased in fasted but not in ad libitum fed MC4R-ki mice and reduced in wild type mice in response to fasting.

8.3.3.5. Gaussian graphical modeling as a tool to identify VOCs

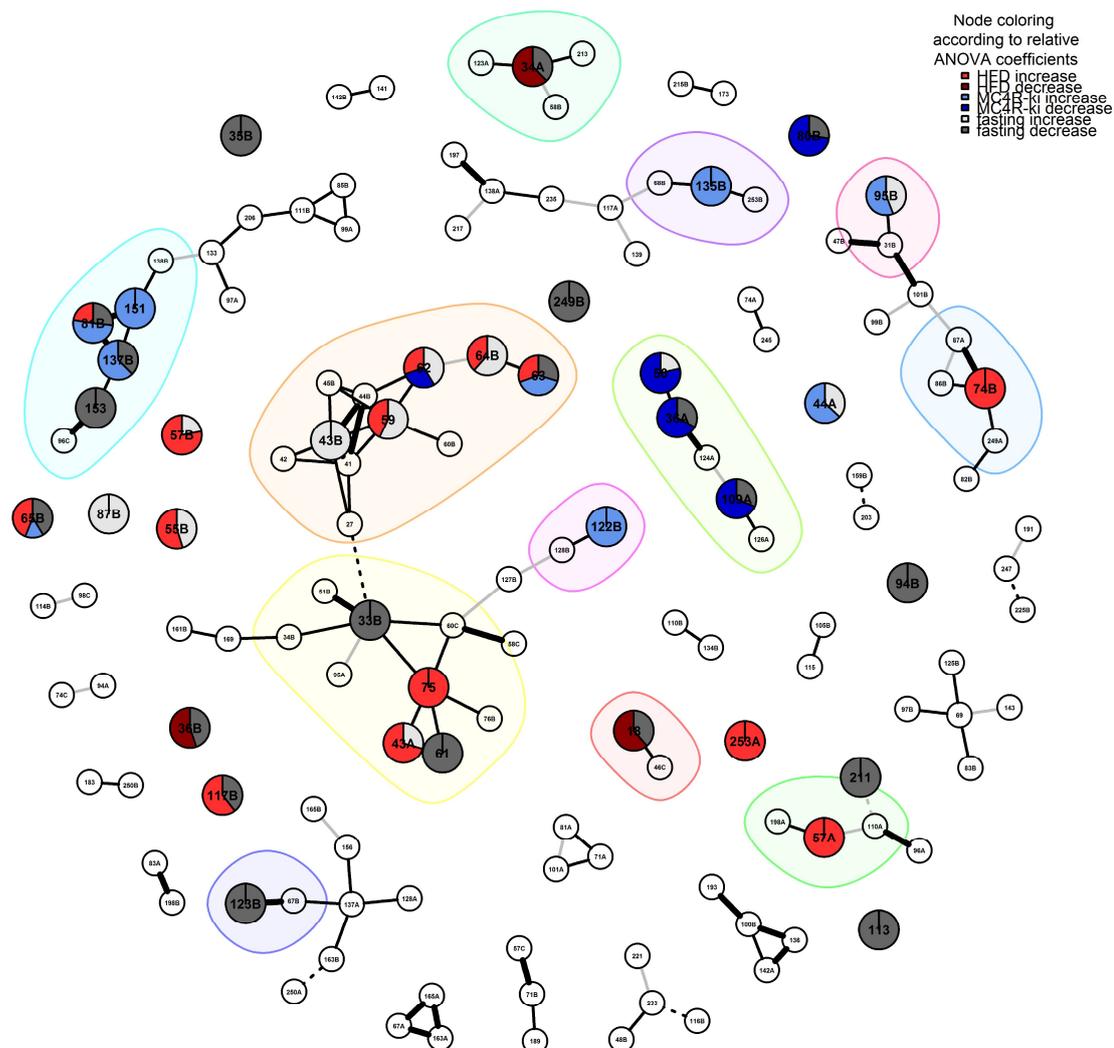


Figure 6: Gaussian graphical model for VOC identification. Gaussian graphical model with nodes corresponding to peaks (labeled with nominal mass and letter for multiple peaks at the same nominal mass) and edges corresponding to shrinkage estimated partial correlation. Highest and lowest 20 % of correlations are highlighted (bold black/ thin gray). Dotted edges indicate negative partial correlation. Peaks with significant mixed effects model main effects (as seen in figures 4 and 5) are shown as pie charts in nodes. Coloring of model coefficients is according to diet increase/ decrease (red/ dark red), genotype increase/ decrease (blue/ dark blue) and fasting increase/ decrease (gray/ dark gray). Mean fasting coefficients are shown if significant in both obesity models. Subnetworks of peaks with significant effects and directly connected nodes with positive partial correlation are plotted on colored background. Peaks without significant edges selected in AUC-RF were added for illustration purposes.

Gaussian graphical models were proposed recently to identify metabolites and model metabolic pathways from metabolomics data (Krumstiek *et al* 2012, 2011). As in a large p, smaller n dataset, a full partial correlation matrix cannot be directly applied, we estimated a partial correlation network using by a shrinkage approach for longitudinal data (Schäfer and Strimmer, 2005; Opgen-Rhein and Strimmer, 2006). In this graphical model, several obesity relevant subnetworks can be detected with additional information on VOC identity (figure 6). In addition to genotype and diet effects, fasting effects are visualized in gray (increase) and dark gray (decrease) pie slices. If both cohorts showed fasting coefficients, we calculated mean coefficients. Fasting coefficients from MC4R cohort contribute to peaks 33B, 35B, 36A, 44A, 61, 62, 63, 65B, 80B, 81B, 87B, 94B, 95B, 109A, 113, 117B, 123B, 137B, 153 and 211 while fasting coefficients from HFD models are contribute to 18, 33B, 34A, 36B, 43A, 43B, 50, 55B, 57B, 59, 62, 64B, 81B and 249B. Interpretations of peak-peak-connections are given in supplementary table 2.

8.3.4. Discussion

The success of personalized medicine approaches for metabolic diseases depends on inexpensive and minimally invasive but also sensitive and specific diagnostic tools. The analysis of VOCs in human breath has the potential to provide such an ‘easy-access’ view to a broad range of metabolic pathways. However, the origin and the link to physiological functions of many volatiles are still unknown, thus hindering the implementation of a breath gas screening in clinical settings. Here in this study, we screened the breath of obese mice for disease related alterations in VOC patterns under controlled and standardized circumstances. The application of so-called -omics technologies generates datasets with a high ratio of measured parameters to the specimen included in the study. This requires adequate statistical methods to account for, e.g. multiple testing. In breathomics, machine learning strategies can be applied for both classification and feature selection (as comprehensively reviewed (Smolinska *et al* 2014)). In our study, we used a RF-based recursive feature selection algorithm, AUC-RF, to derive VOCs altered in obesity. We used multiple iterations of five-fold cross validation to avoid model overfitting. Furthermore, a cut-off of 70% selection probability was set to reduce selection false positive features. It has to be noted that, although RF is rather robust to collinearity in typical breath data (Smolinska *et al* 2014), some highly correlated features might remain undetected by using a selection probability cut-off as they are almost equally likely to be selected. However, when applying this approach to data with randomly permuted class labels, no (HFD) or a single (MC4R-ki) peak is selected in the process. This, in combination with the massive drop in the area under the ROC curve, underlines the robustness of the procedure. Permutations as a tool for model validation are used in the field of breath research (Halbritter *et al* 2012, Hauschild *et al* 2013). Regarding the slightly elevated median and the observed variation of the area under the ROC curves, we support the idea to combine both cross-validations and label permutations to validate classification models. Interesting additions or variations to the applied approach might be the use of bootstrapping procedures or the Boruta algorithm, which conveniently implements benchmarking against random variables in a single all-relevant feature selection algorithm (Kursa and Rudnicki, 2010).

We found that a variety of volatiles were affected in obese mice or varied depending on the feeding status of the animals. For both visualizing and identifying altered patterns of VOCs, we estimated a Gaussian graphical model as a data-driven approach novel in PTR-MS based breath gas analysis. Known from metabolomics studies (Krumstiek *et al* 2011, 2012), graphical modeling here helped to identify contributions to single peaks; for example, isotopes of MTMT (pk62) and $\text{CO}_2\cdot\text{H}_3\text{O}^+$ /DMS (pk63) to peak 64B. In addition, visualizing fragmentation patterns like propanol fragments at peaks 41 and 43B is possible (Schwarz *et al* 2009). In addition to chemical properties, biochemical pathway

information could be included, such as the conversion of acetone to 2-propanol by ADH1 (Lewis *et al* 1984). The combination of this information can shed light on previously unknown peaks and the corresponding volatiles. In addition, the graphical model resembles the hierarchical clustering shown in the two heatmaps in several edges but emphasizes numerous further connections as it can be built on both complete data sets with correction for intra-experimental effects. In the following, the tentatively identified volatiles altered in obese mice are discussed individually with a special emphasis on the available obesity and metabolic disease relevant literature.

8.3.4.1. **A VOC signature altered in the volatilome both obesity models**

In the present study we found that in both diet-induced and genetically induced obesity source strengths of several emitted VOCs were altered. Interestingly, a set of eight peaks was changed in both obesity models, of which four were tentatively identified and four remain unknown.

Acetic acid (pk61, pk43A) A volatile identified in both groups by the feature selection algorithm as potentially relevant was acetic acid (pk61 and fragment pk43A in HFD mice). In this untargeted screening approach mixed effects modeling did not find significantly changed source strengths in acetate but in the fragment pk43A in HFD mice. Interestingly, despite the non-significant increase in source strength, the increased variance and outlying data in both obesity models might be used for stratification of obesity related pathologies such as disturbances in glucose homeostasis. The enzymes acetyl-CoA synthetase and acetyl-CoA hydrolase regulate free acetic acid levels, in addition exogenous sources such as gut fiber fermentation contribute to serum levels predominantly after food intake (Wolever *et al* 1997). Serum acetic acid levels were reported to be inversely correlated to insulin levels in mice and humans (Layden *et al* 2012, Sakakibara *et al* 2009). Acetate reduces glucose-induced insulin secretion via pancreatic free fatty acids receptors (FFAR) 2 and 3. This is likely mediated by pancreas secreted acetate produced from glucose as a negative feedback as well as from overall systemic acetic acid levels (Tang *et al* 2015). It is therefore coherent that acetic acid in breath could be used to model glucose levels during an oral glucose tolerance test and to detect individuals with gestational diabetes (Halbritter *et al* 2012). In addition, short chain fatty acids like acetate have been recognized to induce a PPAR γ -dependent switch from lipid synthesis to lipid utilization in white adipose tissue and liver (den Besten *et al* 2015). Thus, if the variance in acetic acid could be attributed to associated (patho)physiological states in obesity, it might be a relevant non-invasive marker.

Methanol (pk33B) Methanol was selected in both models (peaks 33B and MC4R-ki mice also ¹⁸O-isotope 35B). A massive increase of methanol was found in the MC4R experiment during the fed state. These mice were fed a so-called chow diet comparably high in pectin/fiber content thus affecting VOC signatures mediated by altered microbial digestion as previously described (Kistler et al 2014). We could not detect a general effect of obesity status on methanol. Thus, the effect of diet on methanol breath levels seems to exceed the endogenous variation. This is in accordance to other studies, identifying methanol to originate mainly from microbial digestion of consumed pectins with a smaller fraction from other dietary and endogenous sources as aspartame or S-adenosylmethionine (Axelrod & Daly 1965, Siragusa *et al* 1988, Lindinger *et al* 1997, Dorokhov *et al* 2012). Notably, methanol was used together with a set of VOCs to model blood glucose in type 1 diabetics (Minh *et al* 2011) and was found to be reduced in HFD fed rats (Aprea *et al* 2012), inversely correlated to BMI in humans (Turner *et al* 2006, Halbritter *et al* 2012) but increased in liver cirrhosis (Morisco *et al* 2013). Those findings may be related to differences in lifestyle or eating habits (e.g. reduced fruit (pectin) consumption) affecting gut microbiota in obese patients. In addition, a reduction of methanol detoxification capacity could be present in those states. However, to verify this in a mouse model, it has to be considered that detoxification of methanol in humans is primarily adh1-driven, while in rodents, peroxidative activity of catalase is relevant for degradation (Dorokhov *et al* 2015, Karinje and Ogata 1990).

*Carbon dioxide *H2O / Dimethyl sulphide (pk63, pk64B, pk65)* Another peak elevated in both obese models is peak 63 as well as isotopes at 64B (in HFD mice) and 65, which are probably a mixed signal from carbon dioxide-water cluster and DMS. Carbon dioxide, as a terminal mitochondrial oxidation product of most energy-containing molecules, is directly related to the amount of energy used in the organism. Indeed, the utilized obese models do have an increased overall amount of metabolic active tissue (figure 1). This increase consists not only of fat mass, which is considered to have a lower but not negligible metabolic activity per gram (Kaiyala *et al* 2010), but also of highly active lean mass which is elevated. Hence a higher emission of carbon dioxide in heavier mice is not surprising (Butler and Kozak 2010, Tschöp *et al* 2012).

DMS, the second candidate, was found to be increased in obese rats with steatohepatitis, obese children, liver cirrhotic patients and is a known constituent of the fetor hepaticus (Aprea *et al* 2012, Alkhouri *et al* 2015, Morisco *et al* 2013, Van den Velde *et al* 2008). DMS can be found in breath after

methionine ingestion and is altered in hepatitis and cirrhosis patients showing an increased half-life (Kaji *et al* 1979). Furthermore, rat skeletal muscle cells were observed to be releasing DMS, possibly produced by the transamination pathway out of methionine and cysteine (Mochalski *et al*, 2014). Thus, after further insight in its metabolism, DMS could be used as a non-invasive biomarker of altered systemic or hepatic metabolism of sulfur-containing amino acids.

(Methylthio)methanethiol (MTMT, pk62) The source strength of a fragment of (methylthio)methanethiol (MTMT, pk62 and part of pk64B signal in HFD fed mice) is ~100 times higher in male mice compared to females (as shown for MC4R-ki animals in supplementary figure 3). This long-distance pheromone was initially described to be detected in the main olfactory bulb, being involved in the attractiveness of male urine to female mice and can be reduced in urine by castrating male mice (Da Yu Lin *et al* 2005). Unexpectedly, we found the source strength of MTMT significantly increased in ad libitum fed diet-induced obese mice but reduced in MC4R-ki mice. The tissue and mechanism of endogenous MTMT synthesis are currently unknown. However, a link between the melanocortin 4 receptor and sexual reproduction has been shown (Van der Ploeg *et al* 2002) and reduced mating success is observed in models with reduction in melanocortin production (Faulkner *et al* 2015). In contrast to the MC4R-ki mice, upon HFD feeding an elevation in MTMT in ad libitum state but not in the fasted state is observed. Upon HFD feeding, an acute compensatory activation of MC4R signaling is known (Butler *et al* 2001) and could also modulate MTMT levels via an MC4R-dependent mechanism.

8.3.4.2. Significant HFD specific peaks

Ammonia (pk18) Ammonia is elevated in liver pathologies (Adeva *et al* 2012) and was found increased in obese children (Alkhoury *et al* 2015). In addition, in a rat study featuring diet-induced obesity measured with similar instrumentation, breath ammonia was increased in purified HFD versus low fat standard diet fed rats (Aprea *et al* 2012). Unexpectedly, we found reduced breath ammonia source strength of in HFD-fed mice in comparison to control mice. Generally, values from HFD-fed mice seem to be lower than both groups in the MC4R experiment as well. In the field of breath research, the reproducibility of breath ammonia measurements is in discussion (Blanco Vela and Bosques Padilla 2011), as aside from changed blood ammonia concentrations, breath ammonia altered by physical activity level (Solga *et al* 2014), mode of breathing, airway or mouth pH (Solga *et al* 2013) and mouth bacteria expressing urease (Chen *et al* 2014). As mice show a strong preference for

nasal breathing, some of the above-mentioned effects should not be present here. A higher dietary protein load in this particular HFD (24.1% in HFD versus 20.8% in LFD) can contribute to a metabolic acidosis. This in combination with the ketoacidosis in ad libitum HFD and fasting could lead to an increased urinary ammonia excretion to compensate acidosis and therefore reduced breath ammonia.

Acrolein (pk57B) Interestingly, another VOC increased in HFD-fed mice both in ad libitum as well as in fasted state but not selected as relevant in MC4R-ki mice is likely acrolein. In humans, acrolein exposure from exogenous sources as diet as well as inhalation of polluted air and smoking are known to be relevant. In addition, endogenous production from lipid peroxidation in oxidative stress, degradation of methionine/ threonine and spermine/ spermidine can contribute to the observed concentrations (Stevens and Maier 2008). Notably, in HFD mice a slight reduction in fasted state is observed, probably indicating that both directly diet-derived and endogenous produced acrolein contribute to the elevation compared to LFD fed littermates. Acrolein is contributing to metabolic pathologies via a wide range of mechanisms and target tissues, including protein adduction, induction of oxidative stress, mitochondrial dysfunction, inflammation and immune alterations, ER stress, structural and membrane effects and deregulated signal transduction as reviewed by Moghe *et al* (2015). Hence it can be an interesting breath resource for monitoring carbonyl stress and redox state.

Methyl acetate (pk75) An increase in pk75 in HFD mice was observed, which is possibly predominantly methyl acetate as suggested by the Gaussian graphical model. In addition to the individually discussed literature on acetate and methanol, methyl acetate in breath is only described to be increased acutely after exercise (King *et al* 2010). In obese patients with non-alcoholic fatty liver disease, an increase in various fecal volatile esters including methyl acetate could be observed and associated to a gut microbial shift (Raman *et al* 2013). Emitted methyl acetate therefore indicates such a shift, or alternatively can be created from acetate and methanol within the mouse metabolism.

¹⁸O¹⁶O oxygen (pk34A) An increased consumption of oxygen isotope ¹⁸O-¹⁶O is observed in HFD-fed mice. This can likely be explained by a higher amount of metabolic active tissue and therefore higher absolute oxygen demand in the heavier HFD mice (Tschöp *et al* 2012).

H₃O⁺.(H₂O)₂ water cluster and fragments of aldehydes (pk55B) Pk55B is increased in HFD mice and is likely to consist of both H₃O⁺.(H₂O)₂ water cluster and fragments of aldehydes e.g. butanal, hexanal, octanal or nonanal (Buhr *et al* 2002). Although it is unclear why humidity and water clustering should be increased in obese mice, aldehydes in breath (and breath condensate) can be increased in oxidative stress pathologies with associated lipid peroxidation (Amann *et al* 2014), which can be elevated in obese state.

Acetone and propanol (pk59, pk43B). In HFD-induced obesity, energy demands are to an extended portion satisfied by lipid oxidation and hepatic ketogenesis. One of the ketone bodies is acetone, which is thought to be produced by spontaneous decarboxylation of acetoacetate. Therefore, in both states of increased fatty acid oxidation, namely HFD and food restriction, acetone source strengths are elevated. In humans, fasting breath acetone levels were shown to be highly correlated to β -hydroxybutyrate and acetoacetate blood concentrations (Qiao *et al* 2014, Musa-Veloso *et al* 2006). Both were associated with increased fasting and 2 h plasma glucose levels and acetoacetate could be used to predict both an increased OGTT AUC and 5-year diabetes incidence (Mahendran *et al* 2013). Notably, in a subnetwork of fasting responsive volatiles (figure 6), acetone showed a significant partial correlation to both propanol fragments at nominal masses 41 and 43B (which showed a significant fasting but no diet effect). The conversion from acetone to iso-propanol is known and can be enhanced in a ketogenic setting (Lewis *et al* 1984, Petersen *et al* 2012). As conversion from propanol to acetone is also possible and breath propanol is highly correlated to environmental concentrations in a clinical setting (Ghimenti *et al* 2013), this can be one reason why in human breath analysis high variance in acetone levels is observed.

8.3.4.3. **MC4R-ki specific peaks**

Cluster of unknowns 151 (F), 153 (G), 137B (N) and 81B (C) In the gaussian graphical model, a subnetwork affected in MC4R-ki mice was observed, featuring peaks 151 (F), 153 (G), 137B (N) and 81B (C). A literature search on similar PTR-MS fragmentation patterns revealed that monoterpenes like α - and β -pinene, 3-carene, limonene and camphor produced fragment ions of masses 67, 81 and 95, as well as a protonated molecular ion of mass 137 or 153 (Tani *et al* 2003). Notably, in a human lipid infusion study to predict plasma TG and FFA levels from breath volatiles, β -limonene and β -pinene were relevant for the models (Minh *et al*, 2012). Also monoterpenes (137.137) and terpene-related peak (135.119) have been found as breath markers for liver cirrhosis in a human study

(Morisco et al, 2013). An altered diet composition, as one explanation the authors named, can be excluded here. So either increased food consumption in MC4R-ki mice or the suggested alteration hepatic terpene metabolism in this study can explain the elevated levels in obese mice. However, it has to be noted that the observed peaks do not match the theoretical masses exactly. Possibly, the fact that those peaks are far from the internal calibration masses typically applied in PTR-MS measurements using $\text{H}_3^{18}\text{O}^+$ (21.02), NO^+ (30.00) and protonated acetone (59.05) should be causing this mass shift. Especially in the used PTR-TOF-2000 instrument with a resolution of $\leq 2000 \text{ m}/\Delta\text{m}$, an added high molecular internal calibration gas can be useful to be included in future studies.

8.3.4.4. **Unassigned volatiles**

In addition, three further volatiles - namely peaks 50, 81B and 117B - were identified to be altered in both obese mouse models. Unidentified peaks were also found in HFD-fed mice, including peaks 249B, 253A, 57A, 36B and 74B. Even more volatiles are considered unknown in MC4R-ki mice, including peaks 44A, 123B, 135B, 36A, 122B, 137B, 113, 94B, 80B, 87B, 211 and 95B. Those candidates are worth further exploration using complementary methods like classical pre-concentration combined with gas chromatography–tandem mass spectrometry or a novel combination of a fast-GC device to the PTR-MS for additional chemical information (Romano *et al* 2014). However, altered experimental settings leading to increased VOC concentration can be necessary due to the lower sensitivity of those methods. Alternatively, nose mask sampling could be applied despite its obvious need for extensive acclimation of rodents to avoid stress induced effects on the measured volatiles (Aprea *et al* 2012). Interestingly, data-driven models such as the applied Gaussian graphical model can at least in part contribute to the identification of VOCs. In addition to the volatiles showing effects in mixed effects models, peaks 126A and 48B did not show effects and are considered false positives in the selection process.

8.3.5. **Conclusion**

In this study we characterized alterations in exhaled volatile organic compounds in both diet-induced and mono-genetic obese mouse models and aimed to evaluate whether a common pattern of VOCs altered in obesity can be determined. Alterations in the volatilome could be detected with a common obesity VOC signature. Notably, different adiposity models create distinct shifts in the volatilome as well, thus showing the potential of VOC analysis to monitor and distinguish different obesogenic

mechanisms. Identified VOCs originate from various metabolic pathways and biological processes including ketone body metabolism, lipid peroxidation and pheromones; allowing a broad overview over metabolic state in a fast and non-invasive way. In addition, we suggest Gaussian graphical models as a helpful tool in understanding and characterizing the volatilome. Thus, the analysis of the volatile metabolome has the potential to contribute to a personalized medicine by aiding in the stratification of patients with heterogeneous metabolic phenotypes and risk profiles.

Author contributions

MKi conceived and designed the experiments, researched data, reviewed and analyzed the data and wrote the manuscript. NR and AM took care of animal management, researched data (AM), reviewed and edited the manuscript. WS and JR conceived and designed the experiments, reviewed the data, wrote (JR), reviewed and edited the manuscript. WW contributed to mouse line generation. CH, M Kl, HF, VGD, WW and MHA contributed to discussion, reviewed and edited the manuscript.

Acknowledgments

The authors declare that they have no conflict of interest. We thank B Herrmann, AE Schwarz and all animal caretakers in the GMC for their technical contribution to mouse care and phenotyping. Furthermore, our thanks go to F Bolze and R Kühn, who contributed to generation and initial characterization of the MC4R-ki mouse line. This work was partly funded by the FP7 Marie Curie Initial Training Network PIMMS, Grant Agreement No. 287382, by German Federal Ministry of Education and Research (Infrafrontier grant 01KX1012) and by the German Center for Diabetes Research (DZD).

References

- Adeva MM, Souto G, Blanco N & Donapetry C (2012) Ammonium metabolism in humans. *Metabolism*. **61**: 1495–1511
- Alkhoury N, Cikach F, Eng K, Moses J, Patel N, Yan C, Hanouneh I, Grove D, Lopez R & Dweik R (2014) Analysis of breath volatile organic compounds as a noninvasive tool to diagnose nonalcoholic fatty liver disease in children: *Eur. J. Gastroenterol. Hepatol.* **26**: 82–87
- Alkhoury N, Eng K, Cikach F, Patel N, Yan C, Brindle A, Rome E, Hanouneh I, Grove D, Lopez R, Hazen SL & Dweik RA (2015) Breathprints of childhood obesity: changes in volatile organic compounds in obese children compared with lean controls. *Pediatr. Obes.* **10**: 23–29
- Amann A, Miekisch W, Schubert J, Buszewski B, Ligor T, Jezierski T, Pleil J & Risby T (2014) Analysis of Exhaled Breath for Disease Detection. *Annu. Rev. Anal. Chem.* **7**: 455–482
- Aprèa E, Morisco F, Biasioli F, Vitaglione P, Cappellin L, Soukoulis C, Lembo V, Gasperi F, D'Argenio G, Fogliano V & Caporaso N (2012) Analysis of breath by proton transfer reaction time of flight mass spectrometry in rats with steatohepatitis induced by high-fat diet. *J. Mass Spectrom.* **JMS 47**: 1098–1103
- Axelrod J & Daly J (1965) Pituitary gland: enzymic formation of methanol from S-adenosylmethionine. *Science* **150**: 892–893
- Baranska A, Tigchelaar E, Smolinska A, Dallinga JW, Moonen EJC, Dekens JAM, Wijmenga C, Zhernakova A & Schooten FJ van (2013) Profile of volatile organic compounds in exhaled breath changes as a result of gluten-free diet. *J. Breath Res.* **7**: 037104
- Basanta M, Ibrahim B, Douce D, Morris M, Woodcock A & Fowler SJ (2012) Methodology validation, intra-subject reproducibility and stability of exhaled volatile organic compounds. *J. Breath Res.* **6**: 026002
- Benjamini Y & Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**: 289–300
- den Besten G, Bleeker A, Gerding A, van Eunen K, Havinga R, van Dijk TH, Oosterveer MH, Jonker JW, Groen AK, Reijngoud D-J & Bakker BM (2015) Short-Chain Fatty Acids Protect Against High-Fat Diet-Induced Obesity via a PPAR γ -Dependent Switch From Lipogenesis to Fat Oxidation. *Diabetes* **64**: 2398–2408
- Blanco Vela CI & Bosques Padilla FJ (2011) Determination of ammonia concentrations in cirrhosis patients-still confusing after all these years? *Ann. Hepatol.* **10 Suppl 2**: S60–65
- Bolze F, Rink N, Brumm H, Kühn R, Mocek S, Schwarz A-E, Kless C, Biebermann H, Wurst W, Rozman J & Klingenspor M (2011) Characterization of the melanocortin-4-receptor nonsense mutation W16X in vitro and in vivo. *Pharmacogenomics J.* **13**: Available at: <http://www.readcube.com/articles/10.1038%2Ftpj.2011.43> [Accessed February 23, 2015]
- Boots AW, Berkel JJBN van, Dallinga JW, Smolinska A, Wouters EF & Schooten FJ van (2012) The versatile use of exhaled volatile organic compounds in human health and disease. *J. Breath Res.* **6**: 027108
- Buhr K, van Ruth S & Delahunty C (2002) Analysis of volatile flavour compounds by Proton Transfer Reaction-Mass Spectrometry: fragmentation patterns and discrimination between isobaric and isomeric compounds. *Int. J. Mass Spectrom.* **221**: 1–7
- Butler AA & Kozak LP (2010) A Recurring Problem With the Analysis of Energy Expenditure in Genetic Models Expressing Lean and Obese Phenotypes. *Diabetes* **59**: 323–329
- Butler AA, Marks DL, Fan W, Kuhn CM, Bartolome M & Cone RD (2001) Melanocortin-4 receptor is required for acute homeostatic responses to increased dietary fat. *Nat. Neurosci.* **4**: 605–611
- Buuren S van & Groothuis-Oudshoorn K (2011) mice: Multivariate Imputation by Chained Equations in R. *J. Stat. Softw.* **45**: 1–67

- Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J et al (2015) shiny: Web Application Framework for R Available at: <http://cran.r-project.org/web/packages/shiny/index.html> [Accessed February 11, 2015]
- Chen W, Metsälä M, Vaittinen O & Halonen L (2014) The origin of mouth-exhaled ammonia. *J. Breath Res.* **8**: 036003
- Da Yu Lin S-ZZ, Block E & Katz LC (2005) Encoding social signals in the mouse main olfactory bulb. *Nature* **434**: 470–477
- Dorokhov YL, Komarova TV, Petrunia IV, Kosorukov VS, Zinovkin RA, Shindyapina AV, Frolova OY & Gleba YY (2012) Methanol May Function as a Cross-Kingdom Signal. *PLoS ONE* **7**: Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3338578/> [Accessed July 19, 2013]
- Dorokhov YL, Shindyapina AV, Sheshukova EV & Komarova TV (2015) Metabolic Methanol: Molecular Pathways and Physiological Roles. *Physiol. Rev.* **95**: 603–644
- Elliott P, Posma JM, Chan Q, Garcia-Perez I, Wijeyesekera A, Bictash M, Ebbels TMD, Ueshima H, Zhao L, Horn L van, Daviglus M, Stamler J, Holmes E & Nicholson JK (2015) Urinary metabolic signatures of human adiposity. *Sci. Transl. Med.* **7**: 285ra62–285ra62
- Faulkner L, Dowling A, Stuart R, Nillni E & Hill J (2015) Reduced melanocortin production causes sexual dysfunction in male mice with POMC neuronal insulin and leptin insensitivity. *Endocrinology*: en.2014–1788
- Fuchs H, Gailus-Durner V, Adler T, Pimentel JAA, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, Dalke C, Ehrhardt N, Fasnacht N, Ferwagner B, Frischmann U, Hans W, Hölter SM, Hölzlwimmer G, Horsch M, Javaheri A, Kallnik M, Kling E, et al (2009) The German Mouse Clinic: a platform for systemic phenotype analysis of mouse models. *Curr. Pharm. Biotechnol.* **10**: 236–243
- Gentleman RC, Carey VJ, Bates DM & others (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **5**: R80
- Ghimenti S, Tabucchi S, Lomonaco T, Di Francesco F, Fuoco R, Onor M, Lenzi S & Trivella MG (2013) Monitoring breath during oral glucose tolerance tests. *J. Breath Res.* **7**: 017115
- Greiter MB, Keck L, Siegmund T, Hoeschen C, Oeh U & Paretzke HG (2010) Differences in exhaled gas profiles between patients with type 2 diabetes and healthy controls. *Diabetes Technol. Ther.* **12**: 455–463
- Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL & Anis AH (2009) The incidence of comorbidities related to obesity and overweight: A systematic review and meta-analysis. *BMC Public Health* **9**: 88
- Halbritter S, Fedrigo M, Höllriegel V, Szymczak W, Maier JM, Ziegler A-G & Hummel M (2012) Human Breath Gas Analysis in the Screening of Gestational Diabetes Mellitus. *Diabetes Technol. Ther.* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22775148> [Accessed July 12, 2012]
- Hauschild A-C, Kopczynski D, D’Addario M, Baumbach JI, Rahmann S & Baumbach J (2013) Peak detection method evaluation for ion mobility spectrometry by using machine learning approaches. *Metabolites* **3**: 277–293
- Hothorn T, Bretz F, Westfall P, Heiberger RM & Schuetzenmeister A (2014) multcomp: Simultaneous Inference in General Parametric Models Available at: <http://cran.us.r-project.org/web/packages/multcomp/index.html> [Accessed February 10, 2015]
- Kaiyala KJ, Morton GJ, Leroux BG, Ogimoto K, Wisse B & Schwartz MW (2010) Identification of Body Fat Mass as a Major Determinant of Metabolic Rate in Mice. *Diabetes* **59**: 1657–1666
- Kaji H, Hisamura M, Saito N, Sakai H, Aikawa T, Kondo T, Ide H & Murao M (1979) Clinical application of breath analysis for dimethyl sulfide following ingestion of DL-methionine. *Clin. Chim. Acta Int. J. Clin. Chem.* **93**: 377–380
- Karinje KU & Ogata M (1990) Methanol metabolism in acatalasemic mice. *Physiol. Chem. Phys. Med. NMR* **22**: 193–198

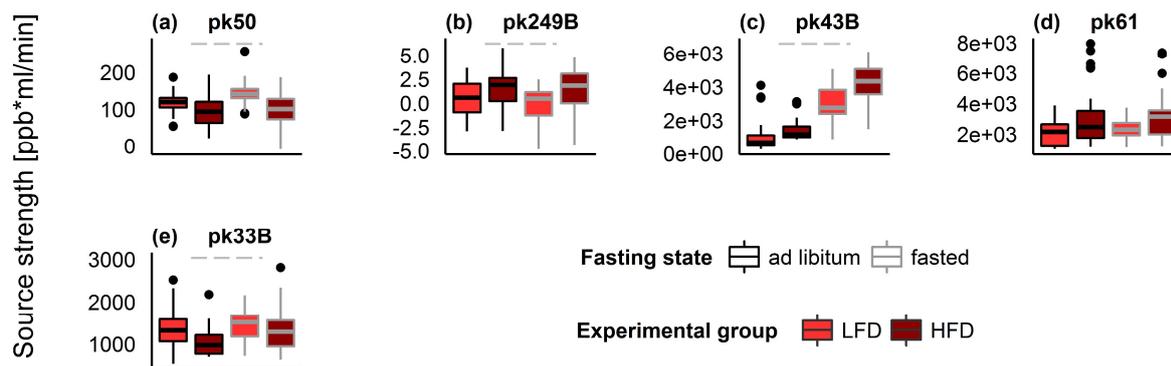
- King J, Mochalski P, Kupferthaler A, Unterkofler K, Koc H, Filipiak W, Teschl S, Hinterhuber H & Amann A (2010) Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. *Physiol. Meas.* **31**: 1169
- Kistler M, Szymczak W, Fedrigo M, Fiamoncini J, Höllriegl V, Hoeschen C, Klingenspor M, Hrabě de Angelis M & Rozman J (2014) Effects of diet-matrix on volatile organic compounds in breath in diet-induced obese mice. *J. Breath Res.* **8**: 016004
- Krumsiek J, Suhre K, Evans AM, Mitchell MW, Mohny RP, Milburn MV, Wagele B, Romisch-Margl W, Illig T, Adamski J, Gieger C, Theis FJ & Kastenmuller G (2012) Mining the Unknown: A Systems Approach to Metabolite Identification Combining Genetic and Metabolic Information. *PLoS Genet.* **8**: Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3475673/> [Accessed July 31, 2013]
- Krumsiek J, Suhre K, Illig T, Adamski J & Theis FJ (2011) Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Syst. Biol.* **5**: 21
- Kursa M & Rudnicki W (2010) Feature Selection with the Boruta Package. *J. Stat. Softw.* **36**: 1–13
- Layden BT, Yalamanchi SK, Wolever TM, Dunaif A & Lowe WL (2012) Negative association of acetate with visceral adipose tissue and insulin levels. *Diabetes Metab. Syndr. Obes. Targets Ther.* **5**: 49–55
- Lee J, Ngo J, Blake D, Meinardi S, Pontello AM, Newcomb R & Galassetti PR (2009) Improved predictive models for plasma glucose estimation from multi-linear regression analysis of exhaled volatile organic compounds. *J. Appl. Physiol.* **107**: 155–160
- Lewis GD, Laufman AK, McAnalley BH & Garriott JC (1984) Metabolism of acetone to isopropyl alcohol in rats and humans. *J. Forensic Sci.* **29**: 541–549
- Lindinger W, Hansel A & Jordan A (1998) On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research. *Int. J. Mass Spectrom. Ion Process.* **173**: 191–241
- Lindinger W, Taucher J, Jordan A, Hansel A & Vogel W (1997) Endogenous Production of Methanol after the Consumption of Fruit. *Alcohol. Clin. Exp. Res.* **21**: 939–943
- Mahendran Y, Vangipurapu J, Cederberg H, Stančáková A, Pihlajamäki J, Soininen P, Kangas AJ, Paananen J, Civelek M, Saleem NK, Pajukanta P, Lusi AJ, Bonnycastle LL, Morken MA, Collins FS, Mohlke KL, Boehnke M, Ala-Korpela M, Kuusisto J & Laakso M (2013) Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. *Diabetes* Available at: <http://diabetes.diabetesjournals.org/content/early/2013/04/01/db12-1363> [Accessed April 11, 2013]
- Martinez-Lozano Sinues P, Tarokh L, Li X, Kohler M, Brown SA, Zenobi R & Dallmann R (2014) Circadian Variation of the Human Metabolome Captured by Real-Time Breath Analysis. *PLoS ONE* **9**: Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4278702/> [Accessed February 20, 2015]
- Minh TDC, Oliver SR, Flores RL, Ngo J, Meinardi S, Carlson MK, Midyett J, Rowland FS, Blake DR & Galassetti PR (2012) Noninvasive Measurement of Plasma Triglycerides and Free Fatty Acids from Exhaled Breath. *J. Diabetes Sci. Technol.* **6**: 86–101
- Minh TDC, Oliver SR, Ngo J, Flores R, Midyett J, Meinardi S, Carlson MK, Rowland FS, Blake DR & Galassetti PR (2011) Noninvasive measurement of plasma glucose from exhaled breath in healthy and type 1 diabetic subjects. *Am. J. Physiol. Endocrinol. Metab.* **300**: E1166–1175
- Mochalski P, Al-Zoairy R, Niederwanger A, Unterkofler K & Amann A (2014) Quantitative analysis of volatile organic compounds released and consumed by rat L6 skeletal muscle cells *in vitro*. *J. Breath Res.* **8**: 046003
- Moghe A, Ghare S, Lamoreau B, Mohammad M, Barve S, McClain C & Joshi-Barve S (2015) Molecular Mechanisms of Acrolein Toxicity: Relevance to Human Disease. *Toxicol. Sci.* **143**: 242–255

- Morisco F, Aprea E, Lembo V, Fogliano V, Vitaglione P, Mazzone G, Cappellin L, Gasperi F, Masone S, De Palma GD, Marmo R, Caporaso N & Biasioli F (2013) Rapid ‘Breath-Print’ of Liver Cirrhosis by Proton Transfer Reaction Time-of-Flight Mass Spectrometry. A Pilot Study. *PLoS One* **8**: e59658
- Musa-Veloso K, Likhodii SS, Rarama E, Benoit S, Liu Y-MC, Chartrand D, Curtis R, Carmant L, Lortie A, Comeau FJE & Cunnane SC (2006) Breath acetone predicts plasma ketone bodies in children with epilepsy on a ketogenic diet. *Nutr. Burbank Los Angel. Cty. Calif* **22**: 1–8
- Novak BJ, Blake DR, Meinardi S, Rowland FS, Pontello A, Cooper DM & Galassetti PR (2007) Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proc. Natl. Acad. Sci.* **104**: 15613–15618
- Opgen-Rhein R & Strimmer K (2006) Inferring gene dependency networks from genomic longitudinal data: a functional data approach. *RevStat* **4**: 53–65
- Petersen TH, Williams T, Nuwayhid N & Harruff R (2012) Postmortem Detection of Isopropanol in Ketoacidosis: POSTMORTEM DETECTION OF ISOPROPRANOL IN KETOACIDOSIS. *J. Forensic Sci.* **57**: 674–678
- Petersson F, Sulzer P, Mayhew CA, Watts P, Jordan A, Märk L & Märk TD (2009) Real-time trace detection and identification of chemical warfare agent simulants using recent advances in proton transfer reaction time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **23**: 3875–3880
- Phillips M, Herrera J, Krishnan S, Zain M, Greenberg J & Cataneo RN (1999) Variation in volatile organic compounds in the breath of normal humans. *J. Chromatogr. B. Biomed. Sci. App.* **729**: 75–88
- Pinheiro J, Bates D, DebRoy S, Deepayan S & R-core (2015) nlme: Linear and Nonlinear Mixed Effects Models Available at: <http://cran.us.r-project.org/web/packages/nlme/index.html> [Accessed February 10, 2015]
- Ploner A (2014) Heatplus: Heatmaps with row and/or column covariates and colored clusters
- Qiao Y, Gao Z, Liu Y, Cheng Y, Yu M, Zhao L, Duan Y & Liu Y (2014) Breath Ketone Testing: A New Biomarker for Diagnosis and Therapeutic Monitoring of Diabetic Ketosis. *BioMed Res. Int.* **2014**: Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4037575/> [Accessed January 22, 2015]
- Raman M, Ahmed I, Gillevet PM, Probert CS, Ratcliffe NM, Smith S, Greenwood R, Sikaroodi M, Lam V, Crotty P, Bailey J, Myers RP & Rioux KP (2013) Fecal Microbiome and Volatile Organic Compound Metabolome in Obese Humans With Nonalcoholic Fatty Liver Disease. *Clin. Gastroenterol. Hepatol.* **11**: 868–875.e3
- R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria Available at: <http://www.R-project.org/>
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C & Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**: 77
- Romano A, Fischer L, Herbig J, Campbell-Sills H, Coulon J, Lucas P, Cappellin L & Biasioli F (2014) Wine analysis by FastGC proton-transfer reaction-time-of-flight-mass spectrometry. *Int. J. Mass Spectrom.* **369**: 81–86
- Rosenthal N & Brown S (2007) The mouse ascending: perspectives for human-disease models. *Nat. Cell Biol.* **9**: 993–999
- Sakakibara I, Fujino T, Ishii M, Tanaka T, Shimosawa T, Miura S, Zhang W, Tokutake Y, Yamamoto J, Awano M, Iwasaki S, Motoike T, Okamura M, Inagaki T, Kita K, Ezaki O, Naito M, Kuwaki T, Chohnan S, Yamamoto TT, et al (2009) Fasting-Induced Hypothermia and Reduced Energy Production in Mice Lacking Acetyl-CoA Synthetase 2. *Cell Metab.* **9**: 191–202
- Schaefer J, Opgen-Rhein R & Strimmer K (2015) GeneNet: Modeling and Inferring Gene Networks Available at: <http://CRAN.R-project.org/package=GeneNet>

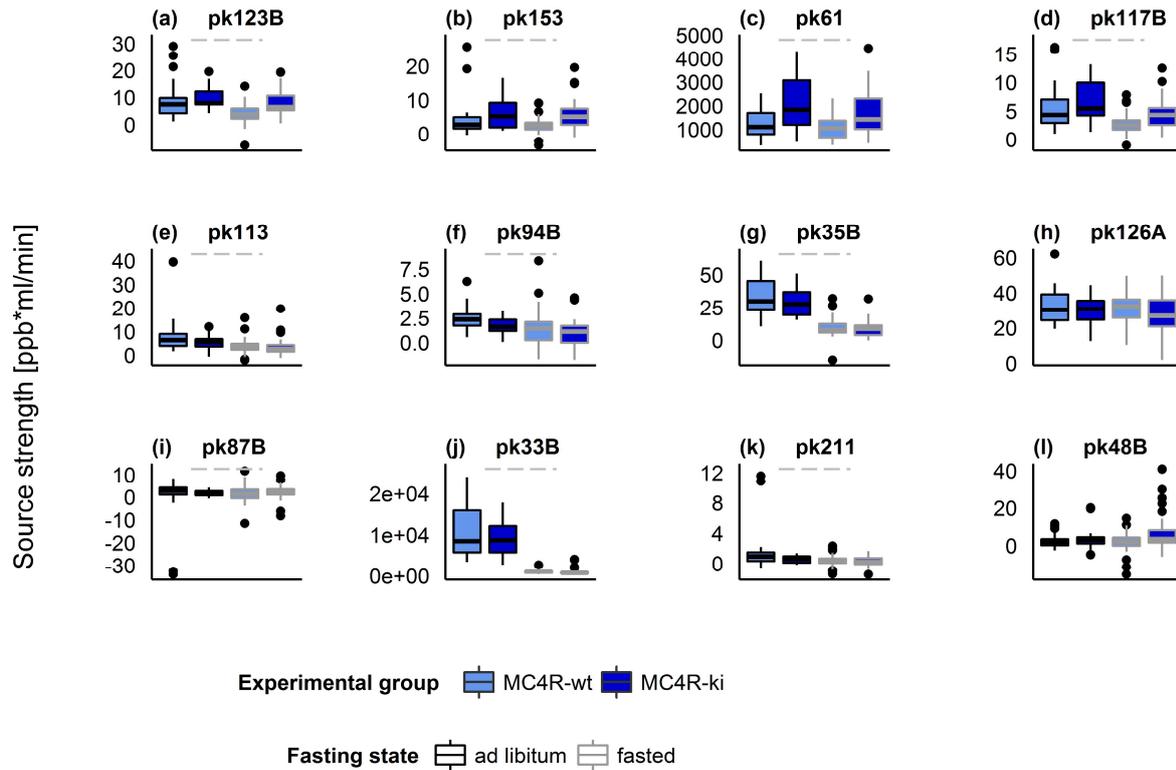
- Schäfer J, Opgen-Rhein R & Strimmer K (2001) Reverse engineering genetic networks using the GeneNet package. *J Am Stat Assoc* **96**: 1151–1160
- Schäfer J & Strimmer K (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* **4**: Available at: <http://www.degruyter.com/view/j/sagmb.2005.4.1/sagmb.2005.4.1.1175/sagmb.2005.4.1.1175.xml> [Accessed August 21, 2014]
- Schwarz K, Filipiak W & Amann A (2009) Determining concentration patterns of volatile compounds in exhaled breath by PTR-MS. *J. Breath Res.* **3**: 027002
- Siragusa RJ, Cerda JJ, Baig MM, Burgin CW & Robbins FL (1988) Methanol production from the degradation of pectin by human colonic bacteria. *Am. J. Clin. Nutr.* **47**: 848–851
- Smolinska A, Hauschild A-C, Fijten RRR, Dallinga JW, Baumbach J & Schooten FJ van (2014) Current breathomics—a review on data pre-processing techniques and machine learning in metabolomics breath analysis. *J. Breath Res.* **8**: 027105
- Solga SF, Mudalel M, Spacek LA, Lewicki R, Tittel FK, Loccioni C, Russo A, Ragnoni A & Risby TH (2014) Changes in the concentration of breath ammonia in response to exercise: a preliminary investigation. *J. Breath Res.* **8**: 037103
- Solga SF, Mudalel M, Spacek LA, Lewicki R, Tittel F, Loccioni C, Russo A & Risby TH (2013) Factors influencing breath ammonia determination. *J. Breath Res.* **7**: 037101
- Stevens JF & Maier CS (2008) Acrolein: sources, metabolism, and biomolecular interactions relevant to human health and disease. *Mol. Nutr. Food Res.* **52**: 7–25
- Szymczak W, Rozman J, Höllriegl V, Kistler M, Keller S, Peters D, Kneipp M, Schulz H, Hoeschen C, Klingenspor M & de Angelis MH (2014) Online breath gas analysis in unrestrained mice by hs-PTR-MS. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **25**: 129–140
- Tang C, Ahmed K, Gille A, Lu S, Gröne H-J, Tunaru S & Offermanns S (2015) Loss of FFA2 and FFA3 increases insulin secretion and improves glucose tolerance in type 2 diabetes. *Nat. Med.* **21**: 173–177
- Tani A, Hayward S & Hewitt CN (2003) Measurement of monoterpenes and related compounds by proton transfer reaction-mass spectrometry (PTR-MS). *Int. J. Mass Spectrom.* **223–224**: 561–578
- Tschöp MH, Speakman JR, Arch JRS, Auwerx J, Brüning JC, Chan L, Eckel RH, Farese Jr RV, Galgani JE, Hambly C, Herman MA, Horvath TL, Kahn BB, Kozma SC, Maratos-Flier E, Müller TD, Münzberg H, Pfluger PT, Plum L, Reitman ML, et al (2012) A guide to analysis of mouse energy metabolism. *Nat. Methods* **9**: 57–63
- Turner C, Spanel P & Smith D (2006) A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol. Meas.* **27**: 637–648
- Urrea V & Calle ML (2012) AUCRF: Variable Selection with Random Forest and the Area Under the Curve Available at: <http://cran.us.r-project.org/web/packages/AUCRF/index.html> [Accessed February 10, 2015]
- Van den Velde S, Nevens F, Van Hee P, van Steenberghe D & Quirynen M (2008) GC-MS analysis of breath odor compounds in liver patients. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **875**: 344–348
- Van der Ploeg LHT, Martin WJ, Howard AD, Nargund RP, Austin CP, Guan X, Drisko J, Cashen D, Sebat I, Patchett AA, Figueroa DJ, DiLella AG, Connolly BM, Weinberg DH, Tan CP, Palyha OC, Pong S-S, MacNeil T, Rosenblum C, Vongs A, et al (2002) A role for the melanocortin 4 receptor in sexual function. *Proc. Natl. Acad. Sci. U. S. A.* **99**: 11381–11386
- Wahl S, Vogt S, Stückler F, Krumsiek J, Bartel J, Kacprowski T, Schramm K, Carstensen M, Rathmann W, Roden M, Jourdan C, Kangas AJ, Soininen P, Ala-Korpela M, Nöthlings U, Boeing H, Theis FJ, Meisinger C, Waldenberger M, Suhre K, et al (2015) Multi-omic signature of body weight change: results from a population-based cohort study. *BMC Med.* **13**: 48

- Wickham H (2009) *ggplot2: elegant graphics for data analysis* Springer New York Available at: <http://had.co.nz/ggplot2/book>
- Wolever T, Josse RG, Leiter LA & Chiasson J-L (1997) Time of day and glucose tolerance status affect serum short-chain fatty concentrations in humans. *Metabolism* **46**: 805–811

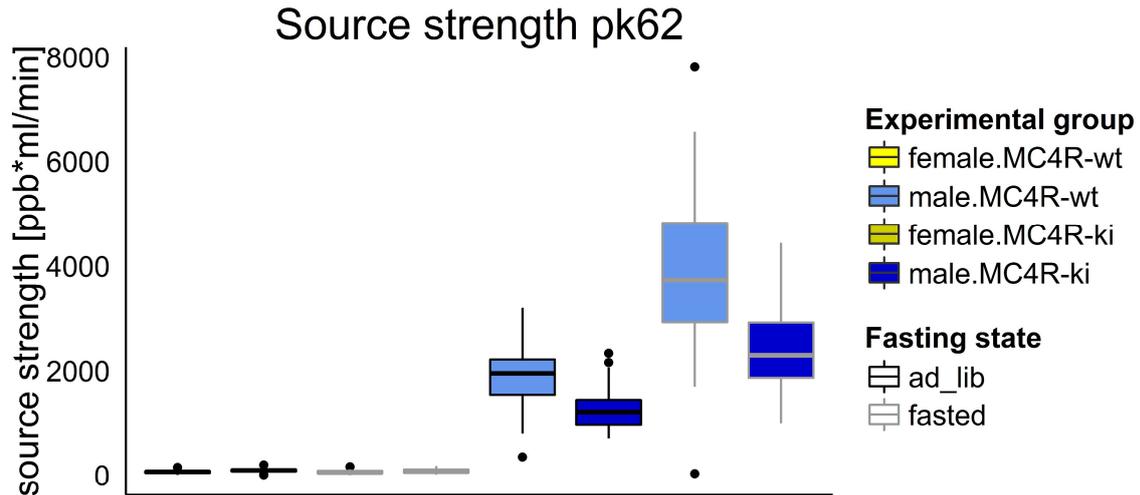
8.3.6. Supplements



Supplementary figure 1: VOC source strengths without significant high fat diet effects. Source strengths for nominal mass-labelled peaks 50 (a), 249B (b), 43B (c), 61 (d) and 33B (e) are shown as boxplots (ordered after selection probability in cross-validated AUC-RF algorithm). Box fill corresponds to diet (red: low fat diet; dark red: high fat diet). Box border corresponds to fasting state (black: ad libitum fed; grey: fasted). Significant main effects in mixed effects model are shown as dotted lines (black: diet, grey: fasting state). In case of interaction, significant group differences are shown as black lines. *P*-values are adjusted for a false discovery rate of 10 %. Group sizes (*n*=9).



Supplementary figure 2: VOC source strengths without significant genotype effects. Source strengths for nominal mass-labelled peaks 117B (a), 61 (b), 33B (c), 123B (d), 153 (e), 113 (f), 94B (g), 35B (h), 126A (i), 87B (j), 211 (k) and 48B (l) shown as boxplots (ordered after selection probability in AUC-RF feature selection). Box fill corresponds to genotype (blue: melanocortin-4-receptor wild type; dark blue: melanocortin-4-receptor W16X knock-in). Box border corresponds to fasting state (black: ad libitum fed; grey: fasted). Significant main effects in mixed effects model are shown as dotted lines (black: genotype, grey: fasting state). In case of interaction, significant group differences are shown as black lines. P-values are adjusted for a false discovery rate of 10 %. Group sizes: MC4R-ki fasted (n=15), MC4R-wt fasted (n=14), MC4R-wt ad lib (n=11), MC4R-ki ad lib (n=9).



Supplementary figure 3: Source strengths of pk62 in males and females. Source strengths of pk62 (MTMT) are shown as boxplots. Box fill corresponds to genotype and sex (blue: male melanocortin-4-receptor wild type; dark blue: male melanocortin-4-receptor W16X knock-in, yellow: female melanocortin-4-receptor wild type; dark yellow: female melanocortin-4-receptor W16X knock-in). Box border corresponds to fasting state (black: ad libitum fed; grey: fasted).

Supplementary table 1: Linear mixed effects model testing showing genotype, diet and fasting effects on VOC source strengths

peak	fasted								ad libitum fed								Lme diet & fasting status			Posthoc-test diet & fasting status				Lme genotype & fasting status			Posthoc-test genotype & fasting status					
	mean LFD		mean HFD		mean MCR-wt		mean MCR-ki		mean LFD		mean HFD		mean MCR-wt		mean MCR-ki		diet	feed state	diet:feed state	ad-lib LowFat	fasted LowFat	fasted HighFat	ad-lib HighFat	genotype	feed state	ff: feed state	ad-lib MCR-BLA	ad-lib BLA-fasted	ad-lib MCR-ki	fasted MCR-ki		
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd																
pk18	745175.43	385208.81	110462.5	194262.69	930685.55	1149901.07	915767.67	1117265.36	161576.84	45077.2	45369.32	23502.63	132291.94	93456.28	202016.44	99732.22	3.90E-05	1.24E-20	4.91E-09	7.50E-13	0.00	E+00										
pk47B	58.04	23.6	120.54	50.68	56.61	36.15	73.53	45.84	67.17	17.64	91.52	28.7	22.2	16.22	32.85	22.74	2.45E-05	1.34E-02	1.13E-04	1.11E-09	1.06E-02	1.39E-03	2.25E-02									
pk45B	23.4	7.06	35.52	13.5	11.12	4.58	13.38	4.07	30.44	9.26	42.99	12.55	7.88	3.25	13.24	7.15	2.66E-02	5.65E-01	6.10E-01													
pk43	1566.95	414.1	1958.68	488.37	1012.43	165.48	1438.3	284.57	1687.35	569.59	2502.49	538.94	845.22	202.71	1295.46	255.99	1.33E-02	3.09E-01	2.37E-01													
pk47S	96.23	31.89	147.73	65.06	185.74	132.46	385.89	257.26	83.12	18.64	117.01	44.55	59.21	49.18	99.84	97.73	4.83E-03	8.84E-02	9.53E-02													
pk462	1027.11	408.79	1952.28	576.09	1591.08	533.48	996.38	225.29	2891.28	1115.26	3121.95	929.57	2886.29	911.32	1599.94	693.56	2.68E-04	7.07E-26	1.18E-06	1.53E-06	0.00	E+00										
pk450	117.52	28.26	93.19	40.42	111.3	91.98	66.2	41.19	144.1	35.54	103.1	47.26	116.03	39.12	90.02	37.87	4.85E-02	3.86E-02	2.87E-01													
pk249 B	0.5	1.72	1.45	1.91	0.34	0.42	0.34	0.52	-0.34	2.2	1.33	2.36	0.18	0.76	0.19	0.81	5.52E-01	3.80E-02	2.66E-01													
pk253 A	0.64	1.55	2.28	1.83	-0.23	1.24	0	0.54	0.76	2.01	1.98	1.98	0.07	0.64	0.24	0.78	6.74E-03	8.78E-01	7.83E-01													
pk117 B	6.23	3.72	10.93	4.2	5.46	3.85	6.79	3.67	5.81	4.02	7.46	4.01	2.75	1.67	4.53	2.6	3.89E-03	3.12E-01	1.22E-01													
pk47A	-0.6	4.53	1.82	4.31	3.28	7.33	2.58	1.16	0.55	5.18	3.4	3.2	1.39	3.05	-0.14	2.2	5.45E-02	2.34E-01	9.36E-01													
pk43A	-494.39	119.32	-603.33	115.66	-759.1	337.44	-656.4	137.48	-558.4	117	-665.69	114	-621.17	174.63	-687.96	199.8	1.88E-03	2.13E-02	9.94E-01													
pk41B	8.22	6.6	12.81	6.72	21.17	30.5	46.02	38.35	5	6.7	8.79	5.49	10.81	11.03	30.1	28.15	5.42E-02	4.14E-03	4.70E-01													
pk47B	2.96	3.47	7.52	3.04	9.49	5.09	12.35	4.65	2.91	4.55	4.57	4.63	6.75	3.87	7.43	4.95	1.53E-02	4.33E-01	2.68E-01													
pk45B	4999.52	2046.66	7316.99	4020.37	1164.13	851.25	934.84	741.57	7022.74	1540.06	8378.98	2882.74	854.03	608.39	913.28	845.59	2.66E-02	2.17E-04	1.48E-01													
pk430	243.8	78.58	176.87	100.91	60.58	48.76	59.34	41.74	189.06	44.65	170.92	35.73	20.57	14.59	26.51	13.75	8.79E-03	1.89E-06	1.58E-04	2.22E-03	3.00E-07	6.55E-01	7.18E-01									
pk41	2678.78	835.65	5094.76	1938.12	1246.08	979.7	2846.36	1179.71	2540.97	701.24	3223.43	1475.24	1041.86	403.09	1639.27	918.77	7.41E-02	1.13E-01	5.41E-01													
pk43A	1330	544.13	2166.76	1556.78	1048.61	522.98	1511.81	866.94	1999.87	490.29	2224.13	1115.73	994.38	435.81	1538.85	973.67	4.50E-02	1.34E-02	1.97E-01													
pk43B	1325.42	479.39	1038.93	346.57	1076.25	996.05	950.39	4344.93	1443.82	562.2	1339.73	495.97	997.49	391.96	957.04	788.17	1.86E-01	3.91E-02	2.79E-01													
pk459	252.32	166.83	747.21	639.59	215.81	101.58	232.03	104.25	812.26	321.44	1047.4	371	822.79	501.83	1365.93	1770.97	1.45E-02	2.14E-10	1.35E-01													
pk44B	1745	7.11	24.25	8.48	15.65	5.41	18.15	4.03	27.73	9.69	33.42	9	20.48	6.75	21.1	6.06	1.97E-02	2.14E-04	4.82E-01													
pk423 B	5.77	5.29	8.37	3.19	8.6	6.95	9.9	4.26	4.53	5.36	6.39	4.2	4.19	2.97	7.81	4.52																
pk451	4.09	4.01	6.65	4.44	8.03	13.06	16.52	15.42	3.83	3.13	5.67	3.64	4.38	5.59	13.2	14.35																
pk4153	4.84	4.26	5.72	4.52	4.91	6.89	6.46	5.05	3.22	4.47	4.08	3.43	2.52	2.26	5.34	4.43																
pk415 B	2.07	4.78	2.68	4.52	4.13	5.53	7.66	7.66	3.31	4.18	2.1	3.76	2.85	3.5	6.82	7.03																
pk436A	97.32	35.21	112.33	26.4	93.85	41.13	70.39	30.18	127.23	28.42	116.45	30.56	77.87	25.99	73.4	39.38																
pk422 B	2	3	4.35	2.42	0.55	1.42	1.52	1.09	4.95	2.64	3.8	3.23	0.41	1.1	1.22	1.05																
pk437 B	3.34	5.58	8.01	6.95	18.07	27.81	40.14	38.89	5.82	4.31	7.43	4.99	8.7	10.27	26.7	27.13																
pk444A	34.65	20.03	52.17	51.14	13.91	21.52	33.19	20.6	51.02	54.42	66.34	45.62	19.08	16.18	35.59	25.51																
pk4113	16.7	5.96	15.6	10.31	8.83	9.23	5.66	3.28	13.97	6.53	16.2	7.9	3.92	3.05	3.26	3.66																
pk494B	2.45	3.73	1.98	4.11	2.46	1.14	1.71	0.84	2.82	3.23	3.41	3.22	1.39	1.68	0.94	1.35																
pk415B	46.17	15.02	39.87	19.75	34.12	15.33	29.65	10.73	48.98	12.59	45.97	16.59	9.31	6.82	8.39	5.94																
pk488B	13.64	7.2	23.93	13.22	5.13	4.47	3.15	1.57	34	12.62	39.14	15.63	4.47	3.58	4.76	4.73																
pk4126 A	55.96	13.37	61	14.63	31.99	9.58	29.9	8.93	74.33	13.35	72.87	14.58	30.76	8.85	27.4	10.35																
pk487B	4.23	4.93	11.26	5.64	0.28	9.75	1.83	1.27	8.88	4.68	8.3	6.75	1.6	3.58	2.16	3.07																
pk4109 A	91.58	32.1	99.61	28.37	65.43	22.79	54.86	20.6	126.49	20.58	126.5	23.55	57.22	17.14	56.09	21.84																
pk495B	7.67	2.87	9.21	6.1	0.9	4.75	3.61	2.63	5.77	4.34	6.96	5.08	2.5	2.45	4.34	5.23																
pk4211	1.74	3.96	2.14	1.61	1.56	2.81	0.55	0.47	2.62	2.48	2.64	2.31	0.36	0.73	0.27	0.64																
pk488B	4.62	8.83	12	16.7	2.44	3.72	3.63	8.18	3.94	5.26	3.91	10.2	1.97	5.09	7.85	11.07																

Supplementary table 2: Tentative assignment of peaks with significant diet, genotype or fasting effects and interpretation of partial correlations to other peaks as depicted in the gaussian graphical model (figure 6).

Subnet	peak 1	formula / assignment(s) (mass)	prot. mass	peak 2	formula / assignment(s) (mass)	prot. mass	partial cor.	Interpretation/ pathway
1 / red	18	NH ₃ .H ⁺ Ammonia (18.03)	18.03	46C	C ₂ H ₇ N structure: CH ₃ CH ₂ NH ₂ .H ⁺ / Ethylamine (46.07)	46.06	0.086	(bio)chemical reaction
	63	CO ₂ .H ₃ O ⁺ / Carbon dioxide water cluster (63.00) C ₂ H ₆ S.H ⁺ / Dimethyl sulphide (63.02)	63.01	64B	¹³ CO ₂ . H ₃ O ⁺ / Carbon dioxide water cluster (64.00) Dimethyl sulphide (64.03) / ¹³ CCH ₆ S.H ⁺	64.01	0.081	C-13 Carbon isotope
	62	CH ₃ SCH ₂ .H ⁺ / (Methylthio)methanethiol fragment (62.02)	62.02	64B	CH ₃ ³⁴ SCH ₂ .H ⁺ / MTMT fragment (64.02)	64.01	0.074	S-34 Sulphur isotope
				44B	¹³ CCH ₃ O.H ⁺ / propanol fragment (44.07)	44.07	0.077	Both increase in fasting/ morning
				59	C ₃ H ₆ O.H ⁺ / acetone (59.05)	59.05	0.083	Both increase in fasting/ morning
				45B	C ₂ H ₄ O.H ⁺ / Acetaldehyde (45.03) C ₃ H ₈ .H ⁺ / Propane (45.07)	45.05	0.078	Unknown/ (Bio) chemical reaction
				42	¹³ CC ₂ H ₄ .H ⁺ / propanol fragment (42.05)	42.05	0.081	Carbon isotope of other propanol fragment
				27	C ₂ H ₂ .H ⁺ (27.02)	27.01	0.084	Fragment
				59	C ₃ H ₆ O.H ⁺ / acetone (59.05)	59.05	0.091	(Bio) chemical reaction catalyzed by alcohol dehydrogenase (adh1)
	2 / orange	43B	C ₃ H ₆ .H ⁺ (43.06) / propanol fragment (after -OH loss)	43.07	44B	¹³ CCH ₃ O.H ⁺ / propanol fragment (44.07)	44.07	0.108
				41	C ₃ H ₄ .H ⁺ / propadiene, propanol fragment (41.04)	41.05	0.204	Alternative propanol fragment
				60B	¹³ CC ₂ H ₆ O.H ⁺ / acetone (60.05)	60.05	0.077	C-13 Carbon isotope
				62	CH ₃ SCH ₂ .H ⁺ / (Methylthio)methanethiol fragment (62.02)	62.02	0.083	Both increase in fasting/ morning
				41	C ₃ H ₄ .H ⁺ / propadiene, propanol fragment (41.04)	41.05	0.089	(Bio) chemical reaction catalyzed by alcohol dehydrogenase (adh1)
				43B	C ₃ H ₆ .H ⁺ (43.06) / propanol fragment (after -OH loss)	43.07	0.091	(Bio) chemical reaction catalyzed by alcohol dehydrogenase (adh1)
				44B	¹³ CCH ₆ .H ⁺ / propanol fragment (44.07)	44.07	0.094	(Bio) chemical reaction catalyzed by alcohol dehydrogenase (adh1)
				45B	C ₂ H ₄ O.H ⁺ / acetaldehyde (45.03) C ₃ H ₈ .H ⁺ / Propane (45.07)	45.05	0.096	(Bio) chemical reaction/ unknown
				59	C ₃ H ₆ O.H ⁺ / Acetone (59.05)	59.05		
3 / yellow					43A	C ₂ H ₃ O ⁺ / acylium ion (43.02)	43.03	0.087
				61	C ₂ H ₄ O ₂ .H ⁺ / acetic acid (61.03)	61.03	0.090	Biochemical pathway / chemical reaction product / educt
	75	C ₃ H ₆ O ₂ .H ⁺ Methyl acetate (75.04)	75.03	33B	CH ₃ OH.H ⁺ / methanol (33.03)	33.04	0.090	Biochemical pathway / chemical reaction product / educt
				60C	C ₃ H ₉ N.H ⁺ / Trimethylamine (60.08)	60.08	0.093	High in chow ad libitum / Both can be affected in presence of urine, e.g. during cleaning behavior
				76B	¹³ CC ₂ H ₆ O ₂ .H ⁺ / Methyl acetate (76.02)	76.02	0.095	Carbon isotope
	61	C ₂ H ₄ O ₂ .H ⁺ / acetic acid (61.03)	61.03	43A	C ₂ H ₃ O ⁺ / acylium ion (43.02)	43.03	0.157	Acetate fragment (-H ₂ O loss)

				95A	C ₂ H ₆ S ₂ .H ⁺ / dimethyl sulfone (94.99)	94.96	0.076	Both high in chow ad libitum / methyl group fragmentation
				27	C ₂ H ₂ .H ⁺ / fragment of propanol (27.02)	27.01	-0.079	Inverse response to fasting in propanol & acetone subnetwork (2) / possible reaction to propanol
	33B	CH ₃ OH.H ⁺ / methanol (33.03)	33.04	75	C ₃ H ₆ O ₂ .H ⁺ / Methyl acetate (75.04)	75.03	0.090	(Bio) chemical reaction
				60C	C ₃ H ₇ O.H ⁺ / ? (60.06)	60.08	0.096	Both high in chow ad libitum / Can be affected in presence of urine, e.g. during cleaning behavior
				34B	¹³ CH ₃ OH.H ⁺ / methanol (34.04)	34.05	0.095	Carbon isotope
				51B	CH ₃ OH.H ₃ O ⁺ / methanol water cluster (51.05)	51.05	0.113	Water cluster formation
4 /olive green	50	Not assigned	50.01	36A	Not assigned	36.04	0.090	Relation unknown
	36A	Not assigned	36.04	50	Not assigned	50.01	0.090	Relation unknown
				124A	Not assigned	123.84	0.108	Relation unknown
	109A	Not assigned	108.87	124A	Not assigned	123.84	0.074	Relation unknown
				126A	Not assigned	126??	0.078	Relation unknown
5 / green	57A	Not assigned	56.94	110A	Not assigned	109.87	0.074	Relation unknown
				198A	Not assigned	197.65	0.079	Relation unknown
6 / aquamarine	34A	¹⁷ O ₂ / oxygen isotope (34.00)	34.00	123A	Not assigned	122.85	0.078	Relation unknown
				213	Not assigned	213.94	0.078	Relation unknown
				58B	C ₂ H ₃ NO.H ⁺ (58.03) CH ₃ N ₃ .H ⁺ (58.04) C ₃ H ₃ O.H ⁺ (58.04)	58.03	0.076	Relation unknown
				151	Not assigned (pot. C ₁₀ H ₁₄ O (151.10))	150.99	0.099	Relation unknown
7 / cyan	81B	Not assigned (pot. C ₄ H ₄ N ₂ .H ⁺ / Pyrazine (81.05), C ₆ H ₈ .H ⁺ / hexenal fragment / monoterpene fragment (81.07))	81.04	137B	Not assigned (pot. C ₁₀ H ₁₆ .H ⁺ / Monoterpenes (137.14))	137.00	0.112	Relation unknown
				138B	Not assigned	138.01	0.080	Relation unknown
				137B	Not assigned (pot. C ₁₀ H ₁₆ .H ⁺ / Monoterpenes (137.14))	137.00	0.094	Relation unknown
	151	Not assigned (pot. C ₁₀ H ₁₄ O (151.10))	151.00	81B	Not assigned (pot. C ₄ H ₄ N ₂ .H ⁺ / Pyrazine (81.05), C ₆ H ₈ .H ⁺ / monoterpene fragment (81.07))	81.04	0.099	Relation unknown
				137B	Not assigned (pot. C ₁₀ H ₁₆ .H ⁺ / Monoterpenes (137.14))	137.00	0.088	Relation unknown
	153	Not assigned (pot. C ₁₀ H ₁₆ O (153.12))	152.99	96C	Not assigned	95.97	0.098	Relation unknown
				153	Not assigned (pot. C ₁₀ H ₁₆ O (153.12))	152.99	0.088	Relation unknown
	137B	Not assigned (pot. C ₁₀ H ₁₆ .H ⁺ / Monoterpenes (137.14))	137.00	151	Not assigned (pot. C ₁₀ H ₁₄ O (151.10))	151.00	0.094	Relation unknown
			81B	Not assigned (pot. C ₄ H ₄ N ₂ .H ⁺ / Pyrazine (81.05), C ₆ H ₈ .H ⁺ / hexenal fragment / monoterpene fragment (81.07))	81.04	0.112	Relation unknown	
8 / blue				249A	Not assigned	248.61	0.079	Relation unknown
	74B	Not assigned (pot. C ₂ H ₃ NS (74.01))	74.00	86B	Not assigned (C ₃ H ₃ NO ₂ (86.02))	86.02	0.095	Relation unknown
				87A	Not assigned (86B C-Isotope)	87.00	0.106	Relation unknown

9 / dark blue	123B	Not assigned (C ₇ H ₆ O ₂ .H ⁺ / Benzoic acid (123.04))	123.00	67B	C ₃ H ₂ N ₂ .H ⁺ (67.03) C ₃ H ₆ .H ⁺ (67.05)	67.04	0.131	Relation unknown
10 / purple	135B	Not assigned	134.98	68B	Not assigned (pot. C ₄ H ₅ N.H ⁺ (68.05))	68.05	0.077	Relation unknown
				253B	Not assigned	252.96	0.081	Relation unknown
11 / light pink	122B	Not assigned	121.98	128B	Not assigned	128.01	0.087	Relation unknown
12 / pink	95B	Not assigned (pot. C ₆ H ₆ O.H ⁺ e.g. Phenol (95.05) / DMSO ₂ (95.01))	95.04	31B	CH ₂ O.H ⁺ / Formaldehyde (31.02)	31.02	0.089	Relation unknown
No significant positive connections to other VOCs	35B	CH ₃ ¹⁶ OH.H ⁺ (35.04)	35.05					
	36B	Not assigned (NH ₃ .H ₃ O ⁺ (36.05))	36.07					
	44A	C ₂ H ₃ O (44.03) / ¹³ CCH ₂ O.H ⁺ (44.03)	44.03					
	55B	H ₃ O ⁺ .(H ₂ O) ₂ / Water cluster (55.04) C ₄ H ₆ .H ⁺ / butadiene as fragment of aldehydes (55.05)	55.05					
	57B	acrolein (2-propenal, C ₃ H ₄ O.H ⁺) (57.03)	57.03					
	65B	Unassigned (pot. C ₂ H ₆ ³⁴ S.H ⁺ / DMS sulphur isotope (65.03) C ¹⁸ OO.H ₃ O ⁺ oxygen isotope (65.01))	65.01					
	80B	Not assigned	80.01					
	87B	Not assigned	87.04					
	94B	Not assigned	94.00					
	113	Not assigned	112.98					
	117B	Not assigned	117.01					
	249B	Not assigned	248.94					
	253A	Not assigned	252.60					