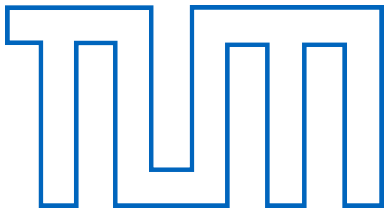


TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Medizin

Institut für Diabetes und Adipositas

RNA SEQUENCING AND DATA INTEGRATION IN THE CONTEXT OF NUTRITION AND WEIGHT LOSS INTERVENTION

VALENTINA SABINIA KLAUS

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende/-r:		Prof. Dr. Radu R. Rad
Prüfer der Dissertation:	1.	Prof. Dr. Paul Th. Pfluger
	2.	TUM Junior Fellow Dr. Jan Krumsiek

Die Dissertation wurde am 07.01.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 12.05.2020 angenommen.

Valentina Sabinia Klaus: *RNA Sequencing and data integration in the context of nutrition and weight loss intervention,* Dissertation, München, Dezember 2019

Dedicated to the loving memory of Hildegard Groß.
1938 – 2016

We are drowning in a sea of data and starving for knowledge.
– Sydney Brenner (et al.)

ABSTRACT

The intake of nutrients is not only one of the fundamental determinants of human health but also a major contributor in disease development and progression. This thesis aims to uncover biological underpinnings of health conditions caused by nutritional dysregulation, such as caloric overload or consumption of hazardous substances as well as the influence of the respective rescue interventions.
The first part of the thesis examines whether the transcriptional induction of genes involved in DNA repair or cancer-related pathways can be used to evaluate the effectiveness of cellular protection after cancerogen exposure and to assess the risk of DNA adducts to translate into tumourigenesis. The results indicate that genes involved in DNA repair are not efficient enough to protect the cell against the increasing numbers of DNA adducts after methylazoxymethanol acetate exposure and an elevated risk of these DNA adducts to manifest into tumour-causing mutations. However, combining the quantification of DNA adducts with time-matched gene expression analysis improves the predictive value for genotoxicant-induced carcinogenesis.
The second part of the thesis examines the molecular underpinnings of hypothalamic leptin sensitivity that was restored by pharmacological-induced weight loss by exendin-4 treatment in contrast to calorie restriction (CR). Hypothalamic gene expression profiles taken from mouse cohorts under five diet settings indicate that calorie-restricted mice undergo fasting-induced changes in fuel utilisation, orexin receptor signaling and transcription factor regulation. Despite similar loss of body weight, these changes in gene expression are absent in mice treated with exendin-4 thus indicating that the prevention of CR-induced hypometabolism might contribute to the superior restoration of leptin sensitivity seen in exendin-4 treated mice.
In the last part of the thesis, an integration method for omics datasets was developed to infer genetic control of metabolic networks, which was named Correlation based Network Integration (CoNI).
First, CoNI was applied to liver metabolome and transcriptome datasets of lean and HFD-fed obese mice to unravel previously hidden gene-metabolite interactions that exert major changes to hepatic metabolite levels under normal dietary conditions and under dietary stress resulting in steatosis. Validation of the selected regulator genes by siRNA-mediated knockdown and by transcriptional profiling in human liver biopsies confirmed that these genes have indeed relevant effects on hepatic metabolite levels. This demonstrates that our fully data-driven method can be used as a flexible and solid tool for multi-

ple omics data integration and interpretation.
Next, CoNI was applied to datasets from livers of mice undergoing weight loss by calorie restriction or Glp-1 agonism to identify genes that might be involved in the acute aggravation of hepatic steatosis by exendin-4 treatment. Two genes with so far unknown function in the regulation of liver metabolism during weight loss were identified that might advance our search for druggable targets.

Overall, this thesis presents three studies that highlight the complexity of molecular systems in response to nutritional dysregulation. They accentuate that the sophisticated integration of data from multiple sources such as transcriptome profiles together with quantified DNA adduct levels or metabolome profiles, as performed with our new data integration approach CoNI, is key to predict the cellular influences of potentially harmful substances or to identify new genes that could be used as druggable targets.

ZUSAMMENFASSUNG

Die Konzentration und Zusammensetzung der mit der Nahrung aufgenommenen Nährstoffe bestimmt nicht nur die Gesundheit, sondern trägt auch wesentlich zur Entstehung von Krankheiten und deren Verlauf bei. Das Ziel dieser Arbeit ist die Untersuchung der biologischen Grundlagen von Gesundheitszuständen, die durch eine dysregulierte Ernährung, wie beispielsweise überhöhte Kalorienzufuhr oder die Aufnahme von gefährlichen Substanzen verursacht wurden. Außerdem wird der Einfluss von entsprechenden Interventionen zur Wiederherstellung des Gesundheitszustandes untersucht.
Im ersten Teil dieser Arbeit wird untersucht, ob die transkriptionelle Induktion von Genen, die eine Rolle bei der DNS-Reparatur oder in Stoffwechselwegen, die bei Krebs reguliert sind, spielen, dazu verwendet werden kann die Effektivität von zellulären Schutzmechanismen zu bewerten nachdem der Organismus mit krebserregenden Substanzen exponiert wurde. Weiterhin soll abgeschätzt werden mit welchem Risiko die gebildeten DNS-Addukte zu Tumorigenese führen. Die neuen Ergebnisse weisen darauf hin, dass die Gene, die an Reparaturmechanismen der DNS beteiligt sind, die Zelle nicht effizient genug gegen eine ansteigende Anzahl an DNS-Addukten als Folge einer Methylazoxymethanol Acetat Exposition und das dadurch ansteigende Risiko einer Umwandlung der DNS-Addukte in tumorauslösende Mutationen schützen. Es konnte jedoch gezeigt werden, dass die Kombination der Quantifikation von DNS-Addukten mit der Analyse der darauf zeitlich abgestimmten Genexpression den Vorhersagewert der durch genotoxische Substanzen ausgelösten Karzinogenese verbessert.
Der zweite Teil dieser Arbeit untersucht die molekularen Grundlagen der hypothalamischen Leptinsensitivität, die durch den pharmakologisch induzierten Gewichtsverlust nach einer Exendin-4 Behandlung wiederhergestellt wurde. Dies wurde mit Kalorienrestriktion allein nicht erreicht. Die Auswertung der hypothalamischen Genexpressionsprofile, der Mäuse aus den fünf verschiedenen Diätregimen weist darauf hin, dass bei den kalorienrestriktierten Mäusen durch das Fasten eine Veränderungen bei der Nährstoffverwertung und der Orexinrezeptor-Signalwirkung sowie eine Regulation von Transkriptionsfaktoren erfolgt. Trotz eines vergleichbaren Gewichtsverlustes treten diese Veränderungen bei den Mäusen, die mit Exendin-4 behandelt wurden, nicht auf. Daraus lässt sich schließen, dass die Behandlung mit Exendin-4 den kalorienrestriktion-induzierten Hypometabolismus verhindert und so vermutlich zu einer verbesserten Wiederherstellung der Leptinsensitivität beiträgt.

Im letzten Teil dieser Arbeit wurde eine neue Analysemethode für Omics Datensätze entwickelt, die es ermöglicht, Rückschlüsse auf die genetische Kontrolle von metabolischen Netzwerken zu ziehen. Die Correlation based Network Integration (CoNI) Methode wurde zuerst auf Metabolom- und Transcriptom-Datensätzen, gewonnen aus Lebergewebe von dünnen Mäusen und dicken Mäusen, die mit einer hochkalorischen Diät gefüttert wurden, angewendet. Mit Hilfe von CoNI konnten vorher verborgene Gen-Metabolit Interaktionen aufgedeckt werden, die zwischen normaler Ernährungsweise und Überernährung, die auch zu einer Lebersteatose führen kann, große Veränderungen in den Metabolitkonzentrationen in der Leber verursachen. Die Validierung der ausgewählten, regulierenden Gene mittels siRNA vermittelten Gen-Knockdown in Zellkultur und durch transkriptionelles Profiling in humanen Leberbiopsien bestätigte, dass diese Gene maßgeblich die Metabolitkonzentrationen in der Leber beeinflussen. Unsere vollkommen daten-getriebene Methode ist somit ein flexibles und stabiles Werkzeug für die Integration und Interpretation verschiedener Omics Datensätze.
Anschließend wurde die CoNI Methode auf Datensätze von Mäusen angewendet, die durch Kalorienrestriktion oder durch die Behandlung mit einem Glp-1 Agonisten an Gewicht verloren, um Gene zu identifizieren, die zu der akuten Verschlimmerung einer Lebersteatose während der Behandlung mit Exendin-4 beitragen könnten. Es konnten zwei Gene identifiziert werden, von denen bisher keine Funktion in der Regulation des Lebermetabolismus während der Gewichtsabnahme bekannt war, die jedoch als potentielle Zielmoleküle für die Behandlung von übergewichts-induzierter Lebersteatose in Frage kommen könnten.

Insgesamt präsentiert diese Arbeit drei Studien, die zeigen wie komplex die Reaktion der molekularen Systeme auf eine Ernährungsdysregulation ist. Die Studien legen außerdem dar, dass die hochentwickelte Integration von Daten aus verschiedenen Quellen, beispielsweise Transkriptionsprofile mit DNS-Adduktkonzentrationen oder Metabolitprofilen, wie sie mit unserer neuentwickelten Methode zur Datenintegration CoNI durchgeführt wurde, entscheidend für die Vorhersage des Einflusses von möglicherweise schädlichen Substanzen auf die Zelle oder für die Entdeckung von neuen krankheitsrelevanten Zielmolekülen ist.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. Dominik Lutter for the continuous support of my PhD study and research, for his patience, encouragement, enthusiasm, and advice. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study.
It is a pleasure to thank Prof. Dr. Paul Pfluger and PD Dr. Kerstin Stemmer who initiated several projects that made this thesis possible. Besides my advisor and Prof. Dr. Pfluger, I would like to thank the rest of my thesis committee: Prof. Dr. Dr. Matthias Tschöp and Dr. Jan Krumsiek for their insightful comments and discussions.
The quality of the provided data is essential for the success of a project. Therefore, I am grateful that all collaboration partners, I had the honour to work with, provided high-quality data. Especially, I would like to thank the following collaboration partners: Dr. Thomas Schwarzmeier for mapping the RNA-Seq hypothalamus reads. Dr. Martin Irmler for providing the microarray quantifications and annotations for the liver samples. Dr. Janina Torkaz for providing the metabolome concentrations of the liver project. Heinke Bastek and PD Dr. Kerstin Stemmer for performing the animal work and wet lab experiments of the MAMAc project. Dr. Sonja Schriever for performing the animal work and wet lab experiments of the weight-loss intervention studies. Dr. Elisabeth Graf and her team for providing protein quantifications and their annotation for the antibody evaluation.
I am grateful for having met and worked with Dr. Sonja Schriever. I thank you for all the scientific discussions, for reviewing this thesis, for your advice in all situations, and for becoming an awesome friend that I would not like to miss.
I thank Kerstin H. and Steffi K. for being great mentors and friends.
Julia, I thank you for being my best friend, always listening, for our discussions and for reviewing this thesis.
Finally I would like to thank Steffi, Chris and especially my family for unlimited support during the time of my thesis.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

alpha CGRP Alpha calcitonin gene-related peptide

AA Amino acids

AAc Aristolochic acid

AB Antibody

Abca5 ATP binding cassette subfamily A member 5

Abhd2 Abhydrolase domain containing 2

AC Acylcarnitines

Acat1 Acetyl-CoA acetyltransferase 1

Acsf2 Acyl-CoA synthetase family member 2

Agrp Agouti-related neuropeptide

Appl2 Adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper containing 2

AR Abundance ratios

ARC Arcuate nucleus

Arhgap24 Rho GTPase activating protein 24

Arrdc2 Arrestin domain containing 2

ATP Adenosine triphosphate

BA Biogenic amine

BH Benjamini-Hochberg

BW Body weight

BWE Body weight measured at the end of the study

BMI Body mass index

CA Class abundance

Calca Calcitonin-related polypeptide alpha

Cart Cocaine- and amphetamine-regulated transcript

Cartpt Cart prepropeptide

Cd82 CD82 antigen

cDNA Complementary DNA

Cdkn1a Cyclin-dependent kinase inhibitor 1A

Ces Carboxylesterase

CNS Central nervous system

Cobll1 Cobl-like 1

ComAB Commercial Phospho-S6 Ribosomal Protein (Ser240/244) antibody

ComAB3P ComAB + 3P peptide

CoNI Correlation-based Network Integration

CR Calorie restriction

Cyba Cytochrome b-245, alpha polypeptide

Cyp Cytochrome P450s

Da Dalton

Ddit4 DNA-damage-inducible transcript 4

Ddx3x DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3, X-linked

DE Differentially expressed

DIO Diet-induced obesity

DNA Deoxyribonucleic acid

Eno1 Enolase 1, (alpha)

Ex4 Exendin-4

Eya3 Eyes absent homolog 3

FDR False discovery rate

FoxO Forkhead Box O

GEO Gene Expression Omnibus

Gk Glycerol kinase

GLM Generalised linear model

Glp-1 Glucagon-like peptide 1

Gm4553 Predicted gene 4553

GGM Gaussian graphical models

GO Gene Ontology

Gpr G-protein-coupled receptor

Grhpr Glyoxylate reductase/hydroxypyruvate reductase

Gzma Granzyme A

H>C HFD>Chow diet switch

HFD High fat diet

Hnrnpm Heterogeneous nuclear ribonucleoprotein M

HOMA-IR Homeostasis model assessment of insulin resistance

Hsd17b12 Hydroxysteroid (17-beta) dehydrogenase 12

Ifi47 Interferon gamma inducible protein 47

IgG Immunoglobulin G

IMPaLA Integrated Molecular Pathway Level Analysis

Inhbe Inhibin βE

INMEX INtegrative Meta-analysis of Expression data

i.p. Intraperitoneal

Jak2 Janus kinase 2

KD Knockdown

KEGG Kyoto Encyclopedia of Genes and Genomes

LPC Lyso-Phosphatidylcholines

LRb Leptin receptor

LRG Local regulator gene

Magix MAGI family member, X-linked

MAM Methylazoxymethanol

MAMAc Methylazoxymethanol acetate

MAPK Mitogen-activated protein kinase 1

Mgmt Methyl-guanin-DNA-Methyltransferase

mRNA Messenger RNA

Mrps12 Mitochondrial ribosomal protein S12

mTOR Mammalian target of the rapamycin

m/z Mass to charge

Myc Myelocytomatosis oncogene

N7-mG N7-methylguanine

NADPH Nicotinamide adenine dinucleotide phosphate

NAFLD Non-alcoholic fatty liver disease

NOAL No-adverse-effect-level

NormA Normalised abundance

Npy Neuropeptide Y

Nrbf2 Nuclear receptor binding factor 2

O6-mG O^6-methylguanine

Ogg1 8-oxoG-DNA glycosylase

OTA Ochratoxin A

PBS Phosphate buffered saline

PBSAB Phospho-S6 Ribosomal Protein (Ser240/244) antibody with the higher concentrated PBS formulation

PBSAB3P PBSAB + 3P peptide

PC Phosphatidylcholines

PCA Principal component analysis

PI3K Phosphatidylinositol 3-kinase

PLS Partial least squares

Plin4 Perilipin 4 (S3-12)

Pnpla2 Patatin-like phospholipase domain containing 2

Pomc Proopiomelanocortin

PRC Phosphorylated ribosome capture

Rapgef4 Rap guanine nucleotide exchange factor (GEF) 4

Rbp Ribosomal binding protein

RCC Renal cell carcinoma

Rhoc Ras homolog family member C

RNA Ribonucleic acid

RNA-Seq RNA-Sequencing

RpS6 Ribosomal protein S6

SerpinA3 Serpin Family A Member 3

Sgk1 Serum- and glucocorticoid-induced protein kinase-1

siRNA Small interfering RNA

SM Sphingomyelins

Smim13 Small integral membrane protein 13

SNP Signel nucleotide polymorphism

Socs3 Suppressor of cytokine signalling 3

Spsb1 SplA/ryanodine receptor domain and SOCS box containing 1

STAT Signal transducers and activators of transcription

Sult1a1 Sulfotransferase family member 1A

TAG Triacylglyceride

Tap1 Transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)

TGF beta Transforming growth factor β

TSC Tuberous sclerosis complex

VLDL Very Low Density Lipoprotein

Vnn1 Vanin-1

WT Wild type

Xpo7 Exportin 7

Zbtb16 Zinc finger and BTB domain-containing protein 16

1 INTRODUCTION

1.1 NUTRITION – THE CRITICAL DETERMINANT OF OUR HEALTH

"Dis-moi ce que tu manges, je te dirai ce que tu es."

Jean-Ánthelme Brillat-Savarin

This citation by Brillat-Savarin [32] translates into: tell me what you eat and I will tell you what you are. Despite being published in 1826 the quote has not lost any of its relevance; the nutritional intake is one of the fundamental determinants not only of the maintenance of human health but also a major contributor in disease development and progression [189]. Diet-induced obesity and its comorbidities such as diabetes and cardiovascular diseases are public health problems of the first order. Lifestyle and diet interventions could prevent these metabolic diseases that are driven by intake of excessive calories, diets with an imbalance of omega 3 and omega 6 fats or high in saturated fats, concentrated sugars, and refined flour products [61, 189]. Next to the directly associated comorbidities, obesity also raises the risk of malignancies since it was estimated that over 90,000 cancer-induced deaths per year could be prevented if the adult population sustained a normal body weight [61]. Therefore, eating too much increases the cancer risk but also consuming the wrong food increases cancer risk. There are several substances that could be present in consumed food items that are carcinogen from the beginning or become carcinogen during the preparation process that induce DNA adducts and increase the risk of developing tumours tremendously [1, 16, 20]. To understand the underlying biology of processes involved in obesity-induced metabolic dysregulation will help to identify new targets for the treatment of obesity related burdens and therefore analysing omics, such as transcriptomics, proteomics and metabolomics are good starting points.

Two of the main causes of death could partly be prevented by the right lifestyle and diet

1.2 CARCINOGENESIS

1.2.1 *Chemical carcinogenesis*

Humans are increasingly exposed to numerous chemicals

The production and use of chemicals are on the rise and humans are increasingly exposed to numerous chemicals present in various sources such as drugs, food additives, cosmetics or pesticides. While the use of many chemicals is highly beneficial for our society, they may bear toxic and carcinogenic side effects. A proper risk assessment of newly developed drugs and chemicals is mandatory in order to identify a potential hazard and to define safe exposure levels for humans. The latter requires detailed information about the mode of action of the test compound.

Chemical carcinogens can be grouped into two groups: (1) Genotoxic substances can covalently bind to DNA thereby causing DNA damage [174]. Subsequent failures in DNA repair mechanisms can cause alterations in the DNA sequence named mutations, which can be initiating events in tumourigenesis [137]; (2) Non-genotoxic carcinogens are cytotoxic or mitogenic substances that do not directly react with the DNA [181]. They either promote cell proliferation by directly stimulating mitogenic signalling cascades or by causing cell death and regenerative cell proliferation [174].

Depending on their classification, carcinogens have been subjected to different regulatory policies. Genotoxic carcinogens are considered to have no safe dose threshold and are assumed to pose a cancer risk at any dose subjected. Accordingly, they are currently banned for human use. In contrast, non-genotoxic carcinogens are thought to have a safe exposure threshold and their use can be permitted if the intake level does not exceed a defined acceptable daily intake level (ADI).

Recently, the above described concept got challenged and the question emerged if also genotoxic carcinogens may bear a safe exposure threshold. The latter is based on the finding that the translation of DNA adducts into tumourigenesis depends on various factors, for example the ability of the cell to repair the adducts. Therefore, the level of DNA adducts alone fails to predict a cancer risk.

Research efforts are centred around the development of short-term assays, which allow to study the mechanism of genotoxic carcinogens in more detail. Accumulating evidence suggests that it is possible to identify the carcinogenic potential of a chemical based on analysis of gene expression analyses (toxicogenomics).

Gene expression profiles can help to identify the toxicity of substances before the histopathological phenotype can be detected

The field of toxicogenomics emerged in the end of the 1990s, when microarray technology was applied to support toxicology experiments [3]. Nowadays, the term toxicogenomics summarises all functional genomics approaches being applied to investigate mechanisms of toxicology [3]. By monitoring thousands of genes simultaneously, toxicogenomics allows to characterise early gene expression patterns induced by environmental toxicants even before histopathological changes can be observed [236]. Accordingly, toxicogenomics has been used to predict toxicological outcomes, such as carcinogenicity.

Together with the numerous bioinformatics approaches that have been applied to analyse toxicogenomics data [3], they are of immense value to determine safety levels for compounds of interest [55], to distinguish genotoxic from non-genotoxic substances, and to identify the mode of action of a chemical [236].

1.2.2 *Renal carcinogenesis*

Kidneys are highly vulnerable to all toxicants circulating through the body

Kidneys play a predominant role in excretion and are therefore highly vulnerable to all toxicants circulating through the organism. The kidneys therefore represent one of the major target sites for chemical toxicity including genotoxic and non-genotoxic carcinogenicity. For 2018, Bray et al. [29] estimated that there will be 403,262 new kidney cancer cases worldwide, which corresponds to 2.2% of all new cancer cases. The most common kidney cancer is renal cell carcinoma (RCC) [213], which sporadically occurs mainly in adults [165]. It originates from the epithelium of the renal proximal tubule [165]. Early diagnoses and as a consequence early treatment of RCC turns out to be challenging and the primary tumour is often only detected due to the occurrence of metastases [165].

Cycasin is converted to the genotoxic metabolite MAM

A well-established renal carcinogen is cycasin, a naturally occurring, toxic azoxyglucoside produced by the cycad plant, which grows for example on the island of Guam. Several western pacific populations used to consume the flour made from cycad nuts and to apply the cycad seed for medical purposes [27]. In vivo, cycasin is converted to the genotoxic metabolite methylazoxymethanol (MAM) which is linked to neurotoxicity, hepatotoxicity, teratogenesis, and carcinogenicity [27, 107]. MAM is further converted to reactive intermediates such as methyldiazonium ions and carbon-centred free radicals [215]. These can interact with the DNA in the O^6-, *N7*-, and C8-positions of guanine, which leads to the accumulation of promutagenic *N7*-methylguanine (*N7*-mG) and O^6-methylguanine (O^6-mG) DNA adducts, causing the detrimental effects of MAM [215].

A long-term observation study conducted by the National Cancer Institute (NCI, Bethesda, MD) showed that MAM induced tumours on several sites, primarily liver and renal cell carcinomas in non-human primates [228]. In rat experiments the feeding of cycad nut flour causes cancer occurring in the digestive tract, liver, and kidney [168]. The synthetic MAM acetate (MAMAc), which is used to perform most studies, is converted into MAM directly after administration [260].

1.2.3 *Eker rats as model animals for kidney cancer*

Eker rats are heterozygous for a loss-of-function mutation in Tsc2 which makes them highly susceptible for RCC

Eker rats are heterozygous for a germ-line loss-of-function mutation in the rat homologue of the tuberous sclerosis complex 2 (*Tsc2*) tumour suppressor gene. A key function of the *Tsc2* gene is to adjust the mammalian target of rapamycin (mTOR) pathway in response to the availability of nutrients, cellular energy and growth factors [145]. In heterozygous Eker rats, spontaneous events or carcinogen exposure can lead to the loss of the second *Tsc2* allele and thus to an uncontrolled activation of mTOR and its downstream effectors [161]. The later has been linked to the development of tumours in several organs with highest incidences in the kidney [206]. In untreated Eker rats the incidence of RCC is complete by one year of age [159, 240]. Based on their heterozygous mutation, Eker rats are specifically prone to genotoxic and non-genotoxic carcinogens. The quantitative nature of lesion formation in the model allows the production of statistically powerful data to analyse the potency and relative degree of effects of short- or long-term exposure to chemicals of interest [159].

1.3 OBESITY, ITS CAUSES AND CONSEQUENCES

1.3.1 *Obesity - the lifestyle burden*

Obesity is a growing global health problem. Since 1975 the number of obese people has tripled resulting in 1.9 billion adults being overweight in 2016, which corresponds to 39% of all adults aged 18 years and older worldwide [239]. Overweight was defined by a body mass index (BMI: body weight in kilograms divided by the body height in meters squared) of 25 kg/m^2 or higher. Within the group of overweight people 13%, corresponding to 650 million people, were classified as obese with a BMI of more than 30 kg/m^2 [239]. Obesity leads to many comorbidities such as type 2 diabetes mellitus (T2DM) and cardiovascular diseases [98, 255]. Obesity is further a strong risk factor for the metabolic syndrome that is diagnosed by several criteria including abdominal obesity, insulin resistance, dyslipidemia, and hypertension [255]. The risk of developing the metabolic syndrome increases from a BMI of 21 kg/m^2 and reduces life expectancy dramatically – by seven years at the age of 40 – and greatly increases the disease and economic burden [98].

Symptoms of the metabolic syndrome reduce life expectancy dramatically and increase the disease and economic burden

It is widely accepted that the majority of obese cases results from a combination of increased food intake along with decreased energy expenditure, resulting in excessive fat accumulation [71]. Obesity is in most cases a preventable disease that depends largely on lifestyle. Nowadays, living a healthy lifestyle has become a challenge: In our everyday life we are faced by elevated levels of stress, the availability of palatable foods that are high in sugar, fat, and calories, also referred to as the "Western diet", and a reduction in physical activity due to the industrialisation [71]. In 2005, obesity was set to overtake smoking as the main preventable cause of illness and premature death in the USA [98]. While this is first and foremost of importance to affected individuals, the associated morbidity is economically damaging for society [98].

So far, behavioural and pharmacological treatments for obesity accomplish only a weight loss of 5 – 10% [209]. This weight loss is almost in all cases regained over time [209]. Therefore, the understanding of factors leading to obesity and its comorbidities, and the resulting identification of druggable targets and pathways can help to develop preventive and more effective strategies against this lifestyle burden and its associated comorbidities.

Treatments for obesity accomplish a weight loss of only 5 – 10%, which is usually regained

1.3.2 *Role of hypothalamus and hormones in the regulation of energy balance*

Energy intake and energy expenditure are the two sides of the energy balance coin, which is important for maintaining body weight. The brain senses the body's nutrient and energy levels and integrates this information to adjust the energy intake and expenditure accordingly [22, 160, 251]. Whole-body energy expenditure is orchestrated by several components, such as regulation of food intake, thermogenesis to maintain body temperature, basal metabolism, the thermic effect of food, and physical activity [22]. Each component has its own effector pathways and neural controls [22].

Brain regions involved in the control of appetite and therefore energy intake are the hypothalamus, the corticolimbic system, and the hindbrain, also referred to as brainstem [22]. Whereas the brainstem is mainly responsible for controlling the size of the meals, the corticolimbic system provides the cognitive, emotional and executive support for ingestive behaviour [22]. Since the neural processing within these core processor units happens outside awareness it is relatively inaccessible to conscious manipulation [22].

The hypothalamus is directly exchanging signals with the peripheral blood circulation and is therefore the main nutrient sensing site in the brain

In contrast, the hypothalamus is directly exchanging signals with the peripheral blood circulation and is therefore the main nutrient sensing site in the brain. Within the hypothalamus many nutrient related signals from both the environment as well as the internal milieu converge and become integrated [160]. In the 1940s, it was shown in rodents that specific nuclei within the hypothalamus alter food intake after electrical stimulation or inhibition by lesions, thus demonstrating the special role of the hypothalamus in the regulation of energy balance [22, 160, 251]. Nowadays, many discrete hypothalamic neuronal populations have been identified that express specific neurotransmitters to mediate particular effects on energy balance in relation to the nutritional state [251].

The importance of the hypothalamus in context of energy homeostasis was also shown for humans in the most recent genome-wide association study that indicated that the majority of genes associated with BMI are expressed in the central nervous system (CNS) and many of them in the hypothalamus [22].

One of the distinct neuronal circuits within the hypothalamus is the arcuate nucleus (ARC), whose collection of neuronal cell bodies occupies almost half of the length of the hypothalamus (Figure 1) [251].

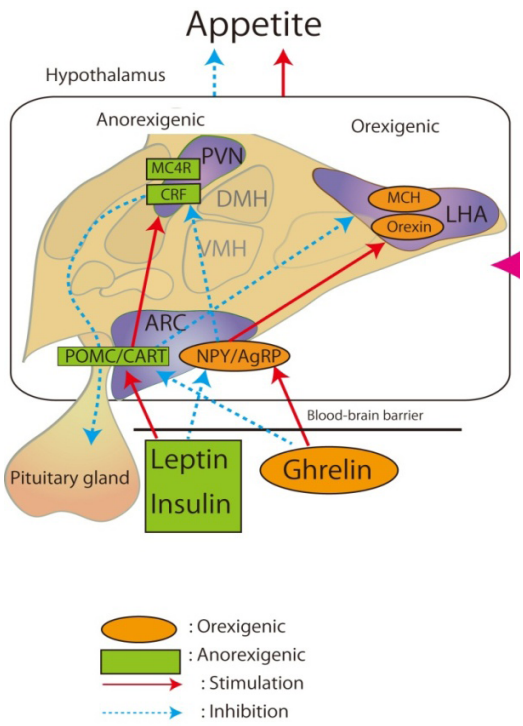


Figure 1: Feeding regulation of neuroendocrine hormones in the hypothalamus.
NPY: neuropeptide Y; POMC: proopiomelanocortin; AgRP: agouti gene-related protein; CART: cocaine- and amphetamine-regulated transcript; ARC: arcuate nucleus; PVN: paraventricular nucleus; VMH: ventromedial hypothalamic nucleus; DMH: dorsomedial hypothalamic nucleus; MCH: melanin-concentrating hormone; LHA: lateral hypothalamus; CRF: corticotrophin-releasing factor; MC4R: melanocortin 4 receptor.
Taken from Yagi et al. [254].

The ARC is situated around the third ventricle and lies above the median eminence [251]. The capillaries in the underlying median eminence are lacking tight junctions, thus the blood-brain barrier is here allowing the exchange with the circulation. As a result, the ARC neurons are readily accessible to circulation messenger such as the orexigenic hormone ghrelin and the anorexigenic hormones insulin and leptin [251]. The ARC has further projections to other hypothalamic regions, including the paraventricular nucleus (PVN), dorsomedial hypothalamic nucleus (DMH), ventromedial hypothalamic nucleus (VMH) and lateral hypothalamus [160, 251].

Within the ARC reside two neuronal populations that are well characterised for their role in the regulation of food intake: one set of neurons that is expressing neuropeptide Y (NPY) and agouti gene-related protein (AgRP) stimulates food intake whereas the second set of neurons expresses proopiomelanocortin (POMC) and cocaine- and

amphetamine-regulated transcript (CART) and inhibits food intake [90, 251].

1.3.3 *Leptin and leptin resistance*

Leptin is an adipocyte derived hormone produced in direct proportion to fat mass

Leptin is a 16 kDa hormone that is produced primarily in white adipocytes in direct proportion to fat mass and indicates the status of long-term energy stores of the body [18]. It targets hypothalamic leptin receptors (LepR), such as the long isoform LRb, and downstream JAK2/STAT3, MAPK and PI3K signaling to regulate food intake and energy expenditure (Figure 2) [169].

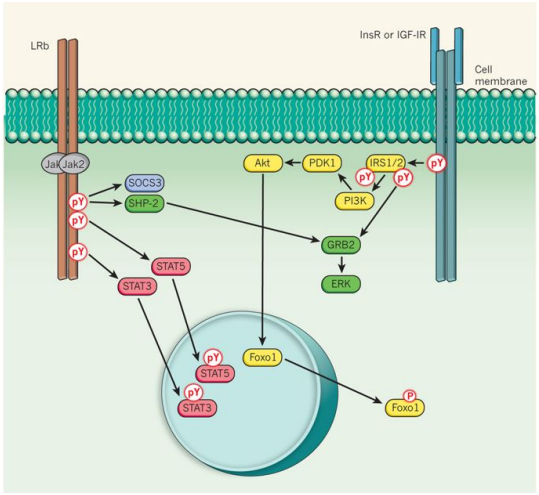


Figure 2: Cellular signalling pathways modulated by leptin. Leptin binds to the leptin receptor (LRb) which in turn activates the associated Jak2 tyrosine kinase. This leads to the phosphorylation of intracellular tyrosine residues on LRb, which leads to an activation of the Extracellular Signal-regulated Kinase (ERK) signaling pathway, a modulation of gene transcription in the nucleus, and the promotion of PI3K pathway signalling. Leptin and insulin signaling converge at multiple steps of their signaling cascade.
Taken from Myers Jr and Olson [169].

The absence of leptin itself or of LRb, the leptin receptor predominantly expressed in the CNS, causes massive hyperphagia and decreased energy expenditure leading to severe obesity [169]. Defects in leptin signaling also result in early-onset insulin resistance, hyperglycaemia and metabolic dysfunctions that are more severe than expected with obesity alone [169]. It was shown that whole-body or CNS-restricted leptin treatment of leptin-deficient ob/ob mice resulted in a rapid restoration of glycemic control independent of changes in food intake or adiposity suggesting that leptin controls blood glucose levels irrespective of energy balance [169].

In obese animals and humans the increased fat mass correlates with elevated circulating leptin levels. Because leptin reduces food intake and body weight, the absence of anorectic and metabolic effects in obesity despite hyperleptinemia is widely interpreted as leptin resistance [18]. There are several hypotheses on the molecular underpinnings of leptin resistance such as impaired LepR trafficking [49, 237], impaired signaling linked to Socs3 expression [25], impaired histone deacetylase 5 activity in the hypothalamus [118], or increased levels of circulating c-reactive protein [38]. But the riddle is not solved yet and further investigations are needed since the solution may offer new targets for the treatment of obesity and related diseases [18].

Leptin resistance is a common trend in obesity and negatively impacts metabolic health

1.3.4 *Exendin-4*

Exendin-4 (Ex4) is a hormone present in the venom of the American poisonous lizard, the Gila monster (Heloderma suspectum) (Figure 3) [74, 167].

Figure 3: Gila monster – *Heloderma suspectum* `http://pixdaus.com/arizona-gila-monster-stalks-its-prey-lizards-reptiles/items/view/275321/`

Exendin-4 mimics the incretin hormone glucagon-like peptide 1 (GLP-1) that is reducing food intake by promoting satiety, reducing fat deposition, and reducing body weight [74, 167]. GLP-1 is secreted from the intestinal K-cells and L-cells respectively and is derived from proglucagon [74]. Exendin-4 is more stable to degradation by peptidases than GLP-1, which has a very short half-live [74]. Like GLP-1, exendin-4 stimulates insulin synthesis, inhibits glucagon secretion, promotes b-cell proliferation, and protects against b-cell apoptosis in response to different insults [74]. It also inhibits gastric emptying which might lead to vomiting and nausea [74]. Müller et al. [167] demonstrated that pharmacological induced weight loss by exendin-4 treatment was able to restore leptin sensitivity in diet-induced obese mice.

Ex4 mimics the incretin hormone GLP-1 in reducing food intake

1.3.5 *The liver as an essential metabolic organ*

The liver is the central organ for various physiological processes governing body energy metabolism and it connects substrate utilisation of several tissues, such as skeletal muscle and adipose tissue [198, 232].

1.3.5.1 *Metabolic processes and anatomy of the liver*

The liver is a hub for various physiological processes governing body energy metabolism

After the digestion of food, nutrients in the form of glucose, fatty acids and amino acids are absorbed into the bloodstream and are transported to the liver and peripheral tissues [103, 198]. These macronutrients are then taken up by the liver and processed thereby generating energy in the form of adenosine triphosphate (ATP) and substrates necessary for the essential metabolic processes performed by this organ such as homeostasis of lipid and cholesterol metabolism, support of the immune system, breakdown of xenobiotic compounds, regulation of blood volume, and endocrine control of growth signaling pathways [232]. The liver can store glucose in the form of glycogen or metabolise it into fatty acids or amino acids, which then provide metabolic fuels not only for the liver itself but also for muscle, adipose tissue, and other extrahepatic tissues during fasting [198]. The liver is further the main organ for protein and amino acid metabolism, as it breaks down amino acids for energy, but it also synthesises the majority of proteins secreted into the circulation and orchestrates the disposal of nitrogenous waste from protein degradation [232].

To perform these demanding metabolic processes, the liver tissue consists of numerous cellular populations. Hepatocytes are the primary epithelial cell type and perform many of the described processes, such as the esterification of free fatty acids to generate triacylglycerol (TAG), which is then stored in lipid droplets in the hepatocytes or is secreted as very low-density lipoprotein (VLDL) particles [198, 232].

Hepatic energy metabolism is mainly regulated by transcription factors and nuclear proteins, whose activity is again controlled by metabolic hormones, such as insulin and glucagon [198]. A dysregulation of hepatic metabolism during nutrient overload contributes to the development of nonalcoholic fatty liver disease (NAFLD) often observed in obese patients [198].

1.3.5.2 *Nonalcoholic fatty liver disease*

NAFLD is referred to as the hepatic manifestation of the metabolic syndrome and is the most common liver disease worldwide in adults

as well as in children and ranges in pathologies from steatosis to nonalcoholic steatohepatitis (NASH), which could progress into liver fibrosis or into liver cirrhosis and in some cases into hepatocellular carcinoma [128, 189, 229]. 20 – 30% of the general population of developed countries are estimated to suffer from steatosis, this number increases to 75 – 100% in obese individuals [128]. Up to date the best therapeutic option to treat steatosis are lifestyle modifications including calorie restriction with or without exercise to reduce body weight [139, 189]. In this context, exendin-4 treatment was demonstrated to have beneficial effects on reducing hepatic inflammation and accumulation of lipids in the liver and thus improving insulin sensitivity [60, 140, 246, 256]. A better understanding of the cellular and molecular mechanisms involved in the development of hepatic steatosis and NAFLD may help to identify novel therapies for these obesity-associated diseases [189].

75 – 100% of obese individuals suffer from NAFLD

1.4 ANALYSING THE TRANSCRIPTOME

The transcriptome is the set of all RNA molecules in one cell or a population of cells. It thus gives a comprehensive view on all active transcripts and their quantity in a given sample at a specific time-point for a specific physiological condition [247]. The information captured in the transcriptome is quite extensive and focuses on different spectra of transcription-involved processes. It enables us to analyse the exon-intron structure, splice variants, and regulatory sequences [164]. Furthermore, noncoding RNA such as micro-RNAs and small interfering RNAs (siRNAs), which are post-transcriptional regulators of gene expression can be discovered and detected by analysing the transcriptome [164]. Additionally, the discovery of transcript rearrangements and profiling of single nucleotide variations are possible applying transcriptomics [164]. One of the main aims of transcriptomics is to capture transcript expression changes under different conditions [247]. Furthermore, transcript expression profiles promise a deep biological insight on molecular processes as they reflect intrinsic and environmental influences [135]. As the focus of this thesis concerning transcriptome analysis was on the analysis of mRNA expression profiles we will concentrate on transcript quantification and the corresponding analysis in the following.

Transcript expression profiles promise a deep biological insight on molecular processes

1.4.1 *Transcript quantification*

Several approaches have been developed to quantify the transcriptome; the most commonly used ones are hybridisation- or sequence-based approaches [247]. In the following we will present two of the most popular technologies to quantify transcripts, which are the hybridisation-based RNA-microarrays and the sequence-based RNA-Sequencing, also referred to as RNA-Seq, that have been used to generate the datasets of this thesis (Figure 4).

1.4.1.1 *RNA-microarrays*

In the 1990s, microarrays were developed as the major technology to quantify transcripts in a sample by analysing the RNA converted into complementary DNA (cDNA) [144]. To detect gene specific transcripts, fluorescently labelled cDNA is incubated with custom-made microarrays or commercial high-density oligo microarrays, which results in the hybridisation between complementary polynucleotides [144, 247]. After the hybridisation of the probes with the target cDNA, the fluorescent signals are scanned and reported [144].

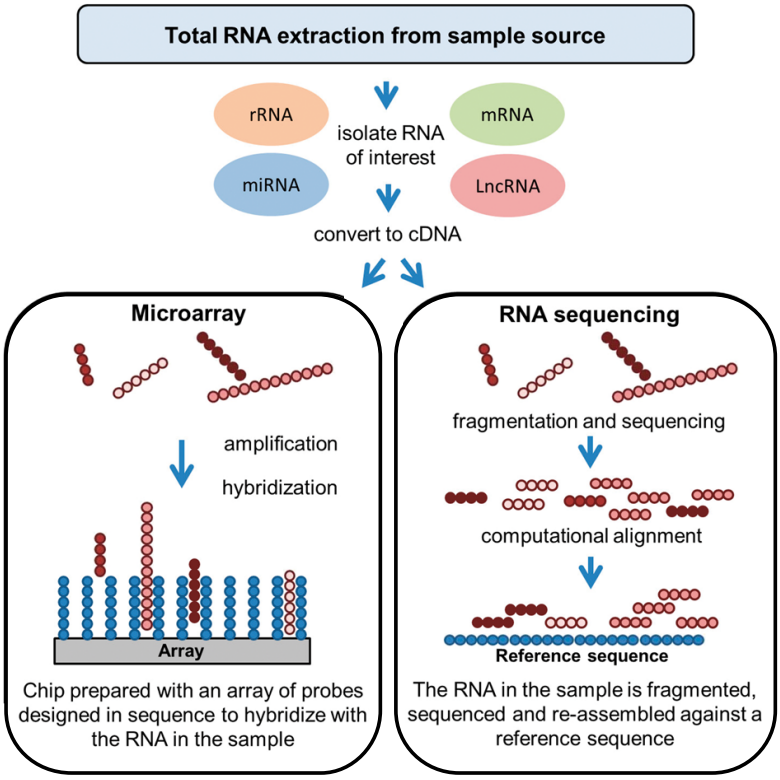


Figure 4: Overview of RNA-microarray and RNA-Sequencing procedures. Addapted from Manzoni et al. [154].

1.4.1.2 *RNA-Sequencing (RNA-Seq)*

In contrast to RNA-microarrays, RNA-Seq is a high-throughput DNA sequencing method for mapping and quantifying transcripts directly to the genome without hybridisation of the cDNA sequence to gene-specific probes, thus allowing unbiased analysis of the whole transcriptome [247]. Since the price for RNA-Seq has dropped in the recent years it has become an attractive alternative to microarrays [144].

Briefly, a population of RNA is converted to a library of cDNA fragments and adaptors are attached to one or both ends of these fragments [247]. Following an optional amplification step, each molecule is sequenced to retrieve short sequences, also referred to as reads, with a length of 30 – 400 base pairs from one or both ends (single-end or paired-end sequencing) [247]. After the sequencing step, reads are aligned to a reference genome or transcriptome or are assembled de novo without a reference sequence [247]. Transcriptomics by RNA-Seq thus results in a genome-scale transcription map consisting of the transcriptional structure and the expression level of each gene [247].

1.4.1.3 *Comparison of usage of RNA-microarrays to RNA-Seq*

RNA-Seq overcomes several limitations observed with the microarray technology

In the comparison of the newer RNA-Seq technology to the established RNA-microarray technology, it becomes obvious that RNA-Seq overcomes several limitations observed with the hybridisation-based microarray technology. First of all, RNA-Seq does not depend on a priori knowledge of the target cDNA sequences, thus it is especially attractive for non-model organisms for which no reference sequences exist [247]. Furthermore, unlike microarrays, RNA-Seq captures the precise location of transcription boundaries (exons), sequence variations, such as single nucleotide polymorphisms (SNPs) and alternative splicing isoforms [144, 247]. Moreover, RNA-Seq has a very low background signal compared to microarrays as it has no hybridisation and cross-hybridisation artefacts and the sequenced fragments can mostly be mapped unambiguously to unique genome regions [166, 247]. Additionally, RNA-Seq has a large dynamic range of expression levels since it does not have an upper limit for quantification whereas microarrays present a lack of sensitivity for genes expressed either at low or very high levels [247]. Last but not least, RNA-Seq has been shown to quantify gene expression levels highly accurately and reproducibly for technical as well as biological replicates [247], which makes the comparison across different experiments easier as no complicated normalisation methods are required. On the other hand, using microarrays is a lot easier and cheaper if the aim of a study is the quantification of known transcripts.

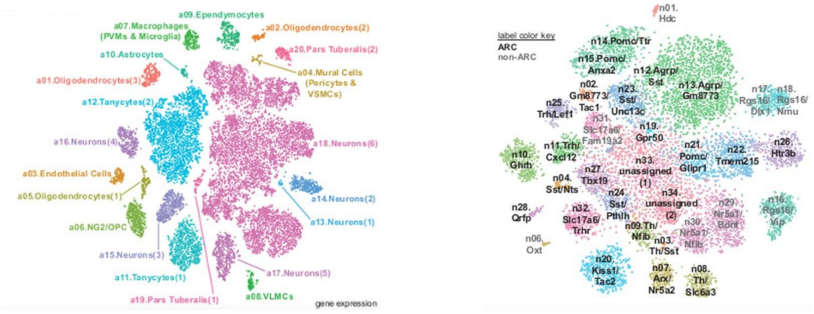
1.4.2 *Interpretation of transcriptome expression levels*

After quantification and identification of individual transcripts, gene expression levels are usually compared between treatment groups or time-points resulting in differential gene expression profiles, which are then linked to the biological context through enrichment and pathway analyses [34]. These enrichment and pathway analyses that allow the formulation of working hypotheses, are performed by mapping significant genes to known biochemical pathways stored in publicly available databases such as the Gene Ontology (GO) [14, 48] or the Kyoto Encyclopedia of Genes and Genomes (KEGG) [121–123]. This mapping approach is limited to the already known biochemical pathways. A second limitation is that the acquired knowledge is stored in numerous databases not necessarily using the same identifier, thus automatic analysis of multiple data sources is impeded. These limitations do not only apply for transcriptome analysis but also to the analysis of proteome and metabolome datasets.

1.4.3 *Phosphorylated ribosome capture technique (PRC)*

Recent studies demonstrate a shift in usage of whole tissue to single cell sequencing

Recent studies demonstrate a shift in the usage of whole tissue to single cell sequencing. With this technique it is not only possible to investigate several cell populations [35] but also to investigate cell subpopulations such as neuronal subpopulations, and how they respond to various stimuli (Figure 5) [35, 101, 110, 138, 257].



(a) Hypothalamic cell populations (b) Neuron populations

Figure 5: Single cell sequencing of hypothalamic cell populations. **(a)** Spectral tSNE plot of all known cell populations in the hypothalamus. **(b)** Spectral tSNE plot of neuron cell populations sequenced from the hypothalamus.
Taken from Campbell et al. [35]

Before single cell sequencing became the state of the art, gathering information on gene expression patterns of discrete cell populations, especially neuronal subpopulations, was more complicated. In 2012, Knight et al. [132] presented an approach to measure gene expression in activated neuronal populations in the brain. The main difficulty here was that neurons share marker genes and are thus difficult to be distinguished [132]. Knight et al. [132] exploited the observation that many stimuli that trigger signaling pathways correlated with neural activity also induce phosphorylation of ribosomal protein S6, which is phosphorylated downstream of PI3-K/mTOR, MAPK, and Protein kinase A (PKA) signaling (Figure 6). Phosphorylated S6 was used as a tag to capture mRNA from activated cells, by immunoprecipitating the phosphorylated ribosomes and their associated mRNA from mouse brain homogenates using antibodies recognising the specific phosphorylation site 244 of S6 [132].

RpS6 was used as a tag to capture mRNA from activated neurons

This technique is highly valuable to understand the biological underpinnings of obesity since it is designed to capture the transcript expression profiles of neuronal subpopulations responding to a stimulus. As described in Section 1.3.2, several neuronal populations play an important role in food intake and energy expenditure. But there are plenty other mRNA quantification based approaches that were and are used in obesity research.

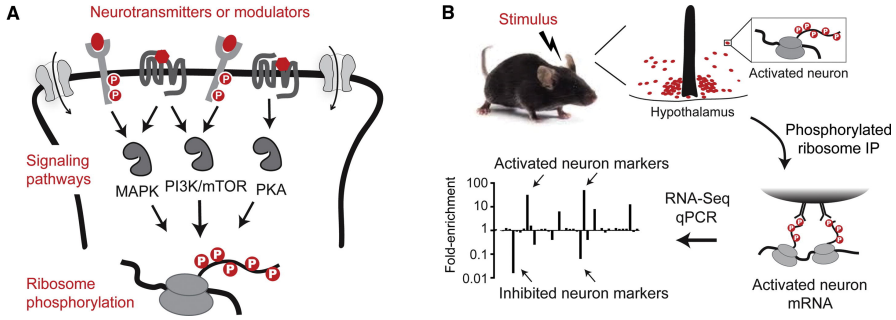


Figure 6: Phosphorylated Ribosome Profiling. **(A)** Neurotransmitters and neuromodulators activate a set of signaling pathways in neurons, which results in the phosphorylation of RpS6, which is a common target of these pathways. **(B)** A stimulus activates a neuron subpopulation (red cells), which shows an increased pS6 concentration. Capturing phosphorylated ribosomes allows the quantification of mRNA expressed in the activated cells.
Taken from Knight et al. [132].

1.4.4 *RNA-Seq in obesity research*

Obesity results from a combination of genetic and environmental factors

Microarray and RNA-Seq technologies, as the most prominently used approaches, have proven to be highly valuable in all areas of biological research including the investigation of the biological underpinnings of obesity, its comorbidities, and treatment opportunities. This is, as any intervention, such as fasting, nutritional deficiency or excessive intake of a nutrient, causes extensive changes in many metabolic parameters [13]. Since gene expression is one of the main determinants of observed metabolic profiles, capturing and analysing changes in gene expression regulation is a promising approach to understand the complex biological underpinnings behind multifactorial diseases such as obesity, which results from a combination of genetic and environmental factors [158].

The broad use of transcriptomics and the public availability of these datasets enable the performance of meta-analysis by combining existing datasets to gain knowledge without having to perform all experiments over and over again. Swindell et al. [226] used such an approach to study the effects of calorie restriction in ten different tissues. They show across multiple datasets that the beneficial effects of calorie restriction on ageing and obesity are at least partly mediated by a set of genes involved in tumour suppression and reduction of oxidative stress across several tissues [226]. In another meta-analysis, Wu et al. [252] examined the gene expression profiles of hypothalami from mice of different ages fed ad libitum or subjected to calorie restriction since weaning. They were able to show that calorie restriction influenced the expression of genes associated with the mammalian target of the rapamycin (mTOR) nutrient sensing pathway, which is

involved in the regulation of energy intake and ageing [252]. Calorie restriction was further shown to alter the expression of hypothalamic genes involved in hunger signaling and circadian rhythms using RNA-Seq and subsequent pathway enrichment analysis [58].

1.5 METABOLOMICS

Metabolites are endpoints of metabolic processes in the cell

Metabolites are small molecules with a size of 50 – 2000 Da that can either be intermediate or end products of metabolism. Metabolites have various functions in the cell, including fuel, structure, signaling, stimulatory and inhibitory effect on enzymes, and are thus an important read out of cells and tissues since they indicate the physiological state of the cell and can be used as biomarkers to predict a phenotype or disease state [11]. Additionally, metabolites are frequently exchanged with the environment for example by inhalation, food intake, and excretion [11]. Furthermore, the metabolome is highly dynamic and time-dependent, as metabolites are sensitive to environmental conditions and changes [11].

It is not possible to capture and analyse all metabolites by only one method due to their high chemical diversity

Metabolomics, the analysis of chemical processes involving metabolites, is a valuable method to study the cellular metabolism, since, as mentioned before, metabolites can be used as biomarkers to predict a phenotype. However, due to the high chemical diversity of the metabolites, measuring the metabolome is not that straightforward and it is not possible to capture and analyse all metabolites by only one method [11, 172]. Therefore, several methods including gas chromatography (GC), liquid chromatography (LC), or nuclear magnetic resonance (NMR) spectroscopy were developed to analyse specific metabolite classes, but there is a lack of consistency in terms of metabolite recovery between these methods [11, 172]. Depending on the research question and prior knowledge, the experimental approach can be targeted to detect a priori selected metabolites by their known masses or non-targeted, where all possible masses are detected and quantified [11]. However, the interpretation especially of non-targeted metabolomics is difficult since of right now only a limited amount of metabolites can be identified and annotated robustly due to technological limitations of the used platforms [34]. Furthermore, the identification of the specific pathway or enzymatic reaction connected to the altered metabolite is difficult as many metabolites are involved in numerous pathways and are the substrates or products of several enzymes or reactions [11]. Another complication in the metabolome interpretation is the tissue-specificity of the metabolites and the limited data availability for only a few specific species and tissues [11]. Other pitfalls are the very low concentrations of intracellular metabolites, their very high turnover rate, their very high susceptibility to chemical modifications, and that many metabolites

do not exist in a free state within a cell [172].

Despite these challenges, metabolomics especially when integrated with upstream cellular processes analysed by genomics, transcriptomics, and proteomics empowers us to understand metabolic pathways and disease mechanisms in more detail. The benefit of an integrative analysis compared to an analysis focusing on individual metabolites or genes, is the conservation of the causative context.

1.6 DATA INTEGRATION OF METABOLOMICS AND OTHER OMICS

Multi-omics integration allows to capture potential causative changes that lead to a complex disease

Each omics type presents another aspect of a phenotype state or disease and its analysis can capture differences between a disease and a control group. Nonetheless, this analysis is restricted to correlations, that mainly reflect reactive processes [97]. The integration of different omics types offers the opportunity to capture potential causative changes that lead to a complex disease, such as diabetes or cancer, as the regulation of a biologic system occurs across multiple omics levels [81, 97].

For the integration of omics datasets, we usually study the more familiar interactions between genes, transcripts, and their translation into proteins following the central dogma of molecular biology (Figure 7) [33].

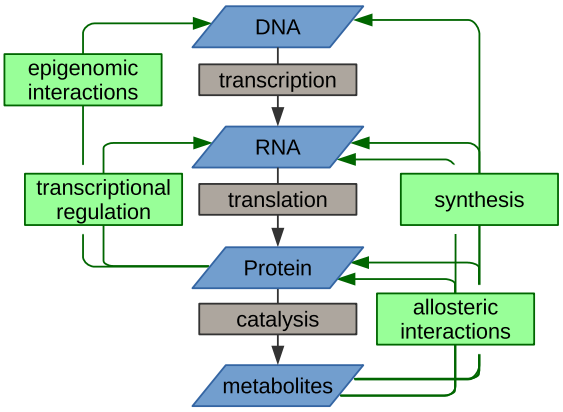


Figure 7: Central dogma of molecular biology. DNA gets transcribed into RNA, which gets translated into proteins, that are then catalysed into metabolites. This central dogma gets extended by relationships (shown in green) that one should keep in mind when integrating omics data.
Taken from Buescher and Driggers [33]

However, for the integration of metabolomics data with genomics, transcriptomics or proteomics, the interactions are less linear and the integration step is rather an disentangling of multiple cycles of functional relationships, as metabolites can interfere with several processes, they can be substrates and products of enzymes, they can allosterically regulate the activity of enzymes and transcription factors, and they are the monomers from which proteins and RNA are synthesised (Figure 7) [33].

For the integration of omics datasets innovative approaches have been invented combining knowledge of multiple disciplines

For the complex integration of omics datasets innovative statistical and machine learning approaches have been invented that combine the knowledge of multiple disciplines such as biology, medicine, statistics, machine learning, and artificial intelligence [261]. Current approaches used to integrate metabolomics with transcriptomics or proteomics datasets can be divided into two groups: analysis tools based on pathway and/or network annotations that depend on prior knowledge versus tools relying more on the numerical aspect of the dataset that is independent from any prior knowledge.

To use existing biological knowledge for the integration of metabolomics and transcriptomics, metabolites and transcripts or their correlations can be mapped to the genome-scale metabolic models of the organism of interest [221, 266]. Su et al. [221] calculated pairs of metabolites and transcripts involved in the same pathways or reactions by correlation analysis of gene expression profiles and metabolite measurements over cell-lines of the NCI60 dataset and mapping known associations from the Edinburgh human metabolic network database to their NCI60 associations. Çakir et al. [266] identified reporter reactions, where metabolites significantly react to perturbations, by mapping metabolites and transcripts to a reduced genome-scale metabolic model, where unmeasured metabolites are eliminated from a published metabolic network.

There are several tools that approach integration through the use of annotated pathway knowledge, such as Integrated Molecular Pathway Level Analysis (IMPaLA [120]) and INtegrative Meta-analysis of Expression data (INMEX [253]) [37]. INMEX is a tool for joint pathway analysis, which can predict if several genes or metabolites are more crucial due to their position in the network and assigns an appropriate weight to these compounds [37, 253]. In general, significantly deregulated genes and metabolites are determined in each set separately and are, in a second step, mapped to pathways of several data collections for over-representation, enrichment, or pathway topology analysis to identify pathways that are differentially regulated on either both levels or only the metabolic or transcriptomic one [37, 120, 253].

The two main methods based more on the numerical aspect of omics data integration are correlation analysis and/or unsupervised multivariate techniques, such as principal component analysis (PCA), partial least squares (PLS) regression, and O2-PLS, which is an extension of the standard PLS model [37, 114, 135]. PLS regression was, for example, used by Griffin et al. [84] to study fatty liver disease by integrating time-resolved metabolomics and transcriptomics data obtained from rat liver tissue [135]. In general, correlation analysis is

simple to implement and used quite frequently to integrate metabolite and transcript datasets [5, 79, 104, 117, 124, 234, 248].

The correlation analysis approach is based on the fact that common regulatory events or substrate-product relationships of structurally similar metabolites derived from the same or related pathways, often show correlations [203]. This correlation between metabolites or between metabolites and genes might not be present at all time-points. In fact, some gene/metabolite correlations only become obvious under particular conditions, such as during a challenge of the system, or at particular developmental stages [202]. Similar to a gene co-regulation analysis, correlation analysis determines significant associations between metabolites and genes, followed by clustering and network visualisation or gene enrichment analysis of functional groups [114]. For instance, Bartel et al. [19] constructed a correlation network that showed which transcripts and metabolites had the same trend of up- and down-regulation in the blood of individuals from a population-based cohort.

Correlation between metabolites or between metabolites and genes might not be present at all time-points

However, the correlation approach might not be the best-suited method to integrate metabolic-transcriptomic data, as metabolites are not directly related to gene transcripts [37]. It has been observed that metabolites known to be closely related in pathways, often do not show a correlative link, while correlations occur connecting compounds with great distances across the network [37]. One approach to exclude these coincidentally identified relationships could be the additional application of partial correlation [37, 73]. Additionally, partial correlation could also solve the problem, that correlation analysis cannot detect dependencies between variables, by incorporating the effect of other variables [124].

One approach to exclude coincidentally identified relationships could be the additional application of partial correlation

One of the best-known and most distributed methods for undirected network inference with partial correlation is called Gaussian graphical model (GGM) [204]. Zuo et al. [265] recently proposed to not only calculate the partial correlation coefficients between each pair of nodes conditional on all other p - 2 nodes (with p being the number of nodes in the network) as it is done in GGM, but to extent the calculation to additionally calculate the zero-th order, which refers to simple correlation, and first order partial correlation coefficients (conditioned on one variable). They then calculate the second order partial correlation (p - 2) only for metabolite-transcript pairs that show zero-th order and first order partial correlations significantly different from 0. This approach is supposed to remove most false positive edges from the inferred networks. Additionally, they claim that it also solves the problem of having biological experiments with a large number of variable measurements but a relatively small

sample size, which usually results in statistical challenges [129, 265].

There exists a heterogeneous jungle of integrative analysis techniques that are often not user-friendly

Next to the data integration approach, there are several other limitations impeding the further advancement of the multi-omics research field. Currently, there exists a heterogeneous jungle of integrative analysis techniques that are often not user-friendly and/or return results that are not easy to interpret. These limitations in data generation and interpretation further complicate the data comparison between different labs that is already difficult due to the limited reproducibility of results obtained with transcriptome, proteome and metabolome research, and missing standards for the data formats of omics data and for metadata [34]. The establishment of standard workflows for metabolomics for data normalisation, quality control, and data analysis, comparable to those implemented in microarray- and sequencing-derived datasets [135] would be very helpful. This standardisation of metabolome analysis could also help to facilitate the establishment of public databases storing metabolomics data, which could then simplify and accelerate future research.

1.7 SCOPE OF THIS THESIS

This PhD thesis aims to use known and to develop new bioinformatics tools to uncover biological underpinnings of conditions caused by nutritional dysregulation, such as caloric overload or consumption of hazardous substances as well as the influence of the respective rescue interventions. First, we will determine the predictive value of combined analysis of pro-mutagenic DNA adducts formation after consumption of carcinogenic food components together with associated early gene expression changes on tumourigenesis. Next, we will use a model of diet-induced obese mice to unravel the molecular underpinnings of pharmacologically induced weight loss versus calorie restriction on hypothalamic leptin sensitivity and hepatic steatosis by transcriptomics and metabolomics.

The PhD thesis is composed of three partially independent projects:

1.7.1 *Time-matched analysis of DNA adduct formation and early gene expression in methylazoxymethanol acetate treated Eker rats*

In the first part of the thesis, differential gene expression analysis and pathway enrichment analysis were applied to examine if transcriptional induction of genes involved in DNA repair or cancer-related pathways can evaluate the effectiveness of methylazoxymethanol acetate treatment on cellular protection. Furthermore, the risk of DNA adduct formation for tumourigenesis on the basis of time-matched transcriptome analysis was evaluated.

1.7.2 *Effects of weight loss interventions on hypothalamic gene expression*

In the second part of the thesis, differential gene expression analysis and pathway enrichment analysis were applied to examine the molecular underpinnings of restoration of hypothalamic leptin sensitivity by pharmacologically induced weight loss in contrast to weight loss induced by calorie restriction.

1.7.3 *Estimating genetic impact on metabolic networks*

In the last part of the thesis, a statistical method for correlation-based network integration, called CoNI, was developed. CoNI is a generic method and can be applied to multiple experimental settings. For the newly developed method, liver transcriptome and metabolome datasets of diet-induced obese mice undergoing weight loss interventions compared to respective controls were combined to infer genetic control of the metabolic networks. Here, we identified local regulator

genes playing a role in hepatic steatosis and in the metabolic adaptation in the liver during weight loss.

2 DATASETS AND METHODS

2.1 MICROARRAY DATASET OBTAINED FOR KIDNEYS OF CARCINOGEN EXPOSED RATS

In the first project of the thesis, gene expression data from kidneys of carcinogen exposed Eker rats were used, as previously published by Stemmer et al. [218]. The study focused on the genotoxic substance aristolochic acid (AAc) and the non-genotoxic substance ochratoxin A (OTA). Paralleled data from MAMAc exposed rats were collected but not published before.

In this study, groups of male $Tsc2^{+/-}$ Eker and wild type (WT) rats at an initial age of 11 to 14 weeks were treated with MAMAc (250 µg/kg body weight (BW)), AAc (10 mg/kg BW) or OTA (210 µg/kg BW) by oral gavage on a daily basis. Time-matched vehicle controls were gavaged with 0.1 M $NaHCO_3$. In case of the WT rats, only one control cohort was kept for all three substances, whereas two control groups were present for the Eker rats: One control cohort for AAc and MAMAc treatment and one for the OTA administration.

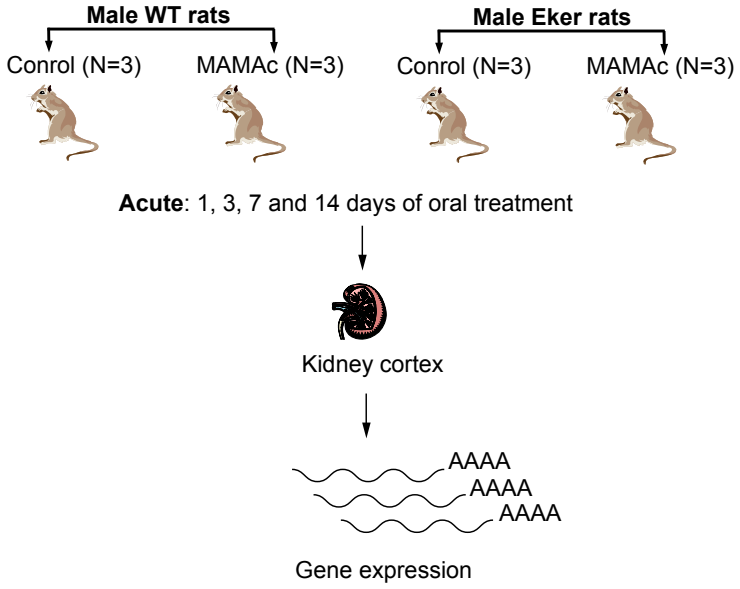


Figure 8: Regimen toxin treatment of rats.

Affymetrix microarray sequencing data were obtained from kidney cortices for four time-points: Day 1, day 3, day 7, and day 14 after the oral treatment with the respective toxin to capture early treatment

Gene expression profiles were obtained from kidney cortices for day 1, 3, 7, and 14 after treatment

effects on gene expression (Figure 8). This resulted in a dataset with three replicates for each sampled time-point, genotype, and treatment. The protocol starting from the RNA isolation and sequencing to the microarray data preprocessing resulted in the probe-ID abundances and was described in Stemmer et al. [218].

2.2 DATASETS OF THE WEIGHT LOSS INTERVENTION COHORT

All experiments of the second and third project of the thesis were performed in adult male C57BL/6J mice purchased from Janvier Labs (Saint-Berthevin, Cedex, France). Mice were maintained on a 12-h light-dark cycle with free access to water and standard chow diet (Altromin, #1314).

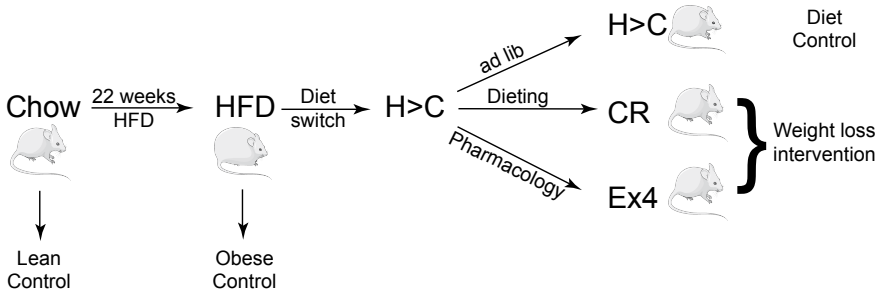


Figure 9: Regimen of weight loss intervention study. The administered diet is indicated on the arrows.

To generate the diet-induced obese (DIO) colony, six weeks old C57BL/6J males were subjected to ad-libitum feeding of 58% high fat diet (HFD) (Research Diets, D12331) for 22 weeks (Figure 9). For the diet intervention study, DIO mice were subdivided into four experimental groups. The group termed HFD was kept on HFD during the study and served as "obese" control. The remaining three groups of DIO mice were switched to chow on day 0 of the study and divided as follows: Diet-switch (H>C) animals received ad libitum access to chow diet. Calorie restricted (CR) mice were restricted to the average food intake of the exendin-4 (Ex4) group and Ex4 treated animals were subjected to daily injections of exendin-4 (0.08 mg/kg body weight) (Tocris biosciences, Bristol, UK) in the morning for ten days. Age-matched mice maintained on chow diet over the 22 weeks and during the diet intervention period were used as a "lean" control group.

Diet-switch from HFD to chow led to a BW decrease of 10%, weight loss intervention to a decrease of 30%

At the end of the weight loss intervention study, mice in the diet-switch control group H>C had lost 10% of their body weight but there was no difference in body weight compared to the obese HFD control. The H>C group was further significantly heavier than the two weight loss groups CR and Ex4 (Figure 10). Dietary and pharmacologically induced weight loss reduced body weight by 30% in the CR and Ex4 group. There was no difference in body weight or body weight loss between the two weight loss groups. CR and Ex4 mice were still significantly heavier than the Chow controls but significantly lighter than the HFD and H>C controls. Mice were randomly assigned to the saline control group or the leptin treatment group

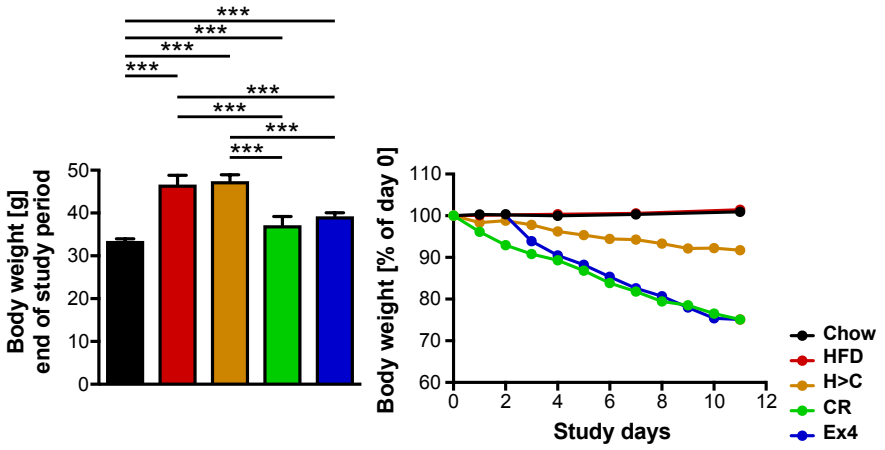


Figure 10: Body weight development of weight loss intervention cohort. The asterisks indicate the significance of the mean difference obtained by Welch's t-test. $***$ indicates a significance level $<$ 2.1e-12.

and received a single intraperitoneal (i.p.) injection of either saline or 3 mg leptin/kg body weight, 75 minutes before being sacrificed by cervical dislocation for organ withdrawal.

Ex4 mice are leptin sensitive while CR mice are still resistant

In an independent cohort of mice that underwent the same weight loss regime, leptin sensitivity was tested by measuring 24h food intake after a single i.p. injection of either saline or 3 mg leptin/kg body weight (Figure 11).

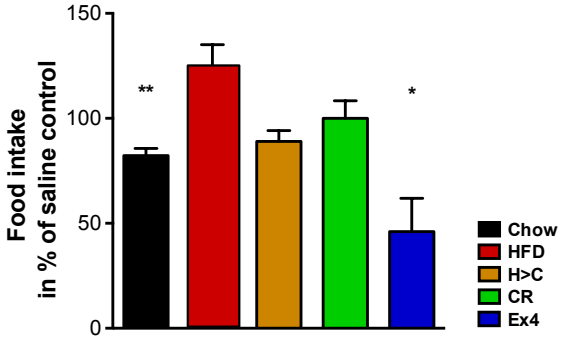


Figure 11: Food intake upon leptin injection of weight loss intervention cohort. The asterisks indicate the significance of the mean difference: $**$: p-value$<$0.002, $*$: p-value$<$0.02.

The lean Chow control and the Ex4 treated mice significantly reduced their 24h food intake after leptin injection relative to their saline injected control group. In contrast, despite similar body weight loss the CR mice did not respond to exogenous leptin by lowering their 24h food intake. The obese HFD group and the H>C control group, which is also still obese, were also leptin-resistant.

2.2.1 *PRC RNA-Seq datasets obtained for the weight loss intervention cohort*

The PRC approach allows the analysis of mRNA bound to the ribosomes exclusively of activated neurons by RNA-Seq. The PRC technique results in two datasets for each experiment: Input samples and their respective IP samples (Figure 12). The Input samples contain sequenced reads of the total mRNA of the hypothalamus. In contrast, the IP samples consist of the sequenced reads of the mRNA, which was isolated from immunoprecipitated ribosomes of activated neurons.

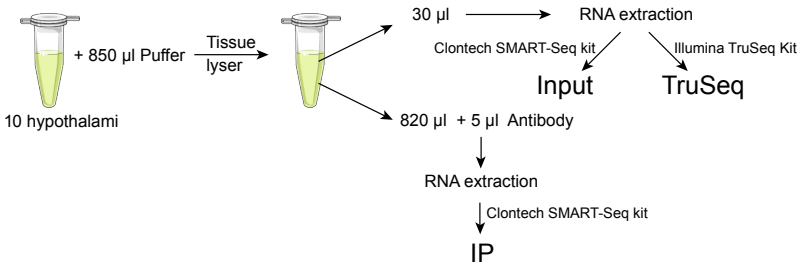


Figure 12: RNA extraction protocols to obtain the three RNA sequencing datasets. Input and IP samples with the PRC technique and Truseq as gold standard control.

The mRNA concentrations of the PRC IP samples (average mRNA concentrations = 125 pg/µl) were below the concentration range of the Illumina TruSeq kit [109], which is used routinely by the RNA-Seq core facility at HMGU research centre. Thus, the Clontech SMART-Seq protocol [43] for very low input samples was applied. In order to have comparable datasets, the respective Input samples (average mRNA concentrations = 70 ng/µl) were diluted to the ultra low RNA levels of the IP samples and were also analysed with the SMART-Seq kit. The sample sets generated with the ultra low input SMART-Seq kit will be referred to as "Input" and "IP" samples.

For the immunoprecipitation step the Phospho-S6 Ribosomal Protein (Ser240/244) antibody (#2215, Cell Signaling Technology (CST), Frankfurt am Main) was used in combination with a 3P peptide that blocks unspecific binding of the polyclonal antibody to other phosphorylation sites of the S6 ribosomal protein (RpS6).

The commercially available formulation of the Phospho-S6 Ribosomal Protein antibody has concentrations of roughly 45 µg/ml in a 50% glycerol solution. However, the IP reaction has to be carried out with 1 µg antibody per reaction and would thus require high volumes of the commercially available antibody. Therefore, a custom made version of the antibody produced by CST with a concentration

of 1.2 mg/ml in a PBS solution was applied. This higher concentrated formulation will be referred to as PBS antibody.

2.2.1.1 *Insulin PRC dataset*

As the PRC approach was a method that was set up for the first time in the lab, a smaller test set was generated before applying the method to the large weight loss intervention study. For this positive control set, adult C57BL/6J mice were injected i.p. with 1.5 U insulin/kg or saline as control 60 minutes before the sacrifice. Two replicates per injection were generated resulting in four Input and four IP samples, two for each treatment. For this test set the commercial Phospho-S6 Ribosomal Protein antibody was used.

2.2.1.2 *Leptin PRC dataset*

The weight loss intervention study was performed to investigate the effect of leptin treatment on hypothalamic mRNA expression as well as the discrepant action of dietary vs. pharmacological weight loss on feeding behaviour. For each of the n=5 diet groups (Chow, HFD, H>C, CR, Ex4) four replicates per leptin or saline injection with 10 hypothalami each were generated resulting in 20 Input and 20 IP samples each. The total RNA-Seq sample set therefore included 40 Input and 40 IP samples. Due to the large sample set, the PRC procedure for the replicates was performed on four consecutive days with each day having one replicate per diet group and treatment. For this large sample set the PBS formulation of the Phospho-S6 Ribosomal Protein antibody was used.

2.2.2 *Gold standard hypothalamic RNA-Seq dataset obtained for the weight loss intervention cohort*

In order to assess the effect of the low input SMART-Seq kit on read abundance and read quality, the standard sequencing Illumina TruSeq kit was additionally used to sequence the Leptin PRC dataset Input samples that reflect the total hypothalamic mRNA. Samples sequenced with this kit will be referred to as "TruSeq" samples. Thus, for the Input samples two datasets – one Clontech SMART-Seq dataset called "Input" and one Illumina TruSeq dataset called "TruSeq" – were generated whereas only one Clontech SMART-Seq dataset was generated for the IP samples (Figure 12).

2.3 LIVER DATASETS OF THE WEIGHT LOSS INTERVENTION COHORT

In the course of the project on hypothalamic leptin resistance and weight loss interventions, peripheral tissues from the differentially treated mice (described in Section 2.2) were extracted additionally to obtain a more holistic view on their health status. In the livers of exendin-4 treated mice elevated triacylglyceride (TAG) concentrations compared to all other groups were observed. This elevation had not been expected as chronic exendin-4 treatment was shown to improve hepatic steatosis before [60, 140, 246, 256]. Therefore, the metabolic underpinnings of the increased liver TAG levels after acute exendin-4 treatment were assessed by metabolomics and transcriptomics.

For the investigation of the metabolic processes in liver the metabolomics and the transcriptomics dataset were combined and a new approach to assess the genetic control of metabolite interactions was developed. The dataset of the five diet groups was divided into two separate sets. The first set was used to test the new integration method and contained the samples from the weight-stable Chow and HFD control groups. The second set contained the samples from the weight loss interventions: diet-switch (H>C) cohort, the calorie restricted (CR) mice, and the exendin-4 treated (Ex4) mice. We applied the newly developed method to this set to identify genes that might be causal for the observed metabolic changes.

2.3.1 *Transcriptome dataset*

Microarray sequencing data were obtained from liver samples of each diet group (n=10 per group) with the Affymetrix Mouse Clariom S Arrays (Affymetrix, Santa Clara, US). Two samples of the HFD-fed mice were excluded from the analysis as these mice had not gained body weight on HFD and their liver transcriptome expression pattern was comparable to control mice (data not shown). Three samples were removed from the Ex4 set and one from the H>C set as these mice were non-responders to the weight loss interventions.

On average, 22,206 transcripts were detected per sample. Transcripts with low expression levels defined as less than 40 counts in 30% of the samples were removed from the datasets. The removal of low expressed transcripts resulted in 10,159 genes in the Chow/HFD dataset and 10,254 genes in the H>C/CR/Ex4 set for the downstream analysis, respectively. Gene expression profiles were log-transformed to fit a normal distribution, which increases the accuracy of applied statistical tests.

The microarray data of the Chow/HFD set has been submitted to the GEO database at NCBI (GSE137923).

2.3.2 *Metabolome dataset*

Targeted quantitative metabolomics analysis of the livers of the weight loss intervention study was performed applying the Biocrates AbsoluteIDQ^{TM} p180 Kit (BIOCRATES Life Sciences AG, Innsbruck, Austria). Combined with mass spectrometry this kit allows quantification of up to 181 metabolites involved in metabolic diseases. The measured metabolites can be assigned to defined classes as shown in table 1.

Table 1: Distribution of measured endogenous metabolites.

METABOLITE CLASS	# MEASURED METABOLITES	# METABOLITES IN ANALYSIS
Hexoses	1	1
Biogenic Amines (BA)	14	12
Amino Acids (AA)	21	20
Acylcarnitines (AC)	40	40
Phosphatidylcholines (PC)	76	76
Lyso-Phosphatidylcholines (LPC)	14	14
Sphingomyelins (SM)	15	12

Metabolites with not-imputable concentration levels were excluded from the downstream analysis (for class distribution of metabolites in downstream analysis see table 1). Metabolite concentration profiles were also log-transformed to fit a normal distribution.

2.4 DATA ANALYSIS OF THE CARCINOGEN EXPOSITION DATASET

2.4.1 *Microarray data processing*

Microarray probe-IDs obtained for the Eker rats and their WT controls were annotated using the R bioconductor package rat2302.db [36, 78, 227]. The gene expression profiles were log-transformed to fit a normal distribution, which increases the accuracy of applied statistical tests. Principal component analysis (PCA), which is an unsupervised statistical procedure to explore the axes of variance in the expression profiles contained in the dataset, was performed on the complete dataset including the AAc and OTA expression profiles.

Since the findings on OTA and AAc treatment were already published, we concentrated on the analysis of the samples relevant for the MAMAc treatment in the following analyses.

2.4.2 *Differential gene expression analysis of microarray probe-IDs*

Limma was applied to find differentially expressed genes between MAMAc treated rats and their untreated controls

To gain more insight in the gene expression changes occurring during the first 14 days of MAMAc treatment, differential gene expression analysis was performed for the expression patterns measured in MAMAc treated rats and the corresponding control cohort. As the gene expression profiles were gathered with the microarray technology, the R bioconductor package Linear models for microarray data (limma) [179, 191] was used for the statistical analysis of gene expression differences. Limma fits a linear model to each gene in the data to estimate the observed variability to be able to distinguish it from the random variation. Then, it determines the expression differences by t-tests. To improve the power and achieve a more stable inference even if only a small number of arrays is available for the experiment, it moderates the standard error of the estimated log-fold changes by parametric empirical Bayes [210, 211]. Empirical Bayes methods estimate the prior distribution from the observed data, whereas in standard Bayes methods the prior distribution is fixed before data observation. In this case, the used empirical Bayes method shrinks the gene-wise residual variances towards a pooled estimate representing the global trend [210, 211].

After the testing step, the false discovery rate (FDR) was controlled by correcting the obtained p-values for multiple testing. The Benjamini-Hochberg (BH) procedure [21] was applied for the correction. Genes with adjusted p-value$<$0.05 were considered to differ significantly, allowing an FDR of 5%.

2.4.3 *Pathway enrichment analysis*

ClusterProfiler [258] identified the metabolic processes, which were affected by the differentially expressed genes. It performed an over-representation analysis of the genes of interest against the KEGG database [121–123] and GO [14, 48] biological process terms. To control the FDR, the Benjamini-Hochberg [21] method was applied and the adjusted p-value cutoff was set to 0.05, allowing an FDR of 5%.

2.5 ANALYSIS OF THE HYPOTHALAMIC RNA-SEQ DATASETS

2.5.1 *RNA-Seq read mapping and quality assessment*

The reads obtained for the hypothalamic RNA-Seq datasets of the weight loss intervention cohort were sequenced as 100 basepair paired-end runs on Illumina HiSeq 2500. The image analysis and base calling were performed with Illumina Real Time Analysis and the reads were demultiplexed with CASAVA 1.8. Reads were aligned with the GEM mapper (version 1.7.1) [155] and standard parameters (except mismatches=0.04 and min-decoded-strata=2) against the mm9 reference genome [47].

The quality of the sequenced, sorted and aligned reads was then assessed using FastQC (version v0.11.2) [10]. The number or mapped reads were obtained using Samtools (version 1.1) [143] and the RNAseqmetrics tool from picard tools (version 1.128) [111] was used to assess the genomic areas, where these reads map to. To obtain the positions of reads mapping to introns, all intronic regions from the UCSC mm9 annotation BED file provided on the RSeQC website [186, 244] were extracted. With RSeQC (version 2.6.1) the RNA-seq read coverages over the gene body were then calculated. The tool takes 100 quantiles from each feature provided in the BED file, in our case these features are introns, and extracts the coverage signals from the provided BAM files. The signals are then normalised and visualised.

2.5.2 *Antibody Evaluation*

Four different antibody combinations were used to produce the IP samples

To assess differences in binding affinities of the two antibodies used for the generation of the IP samples, label-free protein quantification within the immunoprecipitated samples using mass spectrometry was performed. As described above, four different settings were used to produce the IP samples: Commercial Phospho-S6 Ribosomal Protein (Ser240/244) antibody (ComAB), ComAB combined with 3P peptide (ComAB3P), which was applied to produce the IP:Insulin samples (Section 2.2.1.1); Phospho-S6 Ribosomal Protein (Ser240/244) antibody with the higher concentrated PBS formulation (PBSAB), which was used to obtain the IP:Leptin samples (Section 2.2.1.2), and the PBSAB combined with 3P peptide (PBSAB3P). For each set of samples an Immunoglobulin G (IgG) negative control sample was run to confirm that the primary antibody binds to the correct epitope on the expected antigen under the same experimental conditions.

The measurement of the protein masses, their normalisation, and annotation was performed by the Research Unit for Protein Science of the Helmholtz Centre Munich. Mascot [178] was used to identify

proteins from the mass spectrometry peptide m/z values with a minimal confidence score of 14 to report a protein.

For the analysis all proteins that were not measured either in the IgG control or in all four antibody samples were removed. Further all proteins, where only one sample showed a spectral count larger than 0, were removed since that hints to a false positive detection. Additionally, proteins for which only one peptide was measured were excluded as one peptide can be present in more than one protein and the assignment is therefore not reliable.

For the comparison of protein concentrations between different samples, the abundance ratios (AR) from the normalised abundance (NormA) of the protein measured within the respective antibody sample compared to the normalised abundance of the protein measured within the IgG control were calculated:

$$AR = \frac{NormA_{AB}}{NormA_{IgG}} \tag{1}$$

To identify the proteins bound by the antibodies, the amount of protein measured within each sample was assessed for several functional protein classes, f. ex. ribosomal proteins. Then the AR of each protein i=1,..,n in the respective antibody sample was normalised to the protein length of the respective protein L and all proteins of the respective class c were summarised:

$$CA_c = \sum_{i=1}^{n} \frac{AR_i}{L_i} \tag{2}$$

For the pie chart presentation of the identified fractions, the percentage of class abundance CA_c per total abundance of all proteins measured in the sample was calculated. Protein lengths and GO [14, 48] functional annotation were obtained from Ensemble (release 67) [72] using the R bioconductor package biomaRt [31, 64].

2.5.3 *Gene expression analysis of the hypothalamic RNA-Seq data*

The following analyses were limited to the TruSeq RNA-Seq data of the Leptin PRC dataset since the antibody used for the IP samples to capture the mRNA attached to the ribosomes in the PRC approach bound nonspecifically and thus dismantled the Clontech SMART-Seq datasets.

2.5.3.1 *RNA-Seq data preprocessing*

Htseq-count [9] in union mode was applied to obtain read counts from the provided read alignment files. This parameter setting is rec-

ommended in the htseq-count manual [8]. The gene expression of the 40 TruSeq samples was compared using the R bioconductor package DESeq2 [148]. DESeq was specifically designed for differential gene expression analysis. For the comparison between samples, the read counts were normalised by library size and the regularised-logarithm transformation was applied to make sure that the data are more homoskedastic, meaning that genes with very high and very low counts contribute equally to downstream calculations such as sample similarity. To detect overall differences between the samples a principal component analysis (PCA) was performed on the transformed expression values.

DESeq2 was applied for differential gene expression analysis

2.5.3.2 *Batch effect correction*

Batch effects were estimated using surrogate variable analysis implemented in the R bioconductor package sva [142]. It borrows information across genes in a dataset to construct covariates directly from expression data to capture unmodeled, unknown, or latent sources of noise [141]. These surrogate variables are then used to correct the gene counts for visualisation and are incorporated in the differential expression analysis.

We then performed the PCA on the corrected counts to check if the batch effect removal was successful as well as to detect differences between the hypothalamic samples in our dataset.

2.5.3.3 *Pairwise gene expression comparison*

The read count values of the two treatment groups within each of the five diet groups were normalised by the library size and then compared, resulting in five differential expression analyses and the corresponding fold changes. Differential gene expression testing was performed on the read counts only normalised for sequencing depth since DESeq2 considers the count variance on the mean value while estimating the dispersion, which can be seen as the square of biological variation [149]. To control for unwanted variation, the surrogate variables obtained with sva were added to the design matrix. DESeq2 then fits a generalised linear model (GLM) for each gene in the dataset. It assumes that the gene count follows a negative binomial distribution, taking the mean and the dispersion into account. The GLM returns then the log2 fold change between two compared sample groups [148]. The Wald test implemented in DESeq2 was then applied to detect significant differences in the calculated fold changes. The BH method [21] was used to control the FDR by correcting the obtained p-values for multiple testing. Genes with adjusted p-value$<$0.05 were considered to differ significantly, allowing an FDR

of 5%.

Genes with a mean base expression smaller than 50 and an absolute log2 fold change smaller than 0.2 were ignored as Hart et al. [94] had shown that low expressed genes might not play a functional role in the cell since they are often associated with repressed promoters. The R package clusterProfiler [258] was then used to test for overrepresented GO [14, 48] biological process terms among the significantly altered genes as described in Section 2.4.3.

2.5.3.4 *Comparison to gene expression pattern in specific neuronal populations*

Henry et al. [101] produced a valuable resource of single-cell RNA-Seq datasets to compare gene expression patterns in hypothalamic AgRP and POMC neurons of mice in the fasted and fed state. This gene expression data set (GEO Series GSE68177) was downloaded from NCBI's Gene Expression Omnibus (GEO) [17, 66] and DESeq2 was used to compare the samples of the fasted and fed mice for each neuronal population separately.

2.6 ANALYSES OF THE LIVER TRANSCRIPTOME AND METABOLOME DATASETS

2.6.1 *Identification of clinical and metabolic parameters influencing the liver metabolome and transcriptome profiles*

To investigate whether the observed metabolite concentrations and gene expression levels were influenced by any clinical or metabolic parameters, we first performed a PCA on each set to find the main factor separating the samples. Tested parameters were: the administered diet/intervention, the body weight measured at the end of the study (BWE), and the liver TAG level.

The Wilcoxon signed-rank test [250] was applied to assess if the parameters of interest differed between the compared groups.
Then, a relative importance analysis [231] was performed to assess the effect of the intervention, BWE, liver TAG level, and liver Cholesterol level on any metabolite concentration or gene expression measured in our dataset. Since these four regressors are intercorrelated relative importance analysis instead of a linear regression analysis was performed, which is better suited to handle intercorrelated regressors [86]. Here, a linear model was fitted for each gene and metabolite including the regressors "Intervention", "BWE", "Liver_TAG", and "Liver_Cholesterol". To decompose the model variance R's relaimpo package [85] was used. P-values were corrected for multiple testing applying the Bonferroni procedure [28] allowing an FDR of 5%.

2.6.2 *Differential liver gene expression analysis*

The R package limma [179, 191] (Section 2.4.2) was applied to infer transcriptional differences between the cohorts. The analysis was performed on the log2-transformed gene expression profiles. The BH procedure [21] was applied to correct for multiple testing. Genes with an adjusted p-value$<$0.05 were considered to differ significantly, allowing an FDR of 5%. The R package clusterProfiler [258] was then used to test for overrepresented GO [14, 48] biological process terms among the significantly altered genes as described in Section 2.4.3.

2.6.3 *Differential liver metabolite concentration analysis*

TwoGroup from the R metabolomics package [53] was used to compare the log2-transformed concentration levels between the cohorts. The BH procedure [21] was applied for the multiple testing correction. Metabolites with adjusted p-value$<$0.05 were considered to differ significantly, allowing an FDR of 5%.

2.6.4 *Linear regression analysis*

Linear regression analysis was performed to assess the predictive potential of liver metabolites under different dietary conditions. For each metabolite profile linear regression models were fitted for each dietary condition separately. We fitted a model for each of the three dependent variables: plasma cholesterol level, plasma insulin level, and plasma TAG level. We defined a model to be significant if it returned a non-adjusted p-value$<$0.0125. This cutoff was chosen, because not a single model would have been significant after multiple testing correction.

2.6.5 *Data preprocessing for data integration*

The log2-transformed metabolite dataset was scaled using the square root of the standard deviation as scaling factor (Pareto scaling [67]). Genes with high within-group variance (variance$>$0.5) were excluded from the downstream analysis to reduce the number of false positives identified due to noisy expression patterns.

For the first set the data preprocessing resulted in 9,391 gene expression profiles in the Chow dataset and 9,397 profiles in the HFD set. 8,888 of them were present in both sets. For the second set, this resulted in 9,605 gene expression profiles in the H$>$C set, 9,111 profiles in the CR set, and 9,349 profiles in the Ex4 set. 8,405 gene expression profiles overlapped between all three datasets.

2.7 CORRELATION-BASED NETWORK INTEGRATION (CONI)

Integrating data has become an important asset in todays research as it offers the opportunity to capture potential causative links. However, many integration methods are not easy to use or produce results that lack an ease of interpretability (Section 1.6). Here, we developed a framework called Correlation-based Network Integration (CoNI) to combine datasets from two related types of data to uncover hidden relationships and support hypothesis generation.

First, relationships are detected in one of the two datasets that are integrated using Pearson correlation. Next, CoNI exploits the ability of partial correlation to detect dependencies between variables for identifying relationships between factors of the two related datasets. In the last step, networks are generated that are easy to interpret despite capturing complex interactions. Since the method itself does not imply any biological principles or hypotheses, it could be applied to related datasets from all scientific fields such as physics, chemistry, and sociology. In the following, we will describe the method in more detail by a practical example for applying the method.

2.7.1 *Correlation-based Network Integration (CoNI) of liver metabolome and transcriptome datasets*

One purpose the new framework CoNI is highly apt for is to elucidate genetic control of metabolic networks, as metabolites are mainly transformed from one into another and genes control the rates. The framework includes three steps carried out for each of the investigated treatment groups independently:

1. Pairwise correlation analysis on the metabolite dataset
2. Partial correlation analysis combining the metabolite concentrations with the transcript expression profiles
3. Construction of undirected, weighted graphs

Starting from the pairwise metabolite correlations, an additional layer of genetic interactions will be added to the dataset. Then, those metabolite pair – gene triplets will be selected where the gene had a significant impact on the metabolite correlation. From these selected triplets an undirected graph will be built, where the nodes refer to the metabolites and the controlling genes form the edges. In this setting, an edge can consist of multiple genes and a gene can be mapped to multiple edges.

2.7.1.1 *Pairwise correlation analysis*

For each metabolite pair M_i and M_j given M, i = 1,...,n, j = 1,...,n, i≠j, the Pearson correlation coefficients were obtained (Figure 13) using R package Hmisc [93].

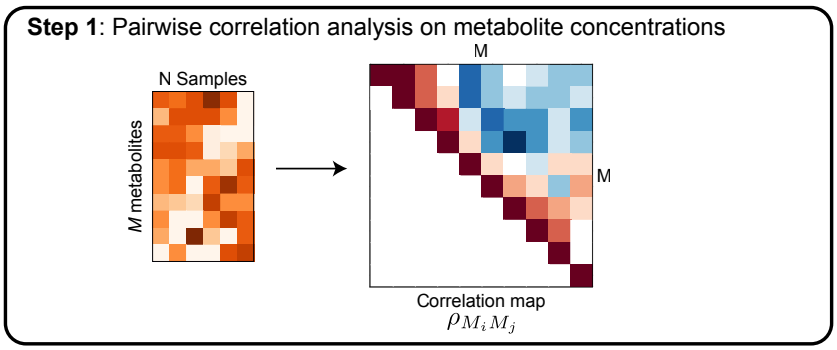


Figure 13: Pairwise metabolite correlation analysis

Metabolite pairs with a p-value<0.05 were chosen for further processing. Multiple testing correction was not performed as it would have restricted the number of possible interactions tremendously.

2.7.1.2 *Partial correlation analysis*

Partial correlation aims to investigate the influence of removing the linear effect of a variable on the correlation of a pair of random variables

Partial correlation aims to investigate the influence of removing the effect of a variable on the correlation of a pair of random variables. In our case, the partial correlation between metabolites M_i and M_j given M, i = 1,...,n, j = 1,...,n, i≠j, reflected the correlation between M_i and M_j after removing the linear effects of transcript G_k, given G, k = 1,...,n (Figure 14).

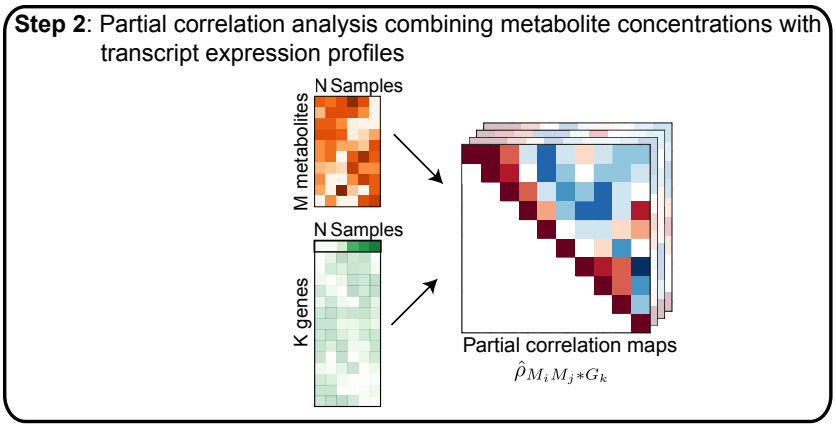


Figure 14: Partial correlation analysis assessing the correlation between two metabolites after controlling for the influence of a gene.

R's package ppcor [201] was used to obtain the Pearson partial correlation coefficients for each combination of selected correlating metabolite pairs and each gene in the noise-reduced set of gene expression profiles. Such a combination will be denoted as triplet in the following.

To select triplets, for which the partial correlation coefficient differed significantly from the correlation coefficient of the respective metabolite pair, Steiger's test [217] was adapted. The original test assesses the significance for the difference between two correlation coefficients that share one variable. If we, for example, compare the correlation coefficients ρ_{ha} and ρ_{hb}, h would be the shared variable. The significance depends on the intercorrelation between variables a and b (ρ_{ab}), which has to be supplied as additional parameter. To assess the difference in the magnitude of the two correlations, Steiger modified the test statistic developed by Dunn and Clark [63] by using the average of the two compared correlation coefficients to calculate the z value [59, 217]:

$$z = ((Z_{jk} - Z_{jh})\sqrt{(n-3)})/(\sqrt{(2-2c)}), \tag{3}$$

where

$$\bar{r} = (r_{jk} + r_{jh})/2 \tag{4}$$

and

$$c = (r_{kh}(1-2\bar{r}^2) - (1/2)\bar{r}^2(1-2\bar{r}^2-r_{kh}^2))/((1-\bar{r}^2)^2) \quad (5)$$

Here, we wanted to compare a partial correlation coefficient and a correlation coefficient. Therefore, we applied the test twice. The provided additional parameter was in the first test the correlation between M1 and G and in the second test between M2 and G: ρ_{M1M2} vs. ρ_{M1M2G} with ρ_{M1G} or ρ_{M2G} as related correlations, where M1 denoted the first metabolite of the pair, M2 the second metabolite, and G a gene. To be selected, the triplet had to reject the null-hypothesis - stating that the correlation coefficients do not differ - significantly (Bonferroni [28] adjusted p-value<0.05, allowing an FDR of 5%) in both tests. The Steiger test is implemented in the function cocor.dep.groups.overlap of R's cocor package [59], which we used to perform the testing.

2.7.1.3 *Undirected graph construction and clustering*

An undirected and weighted graph was generated using the R package igraph [50]. Nodes were formed by metabolites and genes set up the edges (Figure 15). Edges were drawn if a metabolite pair correlated and this correlation was influenced by at least one gene.

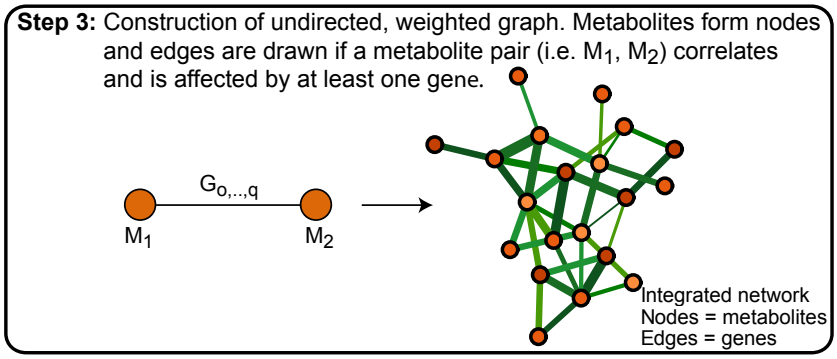


Figure 15: Selected triplets were used to generate an undirected and weighted graph. Metabolites form the nodes and the genes are stored as lists in the edges.

Several genes can be stored in several edges connecting more than two correlating metabolites and a pair of metabolites might also have several genes stored in the edge connecting them, indicating that it can be influenced by more than one gene. The number of genes connecting the two respective metabolite nodes determines the edge weight.

2.7.1.4 *Analysis of undirected and weighted graphs*

The resulting graphs were compared regarding several aspects such as connectivity, number of edges, metabolites (a.k.a. nodes), and genes

present in the network. String protein database [196] was used to quickly check for overrepresented GO [14, 48] molecular function terms among the genes present in the two networks.

2.7.2 *Prioritising candidate genes by local regulator gene (LRG) identification*

Local regulator genes have a specific impact on a local sub-graph

The network stored hundreds of genes within its edges. One way to identify differences in the genetic program of the studied phenotypes, is to select genes with a specific impact on a local sub-graph within the estimated networks. These genes control a densely connected metabolic sub-network and could, in theory, specifically regulate a metabolic pathway. We defined these so-called local regulator genes (LRGs) as genes significantly enriched within a local sub-graph.

For each node in the network, all genes were selected that appeared in the edges to the respective first indirect neighbour following each outgoing edge. To test if a gene can be observed ("drawn") more often than expected by chance, and can therefore be called a LRG, binomial distribution testing was performed. P-values were Bonferroni [28] corrected for multiple testing and genes with adjusted p-value$<$0.05 were selected, allowing an FDR of 5%.

2.7.3 *In vitro validation of obesity-related LRGs*

2.7.3.1 *Database query for associated disease-relevant SNPs for all LRGs*

To validate candidate genes in vitro, one option is to query additional databases for information on these hits, such as their function, pathways they are involved in, or known associations with disease-related SNPs. Therefore, the Type 2 Diabetes Knowledge Portal (`http://www.type2diabetesgenetics.org/`) was queried to see which of the obtained LRGs were associated with SNPs relevant for obesity or related disease markers.

2.7.3.2 *SiRNA-mediated knockdown of selected LRGs*

5 local regulator genes were validated by siRNA-mediated knockdown in HepG2 cells

To validate the results of the developed framework, five LRGs were selected that were either differentially expressed between chow- and HFD-fed mice or had been assigned to SNPs associated with obesity in humans. We wanted to confirm the influence of the selected genes on the metabolome since we hypothesised that these genes were important for a metabolite sub-network. Therefore, changes in the gene expression should change the correlation pattern of the metabolites.

An siRNA-mediated knockdown of each of these five genes was performed in HepG2 cells separately. The silencing mechanism in-

duced by the siRNA should degrade the mRNA of the corresponding gene [176]. RNA was extracted from HepG2 cells after siRNA knockdown using the NucleoSpin RNA isolation kit (Macherey-Nagel, Düren, Germany). Equal amounts of RNA were reverse transcribed to cDNA using the QuantiTect Reverse Transcription kit (Qiagen, Hilden, Germany). Gene expression was analysed using TaqMan probes for APPL2 (Hs01565861_m1), COBLL1 (Hs01117513_m1), GK (Hs02340007_g1), INHBE (Hs00368884_g1), RAPGEF4 (Hs00199754_m1), and HPRT (Hs02800695_m1) as the housekeeping gene with the respective TaqMan mastermix (Thermo Fischer Scientific, Inc., Rockford, IL USA). qPCRs were carried out using a Quantstudio 6 Real Time PCR System (Applied Biosystems). Gene expression was evaluated using the Δ-Δ Ct method.
Gene expression analysis by qPCR confirmed that the siRNA-mediated knockdown was successful as mRNA levels had been reduced by more than 60% compared to non-target siRNA controls.
For each gene, two biological replicates with three technical replicates for each gene specific knockdown as well as for the non-target siRNA control knockdown were obtained. The Biocrates AbsoluteIDQ^{TM} p180 kit was used to quantify metabolite concentrations in the cell samples as described in Section 2.3.2.

The measured metabolite concentrations were first normalised to the relative cell count of the corresponding sample measured by Hoechst staining of nuclei. One replicate of the Cobll1 specific knockdown had to be excluded due to an error during sample preparation. Metabolites with missing values were excluded from the whole dataset due to the overall low sample size. The metabolite concentrations were log-transformed and pareto-scaled as described before (Section 2.3.2).

To assess how the metabolites that are connected by the edges containing the selected LRGs behave in the in vitro setting, Pearson correlation coefficients were obtained for these metabolite pairs for the non-target siRNA control and gene specific siRNA knockdown samples separately. R's cocor.indep.groups function implemented in the cocor package [59] was used to compare the corresponding pairwise correlation coefficients between knockdown and control for each metabolite pair.

A one way ANOVA was performed to compare the log2-transformed and pareto-scaled concentration levels between the siRNA-mediated knockdown and control samples. The p-values were FDR-corrected as introduced by Storey [220] and metabolites with adjusted p-value$<$0.05 were considered to differ significantly.

2.7.4 *In vivo method validation*

2.7.4.1 *Human cohort of subjects with liver tissue samples*

170 liver tissue samples of 106 men and 64 women of European descendent undergoing liver surgery at the Department of General, Visceral, and Transplant Surgery at the University Hospital of Tübingen (Tübingen, Germany) were obtained to analyse the gene expression in human livers (Table 2).

Table 2: Characteristics of subjects who donated liver samples.

	MEANS ± SD	MEDIAN (25-75%)
Age (years)	63.5 ± 11.7	65 (58-72)
BMI (kg/m²)	25.2 ± 4.1	25.2 (22.2-27.5)
Liver fat content (%)	2.16 ± 2.08	1.36 (0.85-2.65)

Before collection of the liver biopsies the subjects were fasted overnight. In a subgroup of 77 subjects (49 men and 28 women) fasting plasma samples were obtained for the calculation of the homeostasis model assessment of insulin resistance (HOMA-IR) as proposed by Matthews et al. [156] (Table 3).

Table 3: Characteristics of subjects with available fasting blood samples.

	MEANS ± SD	MEDIAN (25-75%)
Age (years)	62.1 ± 12.1	64 (55-71)
BMI (kg/m²)	24.7 ± 4.1	24.3 (21.5-26.7)
Liver fat content (%)	1.72 ± 1.35	1.22 (0.85-2.11)
Fasting glucose (mg/dl)	98.5 ± 25.8	92 (82-115)
HOMA-IR (arbitrary units)	2.26 ± 2.96	1.38 (0.71-2.36)

All subjects tested negative for viral hepatitis and did not show signs for liver cirrhosis. Liver samples were taken from normal, non-diseased tissue during surgery, immediately frozen in liquid nitrogen, and stored at -80 °C. The tissue was judged by an experienced pathologist and only samples from normal, non-diseased tissue were included in the study. Informed, written consent was obtained from all participants, and the Ethics Committee of the University of Tübingen approved the protocol (239/2013BO1) according to the Declaration of Helsinki.

2.7.4.2 *Determination of liver tissue triglyceride content*

Liver tissue samples were homogenised in phosphate-buffered saline containing 1% Triton X-100 with a TissueLyser (Qiagen, Hilden, Germany). Triglyceride concentrations in the homogenate were quantified using an ADVIA XPT clinical chemistry analyser (Siemens Healthineers, Eschborn, Germany) to determine the liver fat content. Results were calculated as mg/100 mg tissue weight (%).

2.7.4.3 *Real-Time PCR*

Frozen tissue was homogenised in a TissueLyser (Qiagen) for real-time (RT)-PCR and quantitative RT-PCR analysis of hepatic mRNA expression and RNA was extracted with the RNeasy Tissue Kit (Qiagen) according to the manufacturer's instructions. Total RNA was treated with RNase-free DNase I and was transcribed into cDNA by applying a first-strand cDNA kit. PCRs were performed in duplicates on a LightCycler480 (Roche Diagnostics, Mannheim, Germany). The human primer sequences that were used are shown in table A11. Data are presented relative to the housekeeping gene Rps13 using the Δ-Δ Ct method.

2.7.4.4 *Quantification of blood parameters*

Plasma insulin was determined on the ADVIA Centaur XPT chemiluminometric immunoassay system. Fasting plasma glucose concentrations were measured using the ADVIA XPT Clinical chemistry analyser (both Siemens Healthineers, Eschborn, Germany).

2.7.4.5 *Statistical analyses*

Data were tested for normal distribution applying the Shapiro-Wilk W-test and data that were not normally distributed were log-transformed. Person correlation was used to test for univariate associations between parameters. Multivariate linear regression analyses were applied to adjust the effects of covariates and identify independent relationships. The statistical software package JMP 14.0 (SAS Institute, Cary, NC) was used.

2.7.5 *Particularity in the application of CoNI to weight loss intervention datasets*

To apply CoNI to the second dataset, containing the samples of the weight loss intervention groups Ex4, CR, and H>C, all preprocessing steps and analyses were performed as described above (Section 2.7.1). In contrast to the Chow vs. HFD dataset set, the linear regression analysis was not performed on the weight loss datasets since the PCA already identified the parameter with the largest influence on

metabolite concentration levels.

When selecting the edges for the Ex4 and H>C networks, the numbers of selected edges were too large to be used for further analyses. Therefore, only 503 edges were allowed to be selected for the H>C network and 565 for the Ex4 network by setting the adjusted p-value cutoffs to 1.0E-05 for H>C and to 1.0E-09 for Ex4. String protein database [196] was used to quickly check for overrepresented Reactome [68] pathways among the genes present in the two networks.

3

RESULTS AND DISCUSSION

3.1 TIME-MATCHED ANALYSIS OF DNA ADDUCT FORMATION AND EARLY GENE EXPRESSION IN METHYLAZOXYMETHANOL ACETATE TREATED EKER RATS

Carcinogenicity testing is a critical component of the safety assessment of newly developed drugs and chemicals. Appropriate regulation of a carcinogenic compound requires a definition of its mode of action in order to clarify whether a safe no-adverse-effect-level (NOAL) can be defined. Key mechanisms contributing to the NOAL of carcinogenic compounds are carcinogen detoxification and DNA repair. While an NOAL can be usually defined for a non-genotoxic carcinogen, threshold doses for genotoxins are still debated. The absence or presence of DNA adducts can serve as a biomarker to distinguish genotoxic from non-genotoxic carcinogens. However, the quantification of DNA adducts alone cannot predict if the exposure to a hazardous substance causes mutations in critical genes that lead to a future cancer risk. The majority of DNA adducts and related DNA damage may be repaired and their predictive value is therefore limited.

The quantification of DNA adducts alone cannot predict if the exposure to a hazardous substance causes mutations in critical genes

We here hypothesise that the complex interplay of various factors that transform DNA adducts into tumour causing mutations could be reflected by very specific expression changes that can be captured using gene expression measurements on several time-points after genotoxin exposure. In combination, the gene expression pattern and the DNA adduct quantification could serve as an early effect biomarker in response to DNA adduct formation.

Stemmer et al. [218] measured DNA adducts and gene expression profiles after 1, 3, 7 and 14 days of low-dose carcinogen exposure in Eker and wild type rats and also captured the number and incidence of pre-neoplastic lesions after six months of continuous exposure to the genotoxin. They investigated the three toxins ochratoxin A (OTA), aristolochic acid (AAc), and MAMAc. Stemmer et al. [218] published the findings on OTA and AAc. The not-published MAMAc data gave us the unique opportunity to investigate the predictive value of combined analysis of pro-mutagenic DNA adducts together with time-matched gene expression changes.

In this first part of the thesis, differential gene expression analysis and pathway enrichment analysis were applied to examine if transcriptional induction of genes involved in DNA repair or cancer related pathways can evaluate the effectiveness of cellular protection and allows to assess the risk of DNA adducts to translate into tumourigenesis. This work is the result of a collaboration with the Chair of Human and Environmental Toxicology, University of Konstanz. The collaboration partners performed all rat and wet lab experiments. Excerpts and figures presented here have been published previously in Archives of Toxicology (Klaus V*, Bastek H* et al 2017; *Equal contribution).

3.1.1 *Expression profiles differ between WT and Eker rats*

PCA separated WT and Eker samples

The PCA performed to get an overview of the gene expression profiles of all 108 samples showed that the genotype caused the strongest differences (Figure 16). The first principal component, which captured 16.21% of the variance, clearly subdivided the samples into WT and Eker clusters. The second principal component capturing 6.29% of the variance subdivided the samples of rats treated with OTA and their Eker controls from the other samples.

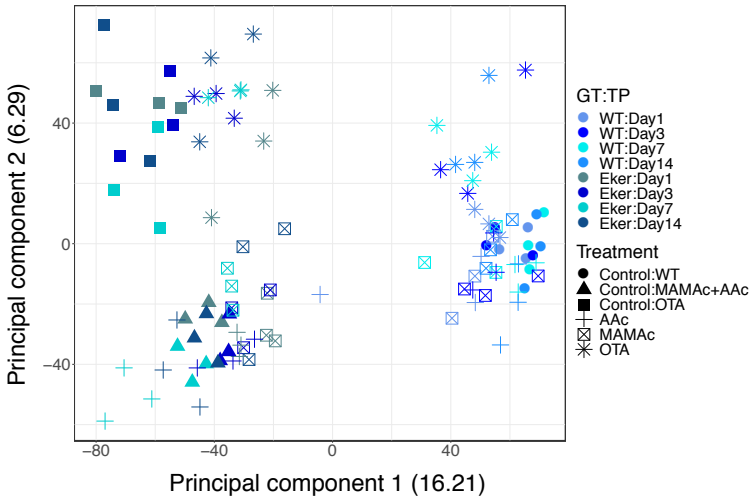


Figure 16: PCA of gene expression profiles of WT and Eker rats treated with toxins AAc, MAMAc, and OTA. The shape of the data points indicates the administered treatment. The colouring shows the respective genotype (GT) and the sampled time-point (TP).

One explanation for the separation of the OTA treatment samples from the others in the Eker rats could be that OTA is in contrast to AAc and MAMAc a non-genotoxic substance. This however would not explain why the controls differed from the AAc and MAMAc controls. Therefore, the most likely explanation would be the pres-

ence of batch effects.

In the Eker samples, there was a slight separation between the MAMAc and AAc treated mice and the corresponding control samples. The samples of the same time-points, treatment, and genotype tended to cluster together, but we could not observe large differences between the time-points of the same genotype and treatment.

3.1.2 *Transient gene expression changes upon MAMAc treatment in Eker rats*

On the first day after MAMAc treatment formed O^6-mG and *N7*-mG DNA adducts could be observed in the Eker rat cohort, which showed that DNA adducts are indeed a sensitive biomarker to detect genotoxin exposure as described in La and Swenberg [137] and Peluso et al. [177]. This is in line with findings of Sohn et al. [212] and Kisby et al. [130], who observed O^6-mG and *N7*-mG adducts after a single high dose of MAMAc (20 mg/kg BW, i.p.) in the liver, kidneys and the brain.

We wanted to assess if MAMAc treatment and the following DNA adduct accumulation correspond to the gene expression profile in the kidney cortices of investigated rats within the first 14 days of administration.

MAMAc treatment and DNA adducts did not alter gene expression for the two weeks after first exposure

No gene expression changes in WT rats after first day of MAMAc exposure

First, we compared each of the four time-points (day 1, day 3, day 7, and day 14) with the preceding one within the respective genotype for the treatment and control group separately. This was done to see if we can observe changes over the 14 days within the cohorts.
We could not observe any gene that changed significantly over time in the WT samples, neither in the control samples nor in the MAMAc treated WT rats.

MAMAc treatment and DNA adducts did not influence gene expression after the first day

Over the period of 14 days, 35 probe-IDs, that correspond to 34 genes, changed significantly in the Eker control group (Table A1) and 22 probe-IDs, that correspond to 21 genes, differed in the Eker cohort treated with MAMAc (Table A2). As six of the 35 probe-IDs differing between time-points in the controls differed also in the treated rats, their expression change seemed to be independent from the MAMAc treatment. This response can only be considered to be very mild. The small number of genes did not allow the performance of a pathway enrichment analysis.

Taken together, the administration of MAMAc and the induced DNA adducts did not hamper the gene expression for the 13 days after first exposure.

MAMAc induced DNA adducts manifest in an instantly altered gene expression profile

Second, we investigated the differences in gene expression for each sampled time-point between treated and control rats of each genotype cohort, to find out which genes differ upon MAMAc exposure on the respective day after treatment initiation.

On the third day no gene was altered between treated and control rats for both genotype groups. For the WT rats we observed one gene that was altered on day 7, and two genes that differed on day 14. For the Eker cohort, we observe one gene to be differentially expressed between treated and control rats on day 7 and no gene to be altered on day 14.

In summary, we observed few genes that were differentially expressed after the first day of MAMAc treatment between treated and control rats of both genotype cohorts. Since the previous analysis already hinted to a stable gene expression pattern over the measured period of time after the first day in the control groups as well as in the exposure groups, we did expect this result.

Comparing the expression patterns for the treated and control rats measured on the first day, we observed two genes that differed in the WT rats and 86 differentially expressed probe-IDs (76 could be annotated with gene symbols) upon MAMAc administration in the Eker rats.

Genes responded in transient manner to MAMAc exposure

The heat map (Figure 17) of the 87 probe-IDs altered at day 1 or day 7, respectively, demonstrates that most genes were regulated in a transient manner, which means that they kept their expression pattern constant across the complete measured time frame. The strongest signal was shown after the first day of MAMAc administration.

We could observe DNA adducts induced by MAMAc treatment after the first day of exposure in the Eker rats. The induction of DNA adducts manifested in an instantly altered gene expression profile. Even if the DNA adduct levels increased over the next 13 days, indicating that adduct formation exceeded the rate of adduct loss, the gene expression patterns of the surrounding tissue remained stable for that time frame.

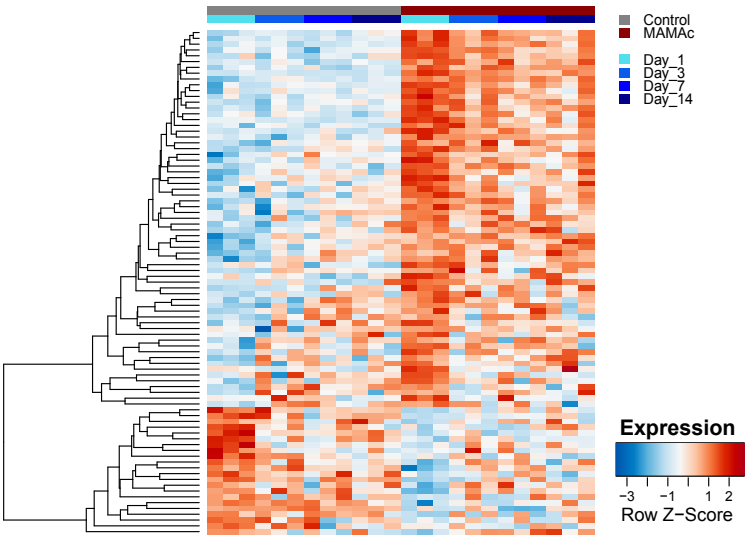


Figure 17: Heat map of probe-IDs altered by MAMAc treatment in Eker rats on days 1 and 7.

This shows that the MAMAc treatment and the following DNA adduct accumulation corresponded to the gene expression profile in the sampled kidney cortices.

3.1.3 *No transcriptional DNA damage response upon MAMAc exposure*

To identify the pathways to which the 76 annotated genes, altered on the first day of MAMAc exposure in the Eker rats, belong to, we performed pathway enrichment analyses on KEGG [121–123] pathways and on GO [14, 48] biological process terms.

10 genes up-regulated upon MAMAc treatment in Eker rats belonged to cancer-related pathways

We observed an effect of MAMAc exposure on 18 KEGG pathways (Table A3). Ten up-regulated genes belonged to pathways related to cancer. We used the String database [196] to find out if the proteins encoded by these ten genes interact. We could observe that these genes were involved in MAPK signaling pathway (KEGG adj. p-value = 0.0017), which plays a role in cancer relevant pathways such as cell proliferation and apoptosis [263], and in FoxO (KEGG adj. p-value = 0.002) and TGFβ signaling pathways (KEGG adj. p-value = 0.016), which are involved in tumour suppression in the healthy state but can lead to cancer progression and tumourigenesis when dysfunctional [69, 263].

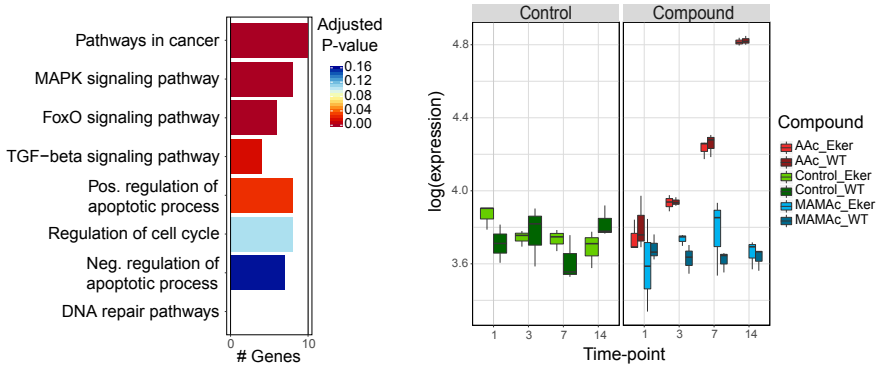
Whether the transcriptional induction of cancer-involved genes was caused by the pro-mutagenic DNA adduct formation needs to be investigated further. However, it may hint to a possible translation of DNA adducts induced by MAMAc into tumour causing mutations,

which could serve as a predictive tool.

In the healthy state, several pathways protect the cells from mutations that might lead to cancer, eventually. MAMAc is a genotoxic substance and the kidney cortex is the tissue that gets in contact with the highest dose of the toxins circulating in the body. Therefore, we would have expected that the treatment with MAMAc would induce transcriptional activation of genes involved in cell protection pathways in the kidney cortex.

We looked at the following protection pathways: DNA repair pathways (DNA replication, mismatch repair, excision repair, homologous recombination, direct reversal), apoptosis (pro- and anti-apoptotic pathway), cell cycle (cell cycle progression, cell cycle arrest), TGFβ signaling pathway, and mTOR signaling pathway.

Eight of the 76 genes with significant expression changes after MAMAc treatment were involved in pro-apoptotic pathways (GO adj. p-value = 0.028) (Figure 18a and Table A4). In contrast, the enrichment of the seven significantly altered anti-apoptotic genes was not significant (GO adj. p-value = 0.16) as well as the enrichment of the eight differentially expressed genes involved in cell cycle regulation (GO adj. p-value = 0.1) (Figure 18a and Table A4).



(a) Selected pathways (b) Mgmt expression

Figure 18: **(a)** Pathway enrichment of genes altered by MAMAc treatment in Eker rats. The number of genes present in a pathway is shown in the x-axis. The bar colour visualises the adjusted p-value. **(b)** *Mgmt* expression after MAMAc and AAc treatment compared in WT and Eker rats.

Even though we observed genes involved in apoptosis or pathways relevant for early cancer stages, the histopathological analyses of the kidneys from the same animals did not show any sign of increased mitotic events or apoptotic cell death in the up to 14 days MAMAc treated Eker rats. This was in line with the only slightly increased

incidence of tumours in MAMAc treated Eker rats compared to the vehicle-treated Eker rats. Taking into account that Eker rats are very sensitive towards genotoxic carcinogens such as aristolochic acid [218, 219] or dimethylnitrosamine [240], we had expected a higher increase. We could observe only few DNA adducts that had manifested in mutation causing tumours, which might be explained by the observed inability of MAMAc to increase cell proliferation. One critical event necessary to convert DNA damage into heritable mutations and for carcinogenesis is increased cell proliferation [45, 184]. Early expression changes of cell cycle regulating genes following up to 14 days of AAc and OTA treatment [218] could predict an increase in cell proliferation after six month of AAc and OTA administration [219]. However, we observed negligible changes of cell cycle regulating genes in the 14 days MAMAc exposed Eker rats and a missing increase of cell proliferation markers upon chronic MAMAc treatment.

Histopathology did not show increase in mitotic events or apoptotic cell death in MAMAc treated Eker rats

MAMAc was not able to increase cell proliferation

No gene involved in any of the DNA repair pathways that should respond to the observed DNA adduct formation was significantly altered upon MAMAc exposure (Figure 18a). Even the primary repair enzyme for O^6-mG adducts, O6-Methyl-guanin-DNA-Methyltransferase (*Mgmt*) did not respond to MAMAc treatment (Figure 18b). Stemmer et al. [218] showed that *Mgmt* is induced in a time-dependant manner in the kidney cortex of male Eker rats following treatment with the positive control genotoxin AAc, which is in line with our observations (Figure 18b). One possible explanation for the discrepant finding could be that DNA adducts have to accumulate to reach a specific threshold level to activate DNA damage response genes to a detectable level. This hypothesis is in agreement with an in vitro study performed by Clewell et al. [42] showing that three tested genotoxic agents increased micronuclei formation significantly but failed to induce expression changes in *p53* and its downstream DNA repair genes.

No gene involved in any of the DNA repair pathways was significantly altered

3.1.4 *Lack of DNA damage response not due to partial loss of Tsc2*

Habib et al. [89] showed that a partial loss of *Tsc2* in kidney cortices of Eker rats reduced the expression levels of the DNA repair enzyme 8-oxoG-DNA glycosylase (*Ogg1*) compared to wild type rats, which contributes to tumourigenesis. Therefore, we wondered if the chronic reduction of DNA repair pathway genes by genetic haploinsufficiency of *Tsc2* in our Eker rats could be the reason for the lack of transcriptional regulation after MAMAc exposure in Eker rats.

In our hands, *Ogg1* did not show expression differences between WT and Eker control rats (Figure 19). The discrepancy might be a result of the difference in the methods used to measure *Ogg1*. While

Habib et al. [89] measured Ogg1 protein expression, we measured its expression on the mRNA level.

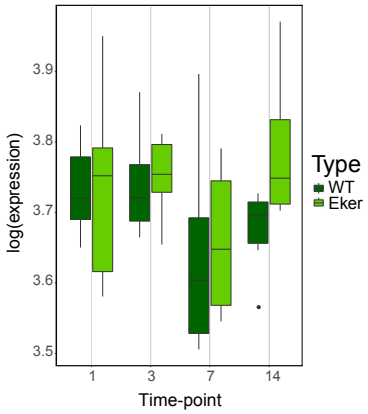


Figure 19: Expression of DNA repair gene *Ogg1* in WT and Eker control rats.

The absent down-regulation of *Ogg1* mRNA levels could be explained by the weak correlation reported between mRNA and protein abundance levels due to post-transcriptional and post-translational factors, and noise and experimental errors that can reduce protein abundance [153].

As mentioned in Section 3.1.2, we did observe five genes that were altered upon MAMAc exposure in the WT rats on all four time-points. None of these genes is involved in any of the DNA repair pathways (DNA replication, mismatch repair, excision repair, homologous recombination, direct reversal, base excision repair, non-homologous end joining, and Fanconi anemia pathway), indicating that WT rats did also not show a significant induction of DNA repair genes following up to 14 days of low dose MAMAc treatment.

Weak effect of partial Tsc2 loss on transcription levels of DNA repair genes

We compared the gene expression profiles for Eker and WT rats on each sampled time-point. We could observe eight probe-IDs that could be assigned to seven genes involved in one of the DNA repair pathways being significantly deregulated on each measured time-point (Table 4).

Table 4: Differentially expressed DNA repair genes in Eker and WT rats.

PROBE-ID	GENE	KEGG PATHWAY	DIRECTION OF REGULATION	ADJUSTED P-VALUE
1390040_at	*Bre*	Homologous recombination	Down	9.037353e-06
1373451_at	*Rnaseh2b*	DNA replication	Down	1.007141e-05
1392807_at	*Lig3*	Base excision repair	Down	1.809460e-05
1372793_at	*Ssbp1*	DNA replication, Homologous recombination, Mismatch repair	Up	5.772972e-05
1388367_at	*Pole4*	DNA replication, Nucleotide excision repair, Base excision repair	Down	1.200040e-04
1374318_at	*Brcc*	Homologous recombination	Up	1.258071e-04
1374044_at	*Stra13*	Fanconi anemia pathway	Up	1.458302e-04
1374010_at	*Lig3*	Base excision repair	Up	8.834112e-04

The direction of the expression change did not show a consistent pattern, as these probe-IDs are up- and down-regulated in equal shares.
This indicates that a partial loss of *Tsc2* affected the overall transcription levels of DNA repair genes only marginally.

3.2 EFFECTS OF WEIGHT LOSS INTERVENTIONS ON HYPOTHALAMIC GENE EXPRESSION

Pharmacologically induced body weight loss by treatment with the *Glp-1* agonist exendin-4 leads to a stabilisation of the lower body weight after weight loss in mice, which is not the case when body weight loss was induced by calorie restriction [167]. In this second part of the thesis, differential gene expression analysis and pathway enrichment analysis were applied to hypothalamus samples of mice injected with saline or leptin after pharmacologically induced weight loss compared to calorie restriction in order to gain more insight into the underlying molecular mechanisms.

This work is the result of a collaboration with three research units at the Helmholtz Centre Munich, namely the Research Unit Neurobiology of Diabetes, the Core Facility Next-Generation-Sequencing, and the Research Unit for Protein Science. The collaboration partners performed all mouse and wet lab experiments.

3.2.1 *PRC RNA-Seq dataset shows bias for IP:Leptin samples*

3.2.1.1 *Read quality assessment and mapping statistics of RNA-Seq reads*

Sample sizes differed massively between the three sample sets

When checking the read quality by applying FastQC no abnormality for the quality of the samples was observed for either the Clontech SMART-Seq or the Illumina TruSeq protocol. However, the sample sizes differed massively between the three sample sets with the IP samples having the highest number of reads (Table 5). At the same time the IP samples had the least percentage of reads mapped to the reference genome.

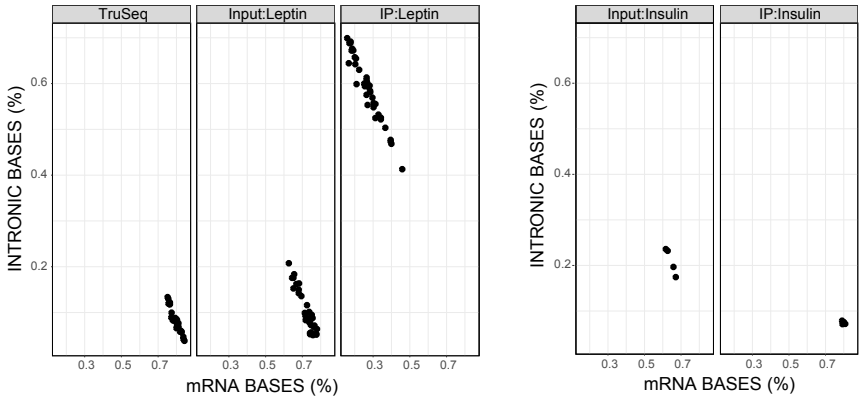
Table 5: Mapping statistics of the three sequencing protocols (m = million).

PROTOCOL	CONTENT	SAMPLE SIZE	MAPPING FRACTION
TruSeq	Whole tissue	57 ± 9m	$98.85 \pm 0.59\%$
Input	Whole tissue	70 ± 11m	$92.42 \pm 2.06\%$
IP	mRNA bound to ribosomes in activated neurons	83 ± 15m	$89.82 \pm 2.0\%$

3.2.1.2 *High proportion of intron-mapping reads observed in IP:Leptin dataset*

Majority of reads of the IP:Leptin dataset mapped to intronic regions

We next assessed the genomic region to which the reads could be mapped, thereby focusing on reads mapping to exonic and intronic regions of the genome. Usually, the majority of reads maps to exonic regions [75, 166, 225], which was true for the Leptin PRC TruSeq and Input:Leptin datasets (Figure 20a) as well as for the Input:Insulin and IP:Insulin datasets (Figure 20b). In contrast, the majority of reads of the IP:Leptin dataset mapped to intronic regions. The average of reads that mapped to intronic regions in the IP:Leptin samples was 58.9% the minimum was 41.3%.



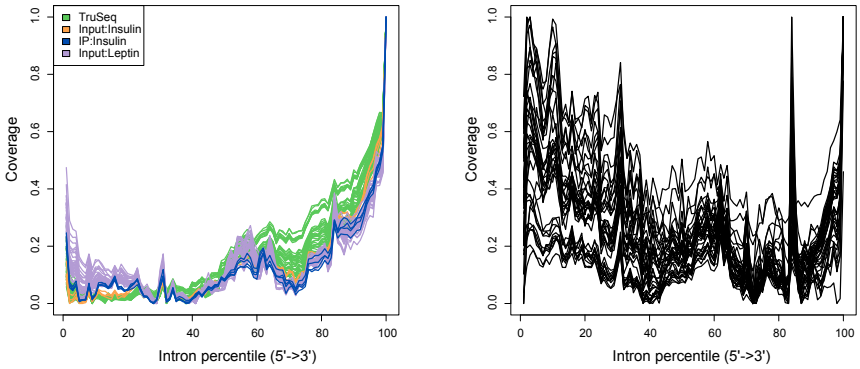
(a) Leptin dataset (b) Insulin dataset

Figure 20: Genomic regions of mapped reads shown for each protocol. The fraction of reads mapped to the intronic regions is shown on the y-axis, the fraction of the reads mapping to the exonic regions is shown on the x-axis. **(a)** shows the fractions for the leptin samples and **(b)** shows the fractions for the insulin samples.

IP:Insulin and IP:Leptin samples differ fundamentally

To check whether these intron-mapping reads showed a positional bias, the distribution of the relative position of reads mapping to introns was assessed for each dataset (Figure 21). All datasets except the IP:Leptin dataset showed a homogeneous pattern which differs drastically from the read distribution observed in the IP:Leptin dataset (Figure 21b). Taken together this hints to fundamental differences between the IP:Insulin and IP:Leptin samples.

Whereas the IP:Insulin samples were used to establish the method in our lab, the IP:Leptin samples were generated with the modified protocol adapted to the large scale sample throughput. The major difference between the applied protocols was the antibody used to capture the mRNA bound to ribosomes of activated neurons.



(a) Intronic read distribution (b) Intronic read distribution: IP:Leptin

Figure 21: Read mapping distribution in intronic regions shown for **(a)** samples of the TruSeq, Input, IP:Insulin, and the **(b)** IP:Leptin datasets. The x-axis shows the read mapping position within the intron respective to the intron length.

3.2.1.3 *Mass spectrometry shows reduced antibody specificity of PBS antibody*

To investigate the differences of the two IP datasets that had been generated independently, mass spectrometry analysis of the proteins pulled down by either antibody formulation with or without addition of the 3P peptide was performed: Commercial antibody (ComAB), Commercial AB + 3P (ComAB3P), PBS antibody (PBSAB), and PBS AB + 3P (PBSAB3P).

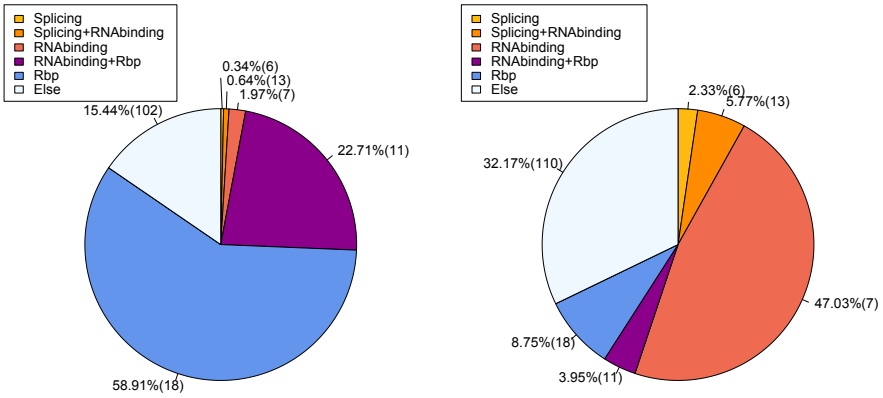
The mass spectrometry analysis quantified 297 annotated proteins that were present in all four IP samples. After removing proteins to which only one peptide could be assigned to and after removing proteins that were not measured reliably, about 170 proteins were used in the downstream analysis.

Both antibody formulations are able to recognise the target protein RpS6

Since the Phospho-S6 Ribosomal Protein (Ser240/244) antibody was designed to specifically bind the two C-terminal sites 240 and 244 of phosphorylated ribosomal protein S6 [132, 267], we first checked whether RpS6 could be detected in all IP samples. The abundance ratio (AR, calculated by dividing the normalised protein abundance of the respective antibody by the normalised protein abundance observed in the iGg control) of RpS6 was highest in the IP samples extracted with the PBSAB with 63.8, followed by the ComAB3P with 48.8. The ARs of the PBSAB3P and ComAB were 27.0 and 28.6, respectively. This indicated that both formulations of the RpS6 antibody were able to recognise the target protein RpS6. However, the total amount of RpS6 protein in the pulldown was rather low (Table A5).

One explanation for the detection of only low levels of RpS6 could be, that the peptide fragments after trypsin digestion were too small to be aligned back to a protein after mass spectrometry. This is a common problem for all alkaline proteins. Additionally, small proteins are in general more difficult to detect. Therefore, not detecting or detecting only low levels of RbS6 was not necessarily a consequence of antibody un-specificity.

The higher number of reads mapping to intronic regions in the PBS:IP samples suggested that the PBSAB was binding nonspecifically and might have extracted mRNA that was not bound to ribosomes and therefore still contained intronic regions. Thus, the functional classes of the extracted proteins were analysed in the different IP samples, focusing on ribosomal proteins as well as proteins involved in splicing and RNA-binding. 165 proteins were analysed for the PBSAB and the PBSAB3P, whereas 142 proteins were taken into account for ComAB and 157 for the ComAB3P. The percentages of the class abundances (Section 2.5.2) of the functional classes for the ComAB3P that had been used to generate the IP:Insulin samples and the PBSAB that had been used to generate the IP:Leptin samples are illustrated in figure 22. The percentages of the class abundances for the ComAB and PBSAB3P are shown in figure A1.



(a) Commercial AB + 3P (b) PBS AB

Figure 22: Comparison of protein abundances detected in samples extracted with the **(a)** Commercial antibody + 3P used to produce the IP:Insulin dataset and the **(b)** PBS antibody used to capture the ribosomes in the IP:Leptin dataset. The number in the brackets gives the number of quantified proteins.

The main difference between the sample sets generated with the two antibody formulations was the class abundance for ribosomal binding proteins (Rbp, Figures 22 and A1, shown in blue) versus the abundance for RNA binding proteins (Figures 22 and A1, shown in salmon). The percentage of class abundance of the commercial anti-

body formulation for Rbps was 59% for the ComAB3P and 48% for the ComAB sample, respectively, whereas it was below 10% in the IP samples generated with PBS formulation. In contrast, samples obtained with the PBS antibodies (Figures 22b and A1b) showed higher percentages of class abundances (47 and 42% vs. 2 and 5%) of proteins involved in RNA binding when compared to the commercial formulation of the antibody (Figures 22a and A1a). Further, the percentages of class abundances of proteins involved in splicing (Figures 22 and A1, shown in yellow) and splicing + RNA binding (Figures 22 and A1, shown in orange) were much higher in the PBSAB samples compared to the ComAB samples.

Taken together, the pattern of extracted proteins of the PBS antibodies indicated that the higher concentrated, specifically synthesised formulation of the antibody bound less specific to ribosomal proteins than the commercially available antibody.

Non-RpS6-specific binding pattern of the PBS antibody could explain high amounts of intronic mRNA in the IP:Leptin dataset

It can be assumed that the less specific binding pattern of the PBS antibody to Rbp (Figure 22b) used to generate the IP:Leptin samples could explain the high amounts of intronic RNA in this dataset (Figure 21b). Ultimately, the mass spectrometry data suggested that processing the IP:Leptin dataset would lead to an analysis of reads rather associated with unprocessed mRNA than with mRNA bound to ribosomal proteins. It was thus decided that only the gold standard whole tissue RNA-Seq dataset of the Input:Leptin samples generated with the TruSeq sequencing kit will be used for the downstream analyses.

3.2.2 *Effect of weight loss intervention and leptin treatment on hypothalamic gene expression*

Despite similar body weight loss, exendin-4 treated mice significantly reduced their 24h food intake after an acute injection of leptin while mice that were calorie restricted did not respond to the leptin injection (Figure 11). These findings are in line with the previously published superior restoration of leptin sensitivity by pharmacologically induced weight loss [167].

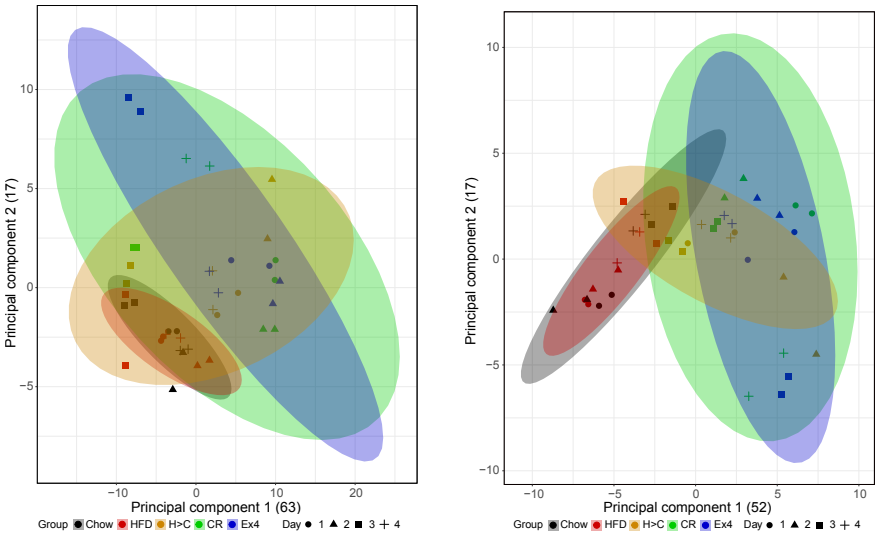
To assess the biological underpinnings of the leptin re-sensitisation by exendin-4 treatment, the hypothalamic gene expression profiles of saline versus leptin treated mice of the two weight loss intervention groups were compared relative to the control cohorts.

3.2.2.1 *Hypothalamic gene expression profiles do not show response to leptin treatment*

The PCA performed to get an overview on the gene expression profiles of all 40 TruSeq samples showed a clear batch effect introduced

by the sampling day (Figure 23a). All samples that had been generated on day 3 formed a cluster on the left side of the graph (Figure 23a). After removing the batch effect the samples split up more according to the administered diet regime. The first principal component, capturing 52% of the variance clearly subdivides the HFD and Chow control groups from the two weight loss intervention cohorts Ex4 and CR (Figure 23b). The diet-switch group (H>C) presents an intermediate phenotype between the Chow and HFD control groups and the two weight loss groups (Figure 23b).

After removing batch effects, samples split up more according to the administered diet regime



(a) Before batch correction

(b) After batch correction

Figure 23: PCA of whole tissue hypothalamus dataset obtained with the Illumina TruSeq kit. **(a)** before and **(b)** after batch correction.

Two samples had been generated for each diet group on each sampling day – one saline and one leptin sample. These two samples were found next to each other in the PCA plot indicating that the gene expression profiles did not differ fundamentally between mice treated with either saline or leptin (Figure 23b).

We compared the two treatment groups within each diet setting to make sure that no small but important leptin-induced changes in hypothalamic gene expression were missed. However, acute injection of leptin only changed the gene profiles of mice on chow diet and of mice that had been calorie restricted compared to their respective saline control group (Table 6).

The G protein-coupled receptor 39 (*Gpr39*) was found to be up-regulated upon leptin treatment in CR mice and down-regulated upon leptin treatment in chow-fed mice, whereas expression levels of granzyme A (*Gzma*), ATP binding cassette subfamily A member 5 (*Abca5*) and suppressor of cytokine signaling 3 (*Socs3*) were changed

Table 6: Number of differentially expressed genes for the comparison of leptin to saline treatment for each diet group. The average expression and the log2 fold changes for the significantly (adjusted p-value < 0.05) differing genes are given in the brackets.

GROUP	TOTAL # OF DE GENES	DE GENES (AVERAGE EXPRESSION, LOG2 FOLD CHANGE)
Chow	4	Gpr39 (0.35, -12.99), Gzma (0.33, -13.49), Socs3 (103.46, 1.16), abca5 (0.40, -15.94)
CR	1	Gpr39 (0.35, 26.77)
H>C	0	
Ex4	0	
HFD	0	

Up-regulation of Socs3 upon leptin treatment in the leptin-sensitive Chow group indicates that the leptin injection per se had been functional

in mice maintained on chow diet (Table 6). The expression levels of *Gpr39*, *Gzma*, and *Abca5* were very low and might not be reliably measurable at all (Table 6). *Socs3* is a known downstream target of leptin that gets activated via STAT3-mediated transcription and plays an important role in the feedback inhibition on the leptin receptor and is a potential mediator of physiologic leptin resistance [24]. The expression of *Socs3* was up-regulated upon leptin treatment in the leptin-sensitive Chow group thus indicating that the leptin injection per se had been functional (Table 6).

However, the absence of more leptin-induced changes in the hypothalamic gene expression patterns points towards a time-dependent effect. The mice had been sacrificed 75 minutes after leptin treatment to capture early gene expression changes in activated neurons. This acute setting might have been too short to induce leptin-mediated gene expression changes, which are usually observed after more than 4h post i.p. leptin injection [6, 57]. Another reason for the absence of a distinct leptin effect could be that mRNA levels of the whole hypothalamus were analysed with the TruSeq method. The hypothalamus is a heterogeneous mix of cells and contains besides neurons, astrocytes, microglia, endothelial cells, and tanycytes [39, 77]. Even if the acute injection of leptin changed the gene expression profiles in leptin-responsive neurons, this effect might have been masked by the unchanged or adverse mRNA levels in the other cell types of the hypothalamic samples [76].

The absence of a leptin effect and the high similarity between the two treatment groups enabled us to combine the leptin and saline samples for each diet. The statistical power of the small dataset was thus increased as now eight replicates per diet group were compared instead of four.

Leptin and saline datasets were combined for each diet group due to absent leptin effect

3.2.2.2 *Pairwise comparisons of intervention groups identified strong deregulation between diet interventions*

To investigate the effect of the administered diets and weight loss interventions on the hypothalamic mRNA levels, we compared the expression profiles of each diet group to the expression profiles of the other diet groups in a pairwise manner using DESeq2. The numbers of differentially expressed genes resembled the image drawn by the PCA: Few genes were altered between the two weight stable control groups Chow and HFD (Table A6). At the same time few genes differed when comparing the three weight-loss groups H>C, Ex4 and CR with each other (Table A6). This quite similar expression pattern between the intervention groups was challenging, as there was no clear difference between the groups.

The numbers of differentially expressed genes resembled the image drawn by the PCA: Few genes were altered between Chow and HFD and between Ex4 and CR

One possible explanation for the high similarity in the hypothalamic expression pattern of chow- and HFD-fed mice is that there is no acute stimulus that would change the gene expression in the hypothalamus. The HFD induced body weight gain had been accomplished by the start of the study. The mice were fed ad libitum and had a stable body weight over the complete study period (Figure 10). In contrast, the weight loss intervention groups CR and Ex4 underwent a massive change in body weight (Figure 10). However, as the loss in body weight between CR and Ex4 treated mice was similar, hypothalamic gene expression was apparently changed likewise. Along this line, more genes were changed between CR or Ex4 compared to the weight stable control groups Chow (1,728 vs. 2,142 DE genes, respectively) and HFD (2,691 vs. 3,079 DE genes, respectively) than between CR or Ex4 and the diet-switch group H>C that also lost some body weight (119 vs. 230 DE genes, respectively) (Table A6).

We took a closer look at the differentially expressed genes between diet groups. For example, one gene that was found to be increased in the HFD-fed mice, when compared to our Chow cohort was Serpin Family A Member 3 (*SerpinA3*) (Figure A2). This increase of *SerpinA3* expression had been reported before by Sergi et al. [205], who proposed that *SerpinA3* expression is involved in hypothalamic inflammation mediated by nutrition. Another gene that showed a higher expression in our HFD-fed cohort compared to the Chow mice was calcitonin-related polypeptide α (*Calca*) (Figure A2). This gene can be alternatively spliced into calcitonin and αCGRP (α Calcitonin gene-

related peptide) in a tissue-specific manner [241]. The expression of αCGRP had been reported to be distributed across the CNS, including the hypothalamus. When Walker et al. [241] prevented the production of αCGRP in their mouse model, they observed that the mice were phenotypically unchanged, but upon HFD challenge they were protected from diet-induced obesity, suggesting that αCGRP is involved in metabolic regulation and lipid management [241].

In the following, we concentrated on the group comparisons that returned small numbers of altered genes but were of high scientific relevance such as the comparisons of HFD to Chow and Ex4 to CR.

3.2.2.3 *Neuropeptides mirror administered diet*

When the hypothalamic gene expression profiles of the weight stable diet groups HFD and Chow were compared, ten genes were found to be altered (Figure A2, Table A6). Three of the significantly regulated mRNAs belonged to the group of neuropeptides that are important regulators in the control of energy expenditure and food intake: Expression levels of the orexigenic mediators *Npy* and *Agrp* were down-regulated in HFD compared to Chow, while the anorexic neuropeptide *Cartpt* (Cart prepropeptide) was up-regulated (Figure A2). These findings are in line with previous work showing a decrease in genes promoting food intake whereas genes inhibiting food intake were increased in diet-induced mice [99].

When the gene expression profiles of all neuropeptides present in our dataset were compared across all five groups, mRNA levels of *Npy* and *Agrp* were found to be similarly expressed in obese (HFD, H>C) mice compared to mice that had lost 30% of their body weight (CR, Ex4) (Figure 24).

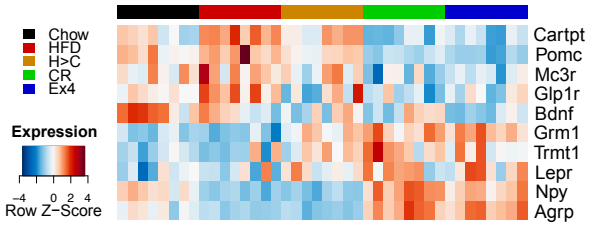


Figure 24: Neuropeptide expression in hypothalami across all diet groups.

Expression of *Npy* and *Agrp* was significantly up-regulated in the weight loss intervention groups CR and Ex4 when compared to the weight stable Chow and HFD groups as well as to the H>C group that had only lost 10% body weight due to the switch to less calorie dense food (Figure 24, Table A7). In contrast to the three control groups that were not restricted in their food intake, Ex4 treatment induced a significant reduction in food intake to which the food intake

of the CR group was matched. The up-regulation of the key orexigenic neuropeptides thus reflected the fasting state of the weight loss intervention groups. This fasting-induced up-regulation of *Npy* and *Agrp* mRNA levels that are co-expressed in the hypothalamus had been shown before [90].

The effect of weight loss interventions was not only reflected in body weight phenotype but also on hypothalamic mRNA levels

In contrast, mRNA levels of the anorexigenic neuropeptide *Pomc* were significantly down-regulated in Ex4 when compared to the HFD and H>C group (Figure 24, Table A7). *Pomc* expression was similarly regulated in CR mice but only significantly down-regulated when compared to HFD (Table A7). Also the second food-intake stimulating neuropeptide *Cartpt*, that was up-regulated in HFD compared to Chow (log2 FC: 0.32, adj. p-value: 0.047), was significantly down-regulated in the weight loss groups CR and Ex4 when compared to H>C (Figure 24, Table A7). As *Pomc* and *Cartpt* are antagonists of *Npy* and *Agrp*, and were shown to decrease their mRNA levels upon fasting [90], the observed findings match the applied food restriction.

Taken together, the neuropeptide mRNA expression patterns showed that the effect of the weight loss interventions was not only reflected in the body weight phenotype but also on the hypothalamic mRNA levels.

3.2.2.4 *Fasting induced gene expression changes*

The mice of both weight loss intervention groups lost the same amount of body weight and fat mass during the study period of ten days. But in contrast to calorie-restricted mice, the exendin-4 treated mice regained leptin sensitivity since they decreased their food intake upon leptin injection. Therefore, the comparison of the hypothalamic gene expression changes between the CR and the Ex4 group will allow us to assess possible targets for weight loss interventions and leptin resensitisation.

8 of the 29 genes differing between CR and Ex4 are induced by fasting in the hypothalamus

When the mRNA profiles of the CR group were compared with the Ex4 group, 29 genes were found to be differentially expressed (Figure 25, Table A8). The majority of these differentially expressed genes showed an up-regulation of the mRNA levels in the CR group compared to the Ex4 group. Interestingly, eight of the 29 altered genes are known to be induced in the hypothalamus by fasting (Figure 25 highlighted in red) and will be discussed in detail in the following.

Among the eight fasting induced genes that were found to be up-regulated in the hypothalamus of CR mice compared to Ex4 treated mice were *Plin4* (Perilipin 4), also referred to as *S3-12*, and *Pnpla2* (Patatin-like phospholipase domain containing 2). *Plin4* belongs to the PAT family of lipid storage droplet proteins and plays a key role in intracellular lipid trafficking [183, 190]. *Plin4* was significantly up-

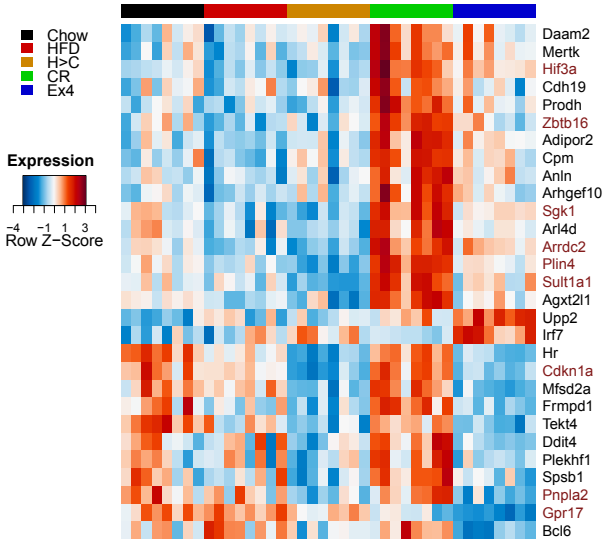


Figure 25: Expression pattern of altered genes between Ex4 and CR. Genes induced by fasting in the hypothalamus are highlighted.

regulated in CR mice compared to all other control groups (Table A8). In Ex4 treated mice, *Plin4* expression was significantly lower than in CR mice but up-regulated compared to the diet-switch control H>C, indicating a weight loss mediated effect that was pronounced by calorie restriction.

The triglyceride lipase *Pnpla2* hydrolyses monoacyl glycerides and glycerol from triglycerides in preparation for beta-oxidation [183]. *Pnpla2* was not altered by exendin-4 treatment when compared to the three control groups. CR, on the other hand, did up-regulate the gene when compared to the three control groups even if the fold changes where rather modest (Table A8).

Both genes are known to increase the utilisation of lipid oxidation instead of glycolysis for energy production [183] and were shown to be induced by fasting in the hypothalamus of mice by Poplawski et al. [183]. Data from human studies further indicated an important role for *Plin4* in the regulation of metabolism as several SNPs assigned to *Plin4* were shown to be associated with obesity related phenotypes [190]. The gene expression changes by weight loss due to calorie restriction therefore indicate a metabolic reprogramming away from glycolysis towards lipid oxidation that is not present when the body weight loss was pharmacologically induced. The metabolic switch from glucose metabolism to lipid oxidation in calorie-restricted mice is further supported by the finding that *Hif3a* (Hypoxia-inducible factor 3, alpha subunit), a transcription factor that has been brought in connection with reduced glycolysis [183], was also up-regulated in

CR mice compared to exendin-4 treated and the diet-switch mice.

Interestingly, two of the eight fasting induced genes that were up-regulated in the CR mice compared to exendin-4 treated mice had previously been linked to the appetite regulating neuropeptide orexin. Orexin, also known as hypocretin, is predominantly produced in the lateral hypothalamus and was shown to regulate arousal, wakefulness, and appetite [15]. Stimulation with orexin increases the craving for food and the meal size by suppressing inhibitory postingestive feedback [15].

The increased mRNA levels of the cellular stress responsive kinase *Sgk1* (serum- and glucocorticoid-induced protein kinase-1) in CR mice was in line with findings from Nonogaki, Ohashi-Nozue, and Oka [175], who showed that *Sgk1* expression is induced upon fasting in the hypothalamus and proposed that hypothalamic *Sgk1* is involved in energy homeostasis in mice. Deng et al. [56] studied the role of *Sgk1* specifically in hypothalamic POMC neurons and saw that *Sgk1* co-localises with POMC neurons. When knocking out *Sgk1* in POMC neurons, mice showed an obese phenotype and a decreased energy expenditure [56]. The opposite phenotype was observed when they over-expressed *Sgk1* in POMC neurons [56]. The work of Koesema and Kodadek [133] then proposed an involvement of orexin receptor signaling in the *Sgk1* mediated regulation of energy homeostasis. The hypothalamic expression of *Sgk1* in exendin-4 treated mice was higher compared to the diet-switch control but unchanged compared to the weight stable control indicating an ameliorated orexin effect in pharmacologically induced weight loss.

The second up-regulated gene in our CR mice that Koesema and Kodadek [133] observed to be highly regulated by orexin was *Zbtb16* (Zinc finger and BTB domain-containing protein 16). *Zbtb16* is a zinc finger transcription factor located in the nucleus that is involved in cell cycle progression, and interacts with a histone deacetylase [224]. *Zbtb16* was shown to be up-regulated upon fasting in AgRP neurons [101]. A loss of function mutation in *Zbtb16* can cause human neurological disorders [26]. Additionally, the reduced expression might contribute to phenotypic properties of the Prader-Willi Syndrome, such as obesity and neurodevelopmental defects [26]. In contrast to the significant up-regulation in the CR group, the expression of *Zbtb16* was unaltered in the Ex4 group compared to all other groups (Table A8). This points towards an involvement of orexin receptor signaling specifically in caloric restriction mediated weight loss.

Another master regulator of food intake is the transcription factor *FoxO1* (Forkhead Box O1 Protein) which is a shared element of path-

ways regulated by the anorexigenic hormones leptin and insulin to regulate food intake and energy efficiency [127, 131, 182]. *FoxO1* promotes opposite patterns of coactivator-corepressor exchange at the POMC and AgRP promoters, resulting in activation of AgRP and inhibition of POMC [131]. Accordingly, induction of the transcription factor *FoxO1* leads to an increase in food intake and as a consequence also to an increase in body weight [127]. In contrast, mice with a knock-out of *FoxO1* specifically in AgRP neurons showed reduced food intake [187]. Along this line of evidence, our CR mice showed a modest up-regulation of *FoxO1* when compared to the three control groups (data not shown) indicating an activation of *FoxO1*. There was no difference in *FoxO1* expression between CR and Ex4 treated mice. However, two targets of *FoxO1*, *Cdkn1a* (Cyclin-dependent kinase inhibitor 1A) and *Gpr17* (G-protein-coupled receptor 17), were significantly altered between CR and Ex4 treated mice (Table A8).

The cyclin-dependent kinase inhibitor *Cdkn1a*, also referred to as *p21*, was shown before to be induced by fasting in the hypothalamus [87, 183, 226, 230]. It had been suggested that *FoxO1*, but not *p53*, contributes to the robust induction of *p21* expression in fasted mice [230]. Indeed, hypothalamic *Cdkn1a* expression was significantly increased in CR compared to Ex4 treated mice and to the obese control groups HFD and H>C. This finding points towards a *FoxO1* dependent effect in calorie restriction since *FoxO1* expression was only up-regulated in CR mice and both the expression of *Cdkn1a* and of *FoxO1* was unaltered between Ex4 and the control groups.

Gpr17 is the Uracil nucleotide/cysteinyl leukotriene receptor that was identified as *FoxO1* target by Ren et al. [187]. AgRP-specific loss of *Gpr17* was shown to lead to a reduction in feeding and improved metabolic status [188]. In contrast to *Cdkn1a*, no change in the hypothalamic expression of *Gpr17* was observed in the CR mice compared to the control groups. But a significant down-regulation of hypothalamic *Gpr17* expression was seen for Ex4 treated mice compared to all other groups.

Taken together, we did not observe a clear *FoxO1* mediated effect in our weight loss intervention due to the observation of a differential regulation pattern for the two *FoxO1* target genes *Cdkn1a* and *Gpr17* between CR and Ex4 treated mice.

The remaining two fasting induced genes in the hypothalamic samples of CR mice were *Sult1a1* (sulfotransferase family member 1A) and *Arrdc2* (arrestin domain containing 2).

The hormonal regulator *Sult1a1* was shown to be up-regulated by calorie restriction in several tissues including the hypothalamus [226]. However, the expression pattern of *Sult1a1* in response to HFD feeding in peripheral tissues was contradictory with increased *Sult1a1* mRNA levels in liver but down-regulation in adipose tissue of HFD-fed rats [88]. In our hypothalamic data set no difference between chow- and HFD-fed mice was observed. However *Sult1a1* expression was up-regulated in both weight loss intervention groups compared to the diet-switch control pointing to an effect specific to body weight loss that was more pronounced in CR.

Not much is known yet about the expression regulation of *Arrdc2* that has been associated with signal transduction [82]. Guarnieri et al. [87] showed an up-regulation of *Arrdc2* upon food restriction in the nucleus accumbens but not in the hypothalamus. In our hypothalamus dataset, *Arrdc2* expression was increased in both weight loss groups compared to the H>C control group. Its expression was also increased in CR when compared to HFD. This might indicate a role of *Arrdc2* in the regulation of body weight loss.

Overall, the described changes in the hypothalamic gene expression profile indicate that calorie-restricted mice undergo fasting induced changes in fuel utilisation, orexin receptor signaling and transcription factor regulation. Here, treatment with exendin-4 seems to ameliorate the fasting induced up-regulation of hypothalamic genes despite similar body weight loss. We hypothesise that this attenuation of the fasting-dependent hypometabolism in CR mice might contribute to the superior restoration of leptin sensitivity in exendin-4 treated mice.

3.2.2.5 *Expression of fasting induced genes in AgRP and POMC neurons*

It has been shown in numerous studies that different neuronal populations are involved in the hypothalamic control of energy expenditure and food intake, as reviewed by Williams et al. [251]. As whole hypothalamus samples were investigated in our dataset that contain various neuronal populations and other cell types such as glial cells and tanycytes, cell type specific gene expression patterns could be masked. In order to investigate whether the 29 genes differentially expressed in the weight loss intervention groups play a role in two of the most important neuronal populations with respect to food intake, the AgRP and POMC neurons (Section 1.3.2), we studied their expression in the dataset of Henry et al. [101] who produced a valuable resource to compare gene expression patterns in AgRP and POMC neurons in the fasted and fed state.
The AgRP and POMC data was analysed separately using DESeq2 to compare fasted and fed mice and to compare the expression changes

observed in the two neuronal populations to the gene expression alterations observed in our whole tissue data set.

7 genes differing between Ex4 and CR are also altered upon fasting in either AgRP or POMC neurons

We could observe that five (*Sgk1*, *Cdkn1a*, *Zbtb16*, *Spsb1*, *Ddit4*) of our fasting genes were significantly differentially expressed in the AgRP neurons upon fasting (Figure 26a, Table A9). All of them showed a higher expression in the fasted state, which is in line with our findings.

Further, an increased expression of *Zbtb16* and *Bcl6* (B-cell CLL/-lymphoma 6) was detected in POMC neurons of fasted mice (Figure 26b, Table A10), which is in line with our data from whole hypothalamus samples. In contrast, *Hr* (Hairless homolog) was down-regulated in POMC neurons (Figure 26b, Table A10), while it was up-regulated in our whole tissue samples.

(a) AgRP neurons (b) POMC neurons

Figure 26: Fasting genes differentially expressed between fasting and feeding in **(a)** AgRP and **(b)** POMC neurons.

Overall, the similar expression pattern of the fasting induced genes in isolated AgRP and POMC neurons from fasted mice indicates, that our analysis of the whole hypothalamus gene expression profile is a good starting point to define genes of interest in the context of fasting. However, further experiments are needed to understand the neuron-specific effects of the weight loss interventions.

3.3 ESTIMATING GENETIC IMPACT ON METABOLIC NETWORKS

The integration of metabolite datasets with transcriptomics and proteomics will help us to gain a more holistic view of molecular underpinnings of physiological and pathopysiological cellular signaling. Furthermore, metabolomics could be used to identify biomarkers and new targets for the treatment of metabolic diseases, such as diabetes and cancer. However, current system-biological approaches integrating metabolomics and other omics datasets often lack user friendliness, output interpretability or are based on prior knowledge on pathways and enzymatic reactions. As a result, the genetic control of metabolic pathways, their dynamics, and dependence on different diets and diseases remains rather unknown. Overall, this complicates the detection of new genes or enzymes, which influence those metabolites that are important for specific pathways or disease states.

In this last part of the thesis, we present a new generic integration method that, here, was applied to integrate metabolomics and transcriptomics data based on pairwise correlation analysis of metabolites coupled to partial correlation combining the metabolite correlations with the transcript expression profiles followed by the construction of undirected, weighted graphs. The idea behind this approach is to link changes in metabolite concentrations to differentially expressed genes between dietary groups or treatment conditions thereby helping to develop hypotheses for genes that influence the observed metabolite network and might therefore play a role for the development of the observed phenotype. In contrast to many other approaches, our method is fully unsupervised and independent from additional prior knowledge. Its straightforward but flexible setting allows for multiple applications, yet the produced networks may convey complex interactions but are easy to interpret.

This work is the result of a collaboration with two research units at the Helmholtz Centre Munich, namely the Research Unit Neurobiology of Diabetes and the Research Unit Molecular Endocrinology and Metabolism. The collaboration partners performed all mouse and wet lab experiments. Excerpts and figures presented here were submitted (Klaus VS*, Schriever SC* et al; *Equal contribution).

Ex4 mice showed massive but transient accumulation of triglycerides in the liver

Pharmacological-induced weight loss by exendin-4 treatment caused a massive accumulation of triglycerides in the livers of mice after ten days of treatment. This acute hepatic steatosis turned out to be a transient phenomenon since the livers of mice treated with exendin-4 for 30 days were clear of triglycerides when compared to the HFD control group (data not shown). To investigate the molecular underpinnings of the acute triglyceride accumulation, hepatic gene expression pro-

files and metabolite levels were measured in mice after ten days of treatment. The metabolome is the functional readout of cellular processes taking place at a defined time-point. These processes are regulated by genes but they can in turn also be regulated by the metabolite levels itself exemplifying one of various complex feed-back systems in biology (Figure 7) [30, 33, 173]. To assess the genetic control of the metabolite networks under each dietary condition, we developed a new data integration approach.

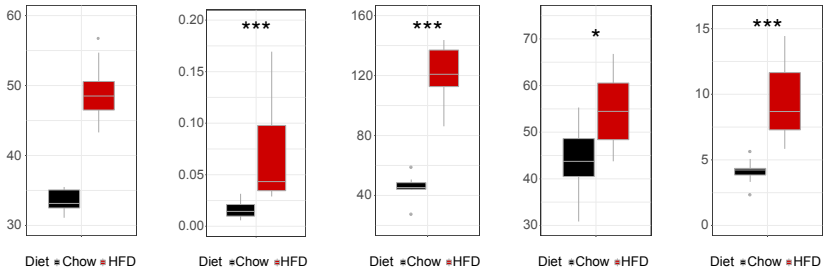
3.3.1 *Method application on Chow and HFD dataset*

For testing the method, the two weight stable control groups, Chow and HFD, were used as the comparison of lean versus diet-induced obese mice is a well-studied setting.

3.3.1.1 *Genetic and metabolic profiling of chow- and HFD-fed mice*

HFD cohort shows typical signs of diet induced obesity

HFD feeding for 22 weeks resulted in the obese phenotype characterised by significant metabolic alterations compared to the lean chow-fed control mice (Table A12, Figure 27).



(a) Body weight [g] (b) Plasma insulin [mg/dl] (c) Plasma cholesterol [mg/dl] (d) Plasma TAG [mg/dl] (e) Liver TAG [µg/mg]

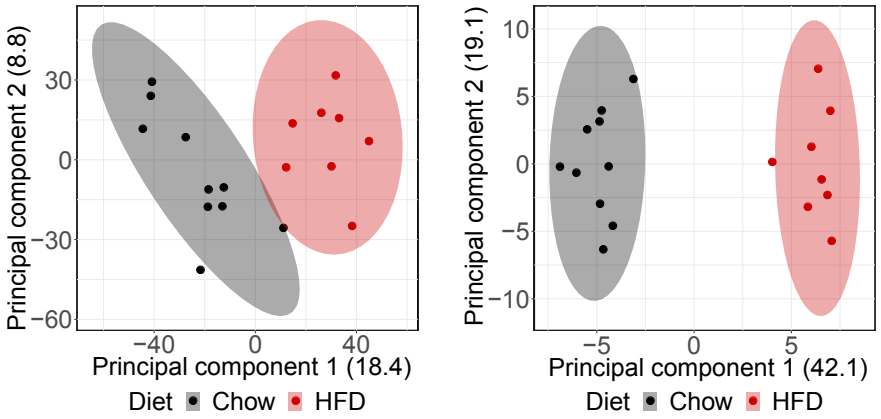
Figure 27: Phenotype monitoring of chow- and HFD-fed mice: **(a)** Body weight measured at the end of the study, **(b)** plasma insulin levels, **(c)** plasma cholesterol levels, **(d)** plasma and **(e)** liver triacylglyceride (TAG) levels. $*$: p <= 0.05; $**$: p <= 0.01; $***$: p <= 0.001

Except of two non-responders with body weights of 33.3g and 37.5g that were excluded from further analyses, the HFD-fed mice had a significantly increased body weight at the end of the study period and elevated plasma insulin levels compared to the Chow controls (Figures 27a and 27b, Table A12), which is in line with the obese phenotype previously reported for the C57BL/6J mouse model [92, 116, 207, 214]. The observed elevated plasma cholesterol and triglyceride levels in the obese mice (Figures 27c and 27d, Table A12) had also been reported before [116, 126]. The analysis of the liver triglyceride concentrations revealed a hepatic steatosis in the HFD-fed mice

(Figure 27e, Table A12) confirming already published data for DIO C57BL/6J mice [119, 207]. Taken together, the two diet groups showed all metabolic perturbations expected after HFD exposure and are thus a good model to study obesity-induced changes in the metabolome and transcriptome.

Diet was main influencing factor responsible for observed variance in the gene and metabolite profiles

The initial analysis of the gene expression profile and the metabolite profile of the chow- and HFD-fed mice by PCAs revealed that the diet was in both cases the main influencing factor responsible for the observed variance in the gene and metabolite abundance profiles (Figure 28). The first principle component explained 18.4% of the variance observed in the gene expression profiles and 42.1% of the variation encountered in the metabolite concentration profiles.

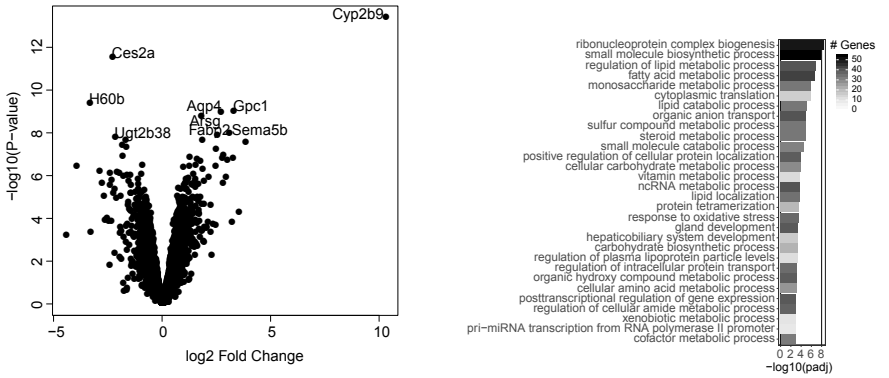


(a) Gene expression profiling (b) Metabolite profiling

Figure 28: PCA showing the main contributor to differences observed in **(a)** gene expression and **(b)** metabolite concentration profiles.

Next, the other two factors – body weight measured at the end of the study (BWE) and liver triacylglyceride levels (TAG) – were also taken into account as possible influencers of the gene expression changes and metabolite levels. BWE and hepatic TAG correlated (ρ=0.86) with each other and were induced by the administered diet. Thus, a relative importance analysis was performed to assess the main influencing factor for each gene and each metabolite in our dataset. The analysis showed that the administered diet explained indeed most of the variance observed in the expression of all genes and metabolites in the dataset (Figure A3, Tables A13 and A14).

The differential gene expression analysis comparing the hepatic gene expression profiles of HFD-fed mice to chow-fed controls identified 989 genes that were significantly deregulated in the livers of obese mice (Figure 29a, Table A15). Of the differentially expressed genes, 501 genes were up- and 488 genes were down-regulated. Among the top 50 significantly differentially expressed genes in liver between



(a) Differentially expressed genes (b) GO term enrichment

Figure 29: Differential gene expression analysis comparing the liver gene expression profiles of HFD-fed mice to chow-fed mice. **(a)** Volcano plot of gene expression alterations. **(b)** Top 30 GO terms enriched by differentially expressed genes. The BH adjusted p-value is shown on the x-axis and the number of genes involved in the respective biological process is indicated by the grayscale.

HFD and Chow were known HFD-regulated genes described in the following.

Cyp2b9 (Cytochrome P450s (Cyp) 2B9) was one of the top up-regulated genes. It belongs to the Cyp2b family members primarily expressed in the liver and is part of the hepatic detoxication system that might be influenced by diets containing unsaturated fatty acids [70]. *Tsc22d1* (Transforming growth factor-β 1 stimulated clone-22 D1), which is suggested to be required for the basic high-density lipoprotein (HDL) cholesterol maintenance in the livers of obese mice [113], was also up-regulated upon HFD feeding in our dataset. Vanin-1 (*Vnn1*) showed a higher expression pattern in the livers of our HFD mice. *Vnn1* is an oxidative stress sensor enriched in the liver regulating several metabolic pathways including hepatic gluconeogenesis [40]. *Ces2a* (Carboxylesterase 2A), for which Ruby et al. [197] showed that it controls a hepatic lipid network, which is dysregulated in human and mouse obesity, showed a significantly lower expression upon HFD feeding. This is in line with the observed decreased hepatic activity of *Ces2a* in obese rodents and humans [197]. This initial analysis of the differentially regulated genes revealed that many genes altered by HFD feeding were involved in lipid metabolism.

Many genes altered by HFD feeding were involved in lipid metabolism

The performed GO term overrepresentation analysis of the differentially expressed genes did support this observation as lipid metabolic pathways were highly significantly regulated and overrepresented in the variety of the metabolic processes deregulated by HFD feeding

(Figure 29b, Table A16).

The differential metabolome analysis comparing the liver metabolite profiles of HFD-fed mice to chow-fed mice identified 91 metabolites to be altered upon HFD feeding. Two thirds of the metabolites (63 of the 91) with changed concentrations by HFD feeding belonged to one of the three measured phospholipid classes: Phosphatidylcholines (PC), Sphingomyelins (SM), and Lyso-Phosphatidylcholines (LPC) (Figure 30, Table A17).

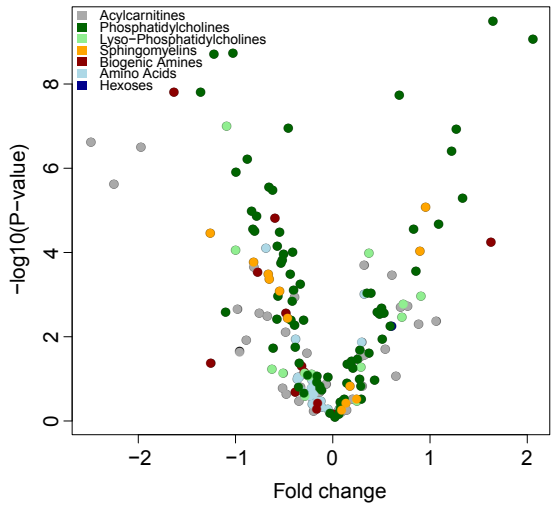


Figure 30: Volcano plot of deregulated liver metabolites upon HFD feeding.

Phospholipids play an important role in the communication between extra- and intracellular space since they are precursors of signaling molecules [46, 52]. Free fatty acids, which are taken up from the liver, are converted into these phospholipids, which are structural components of cell membranes of very low density lipoproteins (VLDL), bile particles, and the surface layer of lipid droplets [198].

Despite higher hepatic triglyceride levels and increased plasma triglyceride concentrations, no clear pattern of regulation was observed for the deregulated phospholipid levels in the livers of HFD-fed mice compared to the Chow control group. It had been shown before that decreased PC levels in the liver lead to an accumulation of intracellular TAG as a result of impaired VLDL secretion due to impairment in PC biosynthesis [46]. On the other hand, in the state of hepatosteatosis and steatohepatitis in mice, however, PC levels are decreased compared to histologically healthy livers [44]. In our dataset, some PCs were decreased, which was in line with the increased liver TAG levels and the observed hepatosteatosis. However, for other PCs increased concentrations were found that could not be explained by

the liver phenotype.

Twenty percent of the metabolites (18 of the 91) with altered concentrations by HFD feeding belonged to the class of acylcarnitines. Acylcarnitines are intermediates of amino acids and fatty acid oxidation [80]. In our dataset, the short-chain acylcarnitines propionyl-carnitine (C3) and valeryl-carnitine (C5) were up-regulated (Table A17). C3 is an intermediate of branched-chain amino acid degradation, while the metabolic origin of C5 is unknown [80]. Giesbertz et al. [80] reported an increase in plasma and hepatic C3 and C5 levels in streptozotocin-induced insulin-deficient mice, which serve as a model of type I diabetes. Plasma branched-chain acylcarnitine levels were further shown to be associated with type 2 diabetes [171], but it is not known if this also applies to their concentration levels in liver. While an increase in long-chain acylcarnitine accumulation in the liver had been linked to HFD-induced NAFLD [242], we did only observe an increase of C12. The levels of most other long-chain acylcarnitines, namely C14:1, C14:1-OH, C14:2, C14:2-OH, C16:1-OH, C16:2, C16:2-OH and C18:2 were significantly decreased in the livers of our HFD-fed mice compared to the Chow controls (Table A17). The demonstrated increase in long-chain acylcarnitines often goes together with an increased expression of cyclooxygenase-2 (*Ptgs2*), a deregulation of long-chain acyl coenzyme A dehydrogenase (*Acadl*), or sirtuin 3 (*Sirt3*) [23, 105, 106, 157, 199, 238]. We did not observe a deregulation of any of these genes induced by HFD feeding in our cohort.

HFD increases the correlation between liver metabolites and clinical blood parameters

In a last step, a linear regression analysis was performed to test if liver metabolites can be used to predict a metabolic phenotype under different dietary conditions. We identified multiple metabolites that correlated with either plasma cholesterol, insulin, or TAG levels (Table A18). In the Chow dataset, only two metabolites could predict either the insulin or the cholesterol level. In contrast, in the HFD dataset three metabolites were observed that correlated with TAG and 17 metabolites that were associated with insulin levels. From the insulin associated metabolites all PC, LPC and AA showed a negative correlation, whereas the AC were all positively correlated. This indicated that the metabolic alterations induced by HFD feeding increase the correlation between liver metabolites and clinical blood parameters.

Taken together, the described changes in hepatic metabolite concentrations indicate that the changes in the gene expression profiles of HFD-fed mice compared to Chow controls were also manifested on the metabolite level.

3.3.1.2 *Correlation analysis of metabolites revealed diet-dependent changes in the hepatic metabolome*

To investigate if and to what extent the interaction between liver metabolites was altered by the respective diet, we calculated for each metabolite pair the Pearson correlation coefficient for the datasets Chow and HFD separately (Figure 31). By allowing a non-adjusted p-value<0.05 for the correlation coefficient significance, we selected 2,488 metabolite pairs for Chow and 2,322 metabolite pairs for HFD. Less than 50% of these metabolite pairs (n=923) correlated significantly in both sets, indicating that the administered diet substantially alters the hepatic metabolism. This finding is in line with previous studies showing that an oversupply of nutrients alters the hepatic metabolism not only in rodents but also in dogs and humans [134].

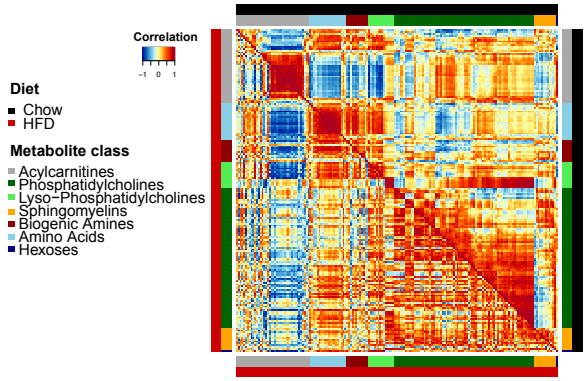


Figure 31: Pairwise correlation matrix. The upper triangle shows the correlation coefficients obtained in the Chow set, the lower triangle shows the coefficients for the HFD set.

When the metabolite classes of the 923 significantly correlating metabolite pairs were analysed, we found that mainly the metabolite class of amino acids was correlated in both diet sets, indicating that amino acids interact independent of dietary conditions (Figure A4).

For metabolite pairs assigned to other metabolite classes we observed substantial changes within their correlations depending on the diet (Figure A4). AC and PC metabolite combinations showed the highest absolute number of changes, with a correlation pattern switching from positive to negative (Figures A4 and 31). For the pairs SM – BA, SM – AA, SM – H, AA – PC, and PC – H no metabolite pair correlated in either diet set.

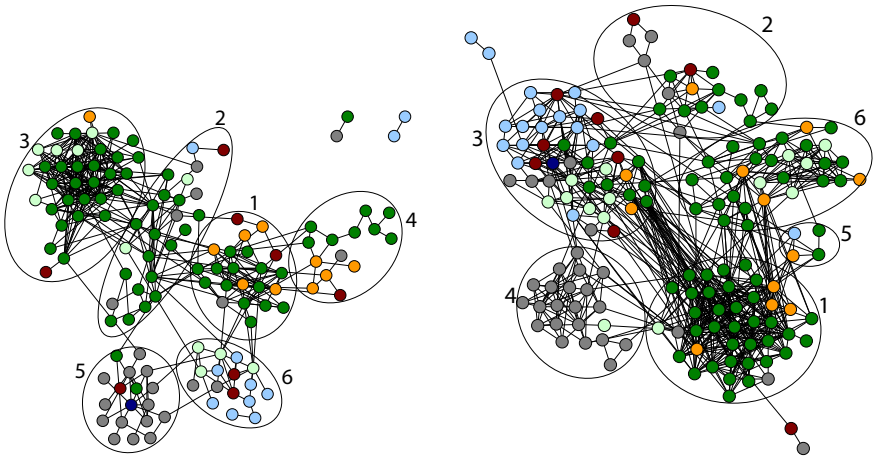
Taken together, the Pearson correlation coefficient analysis revealed diet-dependent changes in the metabolite correlations with the exception of amino acids. This was also observed by others. Dyar et al.

[65] previously reported diet-dependent changes in metabolite correlations in a circadian manner over multiple tissues.

3.3.1.3 *Chow and HFD networks obtained with CoNI reflected the metabolic phenotype*

To assess the genetic impact on the metabolite network formed under each dietary condition, we integrated the gene expression profiles and the metabolite concentration profiles with our newly developed approach called Correlation-based Network Integration (CoNI) (Section 2.7.1).

As first result, we obtained two independent networks for Chow and HFD (Figure 32). The networks were constructed of 485 metabolite pair – gene triplets for the Chow and 1,058 metabolite pair – gene triplets for the HFD set. From the 181 measured metabolites 133 were connected within the Chow network and 164 metabolites within the HDF network, respectively. Of these metabolites that could be connected 127 were found in both networks. When the resulting metabolite pairs were compared between the Chow and HFD dataset, 67 metabolite pairs were found to be connected in both diet networks. In the Chow network, 340 metabolite pairs were exclusively connected, whereas twice as much metabolite pairs (n=655) were connected specifically in the HFD network.

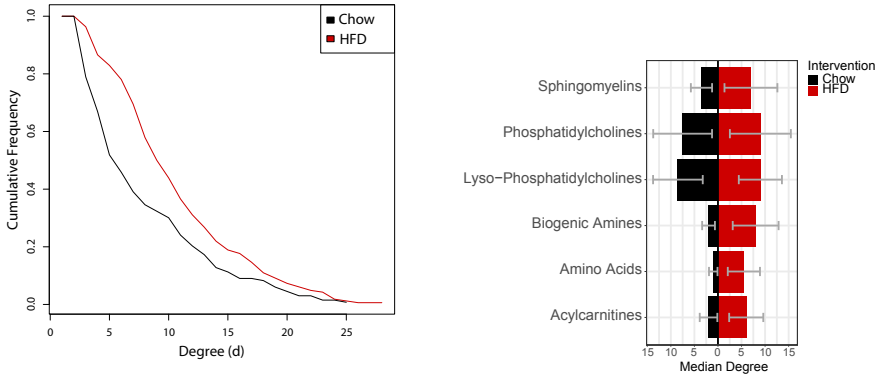


(a) Chow Network (b) HFD Network

Figure 32: Networks obtained with CoNI for **(a)** Chow and **(b)** HFD. Nodes represent metabolites and are coloured according to their metabolite class. Edges are build from genes that influence the respective metabolite pair. The numbered ellipses show the communities obtained for the respective network.

When the node degree distributions of both networks were compared, a trend towards higher degrees was observed for the HFD net-

work. However, both graphs followed a power law distribution (Figure 33a). The comparison of the node degrees for the specific metabolite classes showed that the HFD network had higher degrees for all metabolite classes, which reflects the overall higher number of edges in the HFD network (Figure 33b). In contrast to the elevated but similar distribution of node degrees within the HFD network, the Chow network shows a trend towards increased node degrees for Lyso- and Phosphatidylcholines compared to the other metabolite classes.



(a) Degree distribution

(b) Node degrees per class

Figure 33: Node degree comparison. **(a)** Degree distribution and **(b)** node degrees shown for each metabolite class of the Chow and HFD networks.

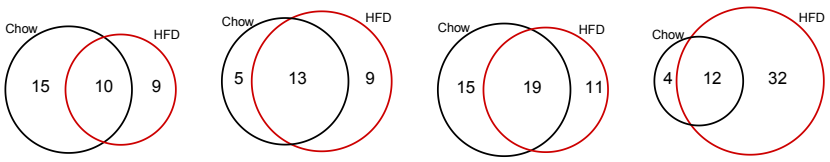
Inferred networks tended to be organised in densely connected sub-networks

A more detailed analysis of the inferred networks revealed that the networks were organised in densely connected sub-networks, also called communities. When the fast greedy modularity optimisation algorithm [41] was applied to the networks, we identified six network communities each in the Chow and HFD graph (Figure 32).

To identify communities with higher diet-dependent regulation the overlap of metabolites within the communities between the Chow and HFD graphs were compared (Figure 34).

For both diets the PCs split up in four communities, whereas the other two communities were mainly characterised by either AC or by AA (Figure A5). Whereas the latter communities (AC and AA) appear to be stable, the PC communities seem to be more reorganised upon the change to HFD feeding (Figure 34, Table A19). One possible explanation could be that amino acid interactions are more conserved, as they are involved in essential cell processes.

Next, we had a closer look at the genes connecting the metabolite pairs. In contrast to the metabolites, the genes in both networks



(a) 2 – 2 (b) 5 – 4 (c) 3 – 6 (d) 6 – 3

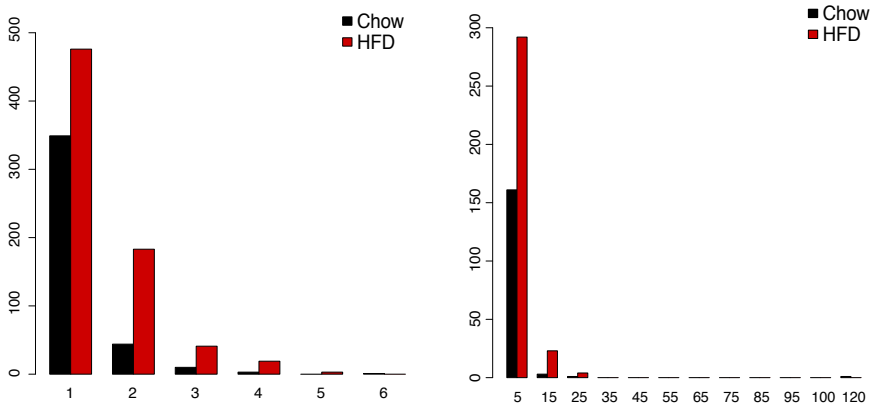
Figure 34: Node comparison between community **(a)** 2 of the Chow network and community 2 of the HFD network. According node comparisons of community **(b)** 5 of the Chow and 4 of the HFD, **(c)** 3 of the Chow and 6 of the HFD, **(d)** and 6 of the Chow and 3 of the HFD networks.

showed greater heterogeneity. Only five genes were present in the edges of both the Chow and the HFD network (*Gm4553, Hnrnpm, Tap1, Xpo7, Eya3*). Of these genes present in both networks, only *Tap1* was differentially expressed between HFD-fed and chow-fed mice. That Tap1 is present in both networks and is differentially expressed is striking since Tap1 was recently found to be a key peripheral contributor to hepatic lipid deposition and the development of diet-induced NAFLD in mice [12]. Within the Chow network, 166 individual genes were identified, of which 15 were differentially expressed between the two diet groups Chow and HFD (Figure A6). In contrast, twice as many individual genes (n=319) were found in the HFD network, of which 33 were differentially expressed in HFD compared to Chow (Figure A7). This, together with the increase in network density, strongly indicates the massive change in metabolite regulation that is caused by dietary stress and a developing hepatic steatosis.

When the number of genes that map to a single edge was counted, we observed that most edges consisted of a single gene (85.75% in Chow, 65.93% in HFD). The maximum number of genes per edge was six genes in Chow and five genes in HFD (Figure 35a). The distribution of genes over the edges (Figure 35b) showed that less genes control more metabolite pairs in the Chow network compared to the HFD network. In contrast, the genes in the HFD network were less distributed. The highest distribution in both networks was found for *Nop16* (NOP16 nucleolar protein), appearing in 113 of 407 edges in Chow and *Cobll1* (Cobl-like 1), appearing in 25 of 722 edges in HFD.

Glycerolipid metabolism and NAFLD were enriched for the HFD network genes

To further classify the genes found in both networks a functional enrichment analysis using KEGG pathways and GO biological processes was performed. For the Chow network we did not identify informative categories for the individual genes. However, the individual genes of the HFD network were enriched in the KEGG categories Glycerolipid metabolism and Non-alcoholic fatty liver dis-



(a) Genes per edge (b) Edges per gene

Figure 35: Gene distribution in Chow and HFD networks. **(a)** shows the number of genes that build up an edge. **(b)** shows the number of edges that could be found for one gene.

ease (NAFLD) (Tables A20 and A21). This finding confirmed that our CoNI approach is reflecting the metabolic phenotype of the HFD group since the mice showed elevated TAG levels in the liver, which is a sign of fatty liver disease.

3.3.1.4 *Local regulator genes impact metabolic sub-networks within the Chow and HFD networks*

The primary motivation for integrating the genetic data into the metabolic network was to identify genes that control hepatic metabolite levels under different dietary conditions. Thus, so-called local regulator genes (LRGs) were defined as genes that were significantly enriched in a local sub-graph within the Chow or HFD network, thereby controlling a densely connected metabolic sub-network (Section 2.7.2).

This prioritisation of candidate genes helped to identify 20 LRGs in the Chow and 59 LRGs in the HFD network, respectively. None of the LRGs was identified in both networks showing that the network-based approach reflected the diet-dependent changes in hepatic metabolism. From these 79 identified LRGs, eight genes were differentially expressed between mice on chow and high-fat diet. In the Chow network *Ddx3x* and in the HFD network *Myc*, *Arhgap24*, *Smim13*, *Rapgef4*, *Cd82*, *Inhbe*, and *Gk* were identified as differentially expressed LRGs (Figures A6 and A7, Tables 7, A22, and A23).

Network-based approach reflected the diet-dependent changes in hepatic metabolism

Among the differentially expressed LRGs were known regulators of hepatic lipid metabolism and genes that had been shown to be regulated by HFD feeding before, whereas others had not been described

in the context of obesity or liver metabolism, yet. The expression of the LRG glycerol kinase (*Gk*) that had been proposed as regulator for several lipids was higher in the HFD group compared to the Chow

Table 7: Characteristics and selection criteria for the genes further subjected to validation experiments, identified in the Chow and HFD setting.

GENE	SELECTION	NETWORK	SI-RNA	LOG2FC HFD VS. CHOW	P-VALUE	ADJUSTED P-VALUE
Arhgap24	LRG	HFD	no	0.992	1.23E-03	2.07E-02
Cd82	LRG	HFD	no	0.500	1.83E-03	2.69E-02
Gk	LRG	HFD	yes	0.729	5.99E-04	1.36E-02
Inhbe	LRG/SNP	HFD	yes	1.200	1.56E-03	2.43E-02
Myc	LRG	HFD	no	-1.309	3.66E-03	4.20E-02
Rapgef4	LRG	HFD	yes	-1.227	1.35E-05	1.38E-03
Smim13	LRG	HFD	no	-0.890	6.44E-04	1.40E-02
Ddx3x	LRG	Chow	no	-0.602	2.72E-04	8.23E-03
Cobll1	SNP	HFD	yes	-0.334	6.58E-02	2.38E-01
Appl2	SNP	HFD	yes	0.632	2.91E-02	1.48E-01
Tap1	DE	Chow/HFD	no	1.043	1.50E-05	1.48E-03

control (Table 7). The increased expression of *Gk* under HFD exposure could be interpreted as adaptive mechanism to handle the increased hepatic lipid load. This hypothesis is supported by the finding that overexpression of *Gk* favours recycling of free fatty acids leading to increased fat storage in rat hepatoma cells [91, 102, 152, 216].

Among the differentially expressed LRGs were known regulators of hepatic lipid metabolism

Another gene that had already been shown to play a role in liver metabolism, was the Myelocytomatosis oncogene (*Myc*) encoding c-myc, which is a pleiotropic transcription factor [83, 125]. *Myc* is not only involved in the regulation of general cellular processes, such as cell growth, cell proliferation, apoptosis, and differentiation [83] but also required for hepatocellular proliferation and liver tumourigenesis [185] as well as the regulation of hepatic glycolysis [192, 193, 235]. Under HFD exposure, *Myc* overexpression in transgenic mice normalises glycemia, insulinemia, and the expression of genes involved in hepatic metabolism [194]. This finding indicates that *Myc* plays an important role in the regulation of liver metabolism under HFD conditions, which is in line with our observed decrease in hepatic expression of *Myc* in the HFD-fed mice compared to the Chow controls (Table 7).

The last differentially expressed LRG with known function in hepatic metabolism was Inhibin βE (*Inhbe*) which encodes an inhibin β subunit and is together with *Inhbc* mainly expressed in the liver [54,

249]. *Inhbe* was observed to be regulated by the nutritional status, such as high fat diet, fasting, and re-feeding in the rodent liver [95, 96, 195]. High-fat diet feeding induced the expression of *Inhbe* in the liver [95]. Data from human livers revealed that the hepatic activin E complex, which consists of two INHBE subunits [163], was potentially linked to insulin resistance, shown by a positive correlation between hepatic *Inhbe* mRNA levels and insulin resistance [223]. Sugiyama et al. [223] demonstrated that a siRNA-mediated knockdown of the *Inhbe* gene decreased fat mass in db/db mice. Along this line of evidence, hepatic expression of *Inhbe* was significantly up-regulated in the HFD group compared to Chow (Table 7).

The other five differentially expressed LRGs encode proteins that are involved in several cellular processes that occur in the healthy state as well as in the pathological state, such as tissue differentiation, cell proliferation, tumour cell metastasis and cell migration. However, no associations to diet-induced obesity or fatty liver disease are known to date.

The LRG *Ddx3x* (DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3, X-linked), identified in the Chow network, is a member of the DEAD-box RNA helicase family and participates in several gene regulation events, including transcriptional regulation, RNA unwinding, splicing, RNA nuclear export, ribosomal biogenesis, and mRNA translation [245]. Expression of *Ddx3x* was reduced by HFD feeding (Table 7) indicating a potential regulating role for hepatic processes.

Also showing a decrease in hepatic gene expression upon HFD exposure was *Rapgef4* (Rap guanine nucleotide exchange factor (GEF) 4) that encodes the protein Epac2 (Table 7). So far, three isoforms have been identified that differ in their tissue expression pattern [222]. Whereas the brain-specific isoform Epac2A was linked to obesity, the functions of the adrenal gland specific Epac2B and the liver specific Epac2C have not been fully unravelled, yet [108, 222]. The tissue specific expression pattern and its regulation by HFD make Epac2C an interesting new target for obesity related deregulation of liver metabolism.

The last LRG with decreased hepatic expression in HFD was *Smim13* (Small integral membrane protein 13, Table 7), whose function is not resolved, yet.

A LRG showing a HFD-induced increase in expression was CD82 antigen (*Cd82*, Table 7), which is one of the most characterised members of tetraspanins. This family of proteins with four transmembrane domains is widely expressed and plays a physiological role in cell adhesion, motility, activation, and proliferation [180] but has so far not been discussed in the context of obesity related fatty liver metabolism.

Another LRG with increased expression in the HFD group was *Arhgap24* (Rho GTPase activating protein 24, Table 7) that is involved

in cell cycle, invasion, apoptosis and other cell processes [262]. It has implications for cancer and is a promising target in proteinuric kidney disease [170, 262] but has not been described a regulator of hepatic metabolism so far.

Additional to the analysis of the differentially expressed genes, the Type 2 Diabetes Knowledge Portal [112] was queried to determine LRGs that contain Single Nucleotide Polymorphisms (SNPs) associated with obesity and related disease markers in human. For our set of 79 LRGs in total, we could identify SNPs for three LRGs of the Chow set (Table A22) and for 17 LRGs of the HFD set (Table A23). Among the LRGs associated most with the human SNPs of interest were Cobl-like 1 (*Cobll1*), Adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper containing 2 (*Appl2*), and the differentially expressed LRG *Inhbe*, which was described above. *Cobll1* is known to bind actin monomers and cadherin but aside from that its function is not fully resolved, yet. So far, *Cobll1* had not been described in the context of high-fat diet feeding, obesity and/or diabetes. In contrast, *Appl2* had been shown to be involved in insulin signaling, endosomal trafficking, adiponectin signaling and other signaling pathways [147]. Ryu [200] demonstrated that a liver-specific knockout of *Appl2* improved insulin sensitivity, increased adiponectin signaling, and induced anti-inflammatory effects in HFD-fed mice, implicating the involvement of *Appl2* in hepatic metabolism.

In a final step, we analysed the metabolites that were directly regulated by the LRGs within the isolated sub-networks (Figure 36). For all selected genes, the regulated metabolites showed substantial differences between Chow and HFD (Table A24) indicating that these genes were indeed influencing the corresponding metabolic sub-network in the respective phenotype setting. The number of metabolites regulated by the selected LRGs and therefore present in the respective isolated sub-networks ranged from six metabolites (*Ddx3x, Smim13* sub-networks) to 14 metabolites (*Cobll1* sub-network). According to the detected abundance and diet-dependent regulation of measured metabolites, most genes within the sub-networks were regulating PC (*Gk* and *Cobll1* sub-networks). However, also metabolites of the classes SM (*Ddx3x, Arhgap24, Cd82, Appl2*) or AC (*Myc*) were found in the respective sub-networks. Interestingly, the sub-networks of the LRGs *Rapgef4* and *Smim13* mainly contained AA, whereas *Inhbe* only regulated AC.

Taken together, several local regulators selected with our network-based approach were genes associated with obesity, type 2 diabetes, HFD feeding, liver metabolism, and cellular processes. This indicated that our approach enabled us to detect genes playing a superordi-

Figure 5

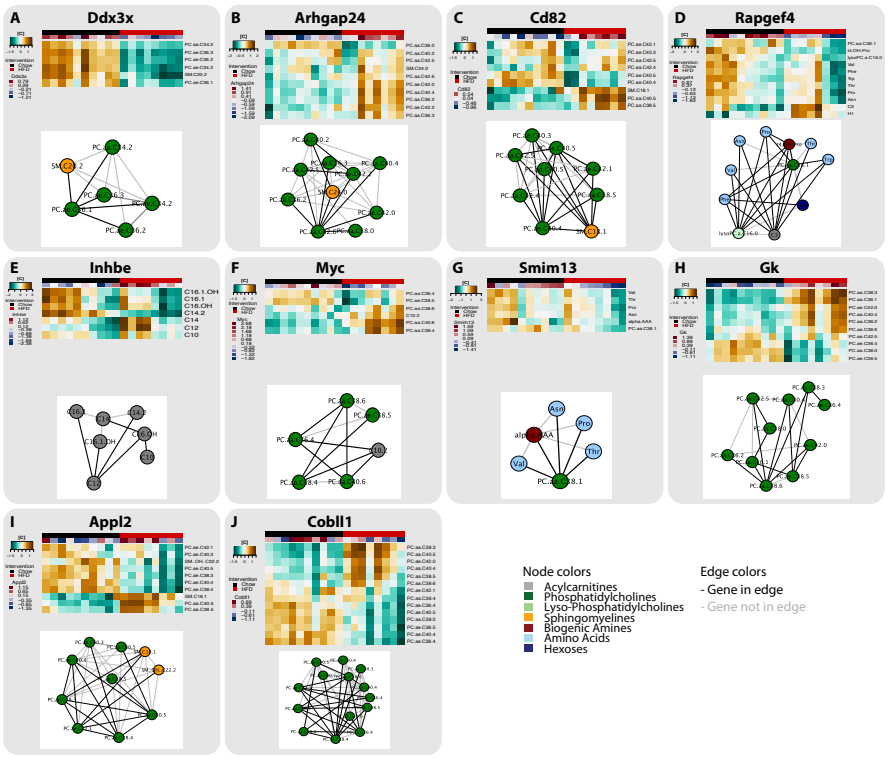


Figure 36: LRGs selected for further investigation for the Chow/HFD setting.

nate role in the regulation of metabolite levels in the context of HFD-induced hepatic steatosis compared to a solely transcriptomics-based approach. Especially the identification of *Rapgef4*, *Inhbe*, and *Cobll1* as important regulators demonstrates that the here presented method is able to identify promising new gene candidates, here in the context of dietary induced liver steatosis.

Initially, we aimed to identify genes that specifically regulate a metabolic pathway. Here, it is important to note that our metabolic data was restricted to 181 metabolites detectable by the used kit, thus, our reconstructed networks are only capable of representing a large abstraction of the underling metabolic network. In our case, no direct gene-metabolite interactions as annotated in KEGG were retrieved with our method. Therefore, we rather expected to identify genes that do not directly control metabolite concentrations, such as enzymes, but more upstream genes with a controlling effect on metabolic pathways. Additionally, only one of the genes selected for validation was among the 100 most significantly regulated genes between Chow and HFD. Thus, the genes identified here would likely have been remained undiscovered by only analysing the transcriptome dataset. The here presented CoNI approach helps to identify otherwise hidden genes that directly influence specific metabolic sub-

networks of the phenotypic setting of interest. Thus, new hypotheses on the molecular underpinnings can be proposed on a more functional level as the metabolite network is included as a second layer to the transcriptome dataset.

3.3.1.5 *Validation of LRGs by siRNA-mediated knockdown showed influence of selected genes on metabolic networks*

Our CoNI approach enabled us to identify eight LRGs differentially expressed in livers of chow- and HFD-fed mice. With the additionally selected two SNP-associated LRGs, and *Tap1*, which was selected due to its presence in both networks and for being differentially expressed at the same time, a total of eleven genes were selected for further investigation (Figure 36, Tables 7 and A24).

These genes were then prioritised for the in vitro validation approach resulting in a set of five genes, for which the siRNA-mediated knockdown (KD) experiments in HepG2 cells were performed and metabolite levels were measured using the AbsoluteIDQ^{TM} p180 kit. All siRNAs successfully knocked down the target genes compared to a non target siRNA (Figure 37).

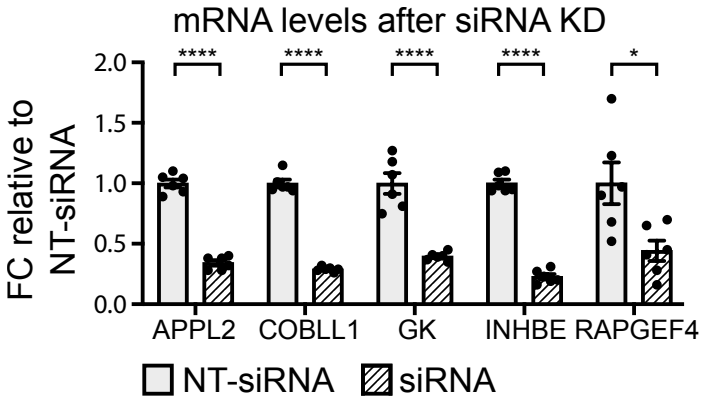
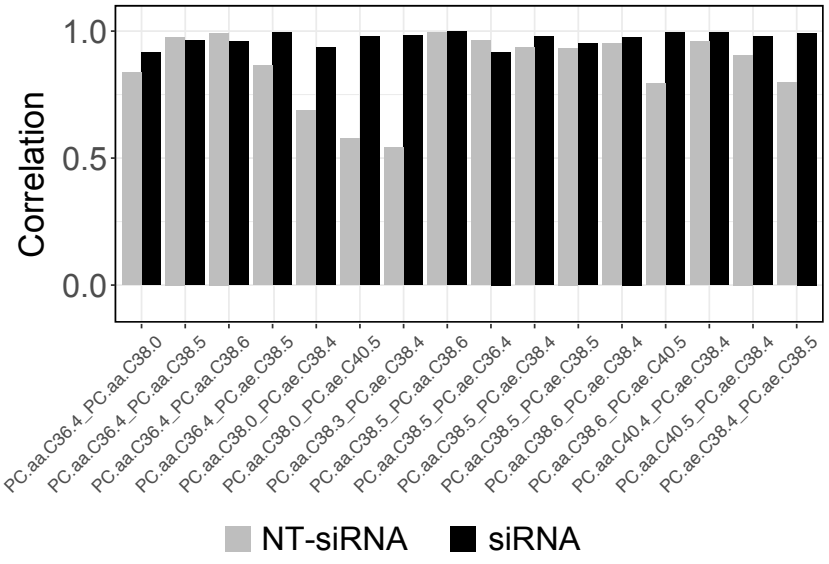


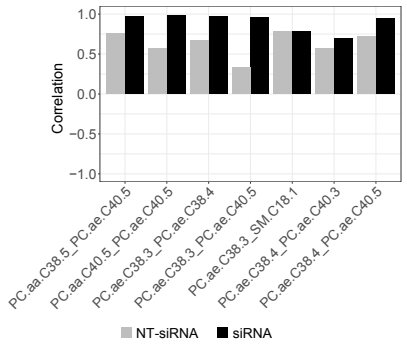
Figure 37: mRNA levels of target genes after siRNA-mediated knockdown in HepG2 cells. Fold changes are shown relative to control samples.

Since we had used pairwise metabolite correlation as a first step to build our CoNI approach for the in vivo datasets, we now looked at the concentration correlation pattern of the metabolites connecting the edges in the sub-graphs of the respective LRG (edges marked in black in Figure 36). In the in vitro validation system, the concentration levels of several metabolites stayed below the limit of detection (LOD). These metabolites included 38 of the 40 measured acylcarnitines, three of the 21 amino acids, 15 of the 21 biogenic amines, 22 of the 90 glycerophospholipids, and three of the 15 sphingolipids. Be-

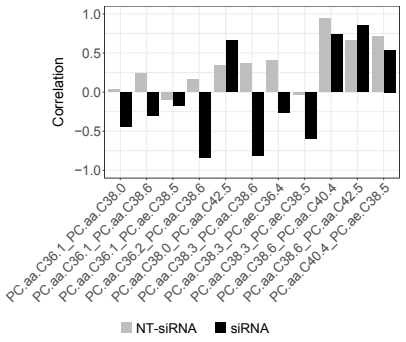
cause *Inhbe* is only associated with acylcarnitines we had to exclude the gene from this analysis.



(a) *Cobll1*



(b) *Appl2*



(c) *Gk*

Figure 38: Correlation coefficients of metabolite pairs present in the **(a)** *Cobll1*, **(b)** *Appl2*, and **(c)** *Gk* sub-networks for control and knockdown samples. Coefficients are only shown for metabolite pairs that build an edge that contains *Cobll1*, *Appl2* or *Gk* in the HFD network, respectively.

Pairwise metabolite correlation differs between gene KD and control samples

The pairwise metabolite correlations detected for *Cobll1* and *Appl2* in vivo were always stronger when the respective gene was regulated by diet-induced obesity compared to the non-target siRNA control setting (Figure 38a and 38b). A similar finding was made for *Gk* (Figure 38c; 8 out of 11 pairs) and for *Rapgef4* (Figure 39; 9 out of 13 pairs). However, this observation was only a trend, as the difference between the correlation coefficients for the pairs observed in the control samples and in the knockdown samples was not statistically significant (Table A25). Nevertheless, the fact that the correlation coefficients differed between KD and control cells still indicated, that the respective

gene influenced the concentration pattern of the respective metabolite pair.

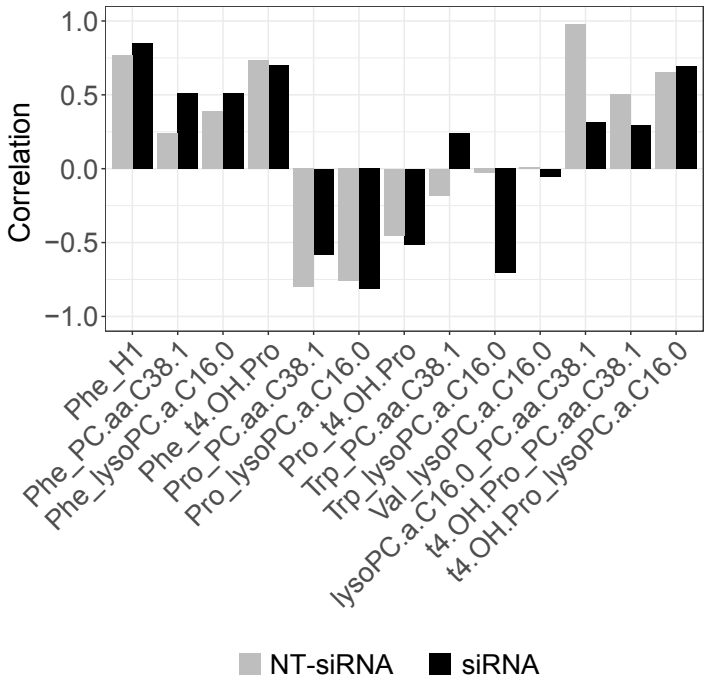
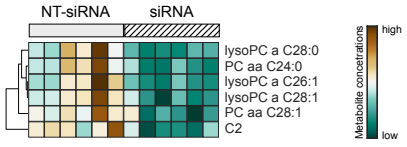
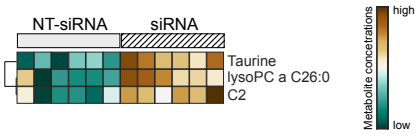


Figure 39: Correlation coefficients of metabolite pairs present in the *Rapgef4* sub-network for control and knockdown samples. Coefficients are only shown for metabolite pairs that build an edge that contains *Rapgef4* in the HFD network.

Next, the comparison of metabolite concentration levels between each knockdown sample set and control samples revealed between two and 58 altered metabolite concentrations for *Cobll1* (Figure 40a), *Gk* (Figure 40b), *Rapgef4* (Figure 41), and *Inhbe* (Figure 42). This points to the predicted regulatory effect of the respective gene on the metabolic network.



(a) *Cobll1*



(b) *Gk*

Figure 40: Metabolites significantly altered by siRNA-mediated knockdown compared to untreated HEPG2 cells for **(a)** *Cobll1* and **(b)** *Gk*.

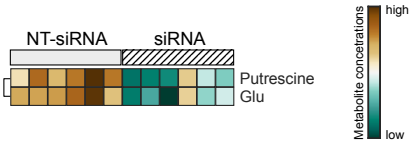


Figure 41: Metabolites significantly altered by siRNA-mediated knockdown compared to untreated HEPG2 cells for *Rapgef4*.

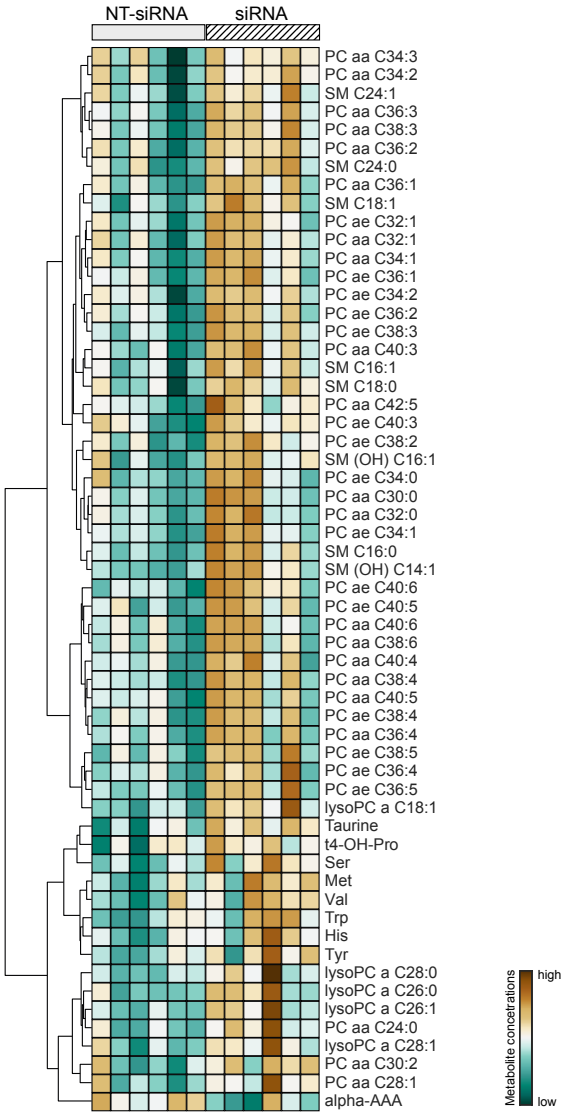


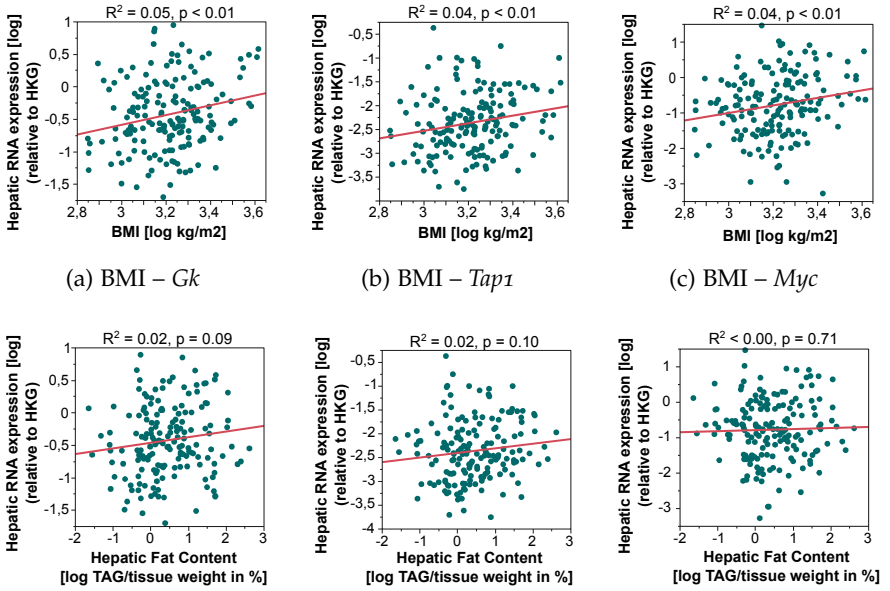
Figure 42: Metabolites significantly altered by siRNA-mediated knockdown compared to untreated HEPG2 cells for *Inhbe*.

Taken together, the analyses showed that the metabolite interactions / networks indeed change upon siRNA-mediated knockdown of at least four of the five candidate genes indicating that these genes significantly impact cellular metabolism.

3.3.1.6 *Validation of method in human samples*

To additionally assess the translational relevance of our developed method, hepatic mRNA expression profiles of the selected LRGs were correlated with clinical parameters, such as hepatic triglyceride content and BMI, in liver tissue biopsies of 170 human volunteers. The subjects covered a wide range of liver fat content (Table 2). Furthermore, in a subgroup of 77 subjects with available fasting blood samples, relationships between metabolic characteristics and the HOMA-IR, which is an indicator for insulin resistance, were analysed (Table 3).

For five of the eleven selected genes, significant associations between gene expression and metabolic traits could be observed (Table A26). *GK*, *TAP1*, and *MYC* were associated with BMI (Figures 43a, 43b, and 43c) but not with liver fat content (Figures 43d, 43e, and 43f).



(a) BMI – *Gk* (b) BMI – *Tap1* (c) BMI – *Myc*

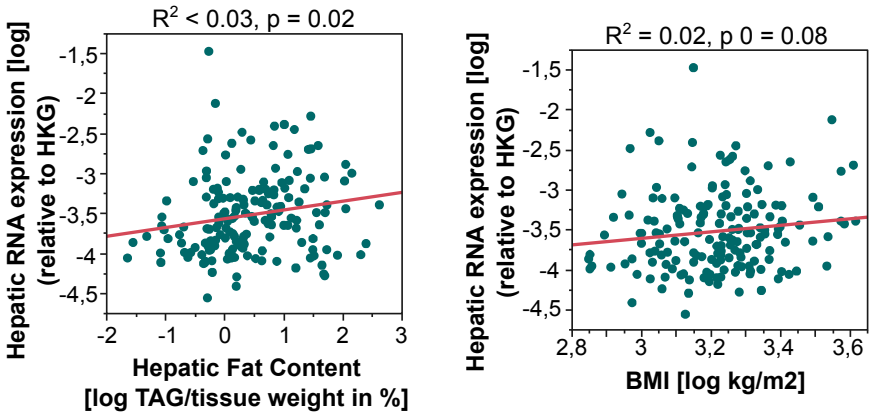
(d) HTC – *Gk* (e) HTC – *Tap1* (f) HTC – *Myc*

Figure 43: Pearson correlation analysis of human hepatic gene expression compared to BMI for **(a)** *GK*, **(b)** *TAP1*, and **(c)** *MYC* and compared to hepatic triglyceride content (HTC) for **(d)** *GK*, **(e)** *TAP1*, and **(f)** *MYC*.

SMIM13 was associated with liver fat content, while it did not show an association with BMI (Figures 44a and 44b).

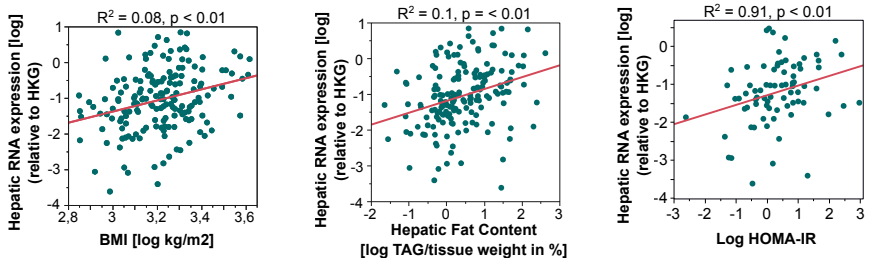
INHBE was associated with BMI, liver fat content and whole-body insulin resistance

Aside from BMI, *INHBE* was also associated with liver fat content and whole-body insulin resistance (Figures 45a, 45b, and 45c). The correlation of hepatic *INHBE* mRNA levels and insulin resistance was also recently shown by Sugiyama et al. [223].



(a) HTC (b) BMI

Figure 44: Pearson correlation analysis of human hepatic gene expression of *SMIM13* compared to **(a)** hepatic triglyceride content (HTC) and to **(b)** BMI.



(a) BMI (b) HTC (c) HOMA-IR

Figure 45: Pearson correlation analysis of human hepatic gene expression of *INHBE* compared to **(a)** BMI, to **(b)** hepatic triglyceride content (HTC), and to **(c)** the HOMA-IR.

Both validation approaches, siRNA-mediated knockdown in HepG2 cells and transcriptional profiling in human liver biopsies, could confirm that seven of the eleven selected genes show a transcriptional association with clinical obesity or show regulatory effects on cellular metabolite levels. Among these seven confirmed genes, four had already been linked to hepatic lipid metabolism, whereas three have not been described in the context of high-fat diet feeding, obesity and/or diabetes before. The method identified among others *Inhbe* and *Cobll1* as important regulators, which could be confirmed. Especially *Inhbe* was significantly correlated with body weight, liver triglyceride levels and whole-body insulin resistance and showed the strongest impact on cellular metabolism. This underlines that CoNI successfully integrates transcriptomics and metabolomics data and enables us to identify regulatory gene candidates in the context of hepatic steatosis induced by HFD.

3.3.2 *Application of CoNI to weight loss intervention datasets*

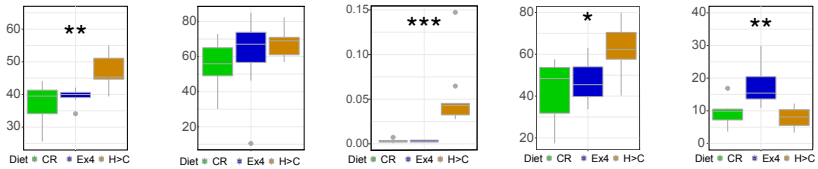
Our biological interest focused on the underlying mechanisms for the observed physiological differences between the two weight loss intervention groups, exendin-4 treatment versus calorie restriction. While both mice cohorts lost the same amount of body weight, only exendin-4 treatment was sufficient to restore hypothalamic leptin sensitivity. On the other hand, this pharmacological-induced weight loss caused a massive accumulation of triacylglycerides in livers of mice after ten days of treatment. This acute hepatic steatosis turned out to be a transient phenomenon since the livers of mice treated with exendin-4 for 30 days were clear of triglycerides when compared to the HFD control group (data not shown). Neither restoration of leptin sensitivity nor acceleration of hepatic steatosis were observed when body weight was lost by caloric restriction.

To investigate the molecular underpinnings of the hepatic triglyceride accumulation in exendin-4 treated mice and to identify novel regulator genes of hepatic metabolism during weight loss interventions, we applied our new CoNI method to the metabolome and transcriptome datasets of the two weight loss groups CR and Ex4 as well as to their diet-switch control group H>C. As the focus of the study is on the two weight loss intervention groups, additional figures and tables for the comparisons of Ex4 or CR, respectively, with the diet-switch control can be found in the appendix.

3.3.2.1 *Genetic and metabolic profiling in livers of mice under weight loss intervention pointed towards a differential regulation pattern in liver metabolism*

Prior to the analysis of the liver-specific genetic and metabolic profiles of mice undergoing weight loss interventions, we assessed basic phenotypic parameters and revealed that both the calorie restricted as well as the exendin-4 treated mice had a significantly lower body weight compared to the diet-switch control group (Table A27 and Figure 46).

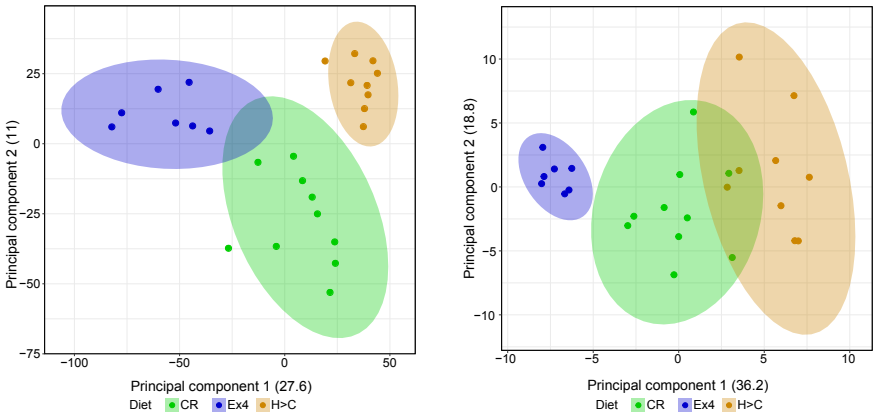
The reduction in body weight was reflected by reduced plasma insulin levels and lower circulating TAG levels compared to the diet-switched mice (Figures 46a, 46c, and 46d), indicating that the mice in the CR and Ex4 groups had overcome the obesity-induced hyperinsulemia. Plasma cholesterol levels were similar for all three groups (Figure 46b). Hepatic steatosis that was assessed by Oil-Red-O-Staining and extraction of hepatic TAG was significantly elevated in mice acutely treated with exendin-4 (Figure 46e).



(a) Body weight [g] (b) Plasma cholesterol [mg/dl] (c) Plasma insulin [mg/dl] (d) Plasma TAG [mg/dl] (e) Liver TAG [µg/mg]

Figure 46: Phenotype monitoring of Ex4 treated and calorie restricted mice and mice of the diet-switch control group H>C: **(a)** Body weight (BW) measured at the end of the study, **(b)** plasma cholesterol levels, **(c)** plasma insulin levels, **(d)** plasma and **(e)** liver triacylglyceride (TAG) levels. *: p <= 0.05; **: p <= 0.01; ***: p <= 0.001

The initial analysis of the gene expression profile and the metabolite profile of the CR, Ex4 and H>C mice by PCAs revealed that the intervention was in all three cases the main influencing factor responsible for the observed variance in the gene and metabolite abundance profiles (Figure 47). The first principle component explained 27.6% of the variance observed in the gene expression profiles and 36.2% of the variation encountered in the metabolite concentration profiles.

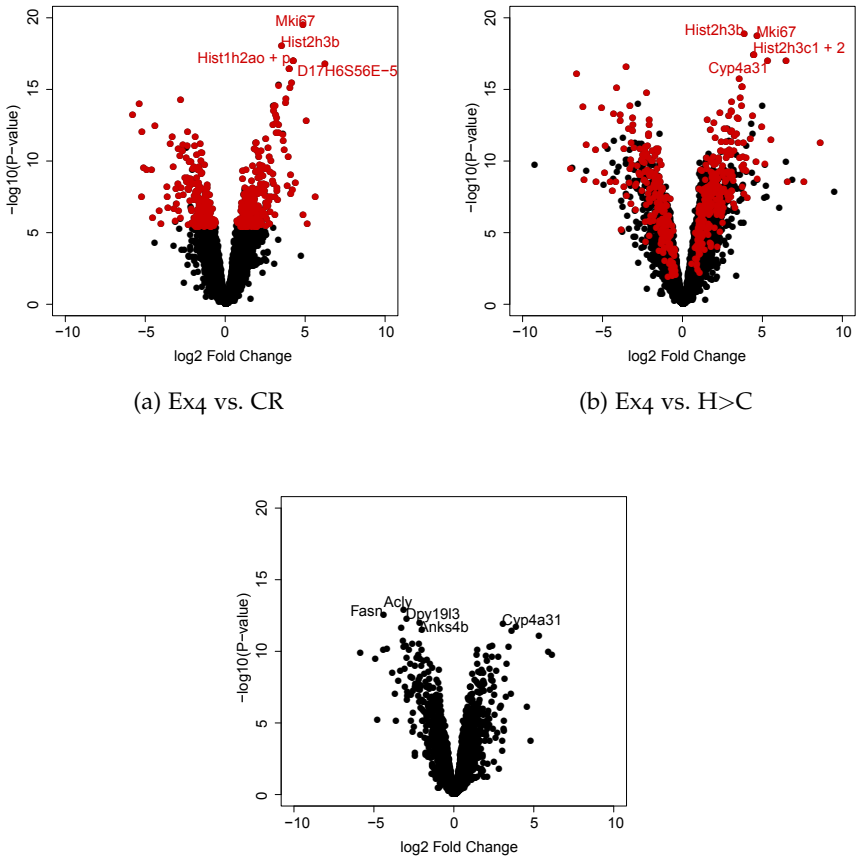


(a) Gene expression profiling (b) Metabolic profiling

Figure 47: PCA showing the main contributor to differences observed in **(a)** gene expression and **(b)** metabolite concentration profiles of the weight loss cohorts.

Next, the differential gene expression analysis comparing the hepatic gene expression profiles of exendin-4 treated mice to CR mice identified 494 genes that were significantly deregulated in the livers of exendin-4 treated mice (Figure 48a and Table A28). Of the differentially expressed genes, 271 genes were up- and 223 genes were down-regulated.

Among the top 50 significantly differentially expressed genes in liver between Ex4 and CR were two genes involved in hepatic lipid metabolism: Hepatic transcript levels of the lipase and esterase *Ces3b* (Carb-oxyl-esterase 3B) were down-regulated in the livers of exendin-4 treated mice compared to calorie restricted mice. Since a high expression of *Ces3b* had been hypothesised to contribute to a reduced hepatic lipid content [136], the lower *Ces3b* expression in exendin-4 treated mice is in line with the increased hepatic lipid content in this group compared to CR mice.
The second differentially expressed gene involved in hepatic lipid metabolism was *Cyp2c54* (Cytochrome P450 2C54) that is known to play a role in liver detoxification processes [146, 243]. The down-regulation of *Cyp2c54* in the livers of exendin-4 treated mice could be a likely consequence of the elevated TAG levels.



(a) Ex4 vs. CR

(b) Ex4 vs. H>C

(c) CR vs. H>C

Figure 48: Volcano plots of deregulated liver genes comparing **(a)** Ex4 treatment to CR, **(b)** Ex4 treatment to H>C, and **(c)** CR to H>C. Genes that are deregulated in the same direction in Ex4 compared to CR as well as to H>C are highlighted in red.

Weight loss itself, independent of the treatment, had a great impact on hepatic transcript levels

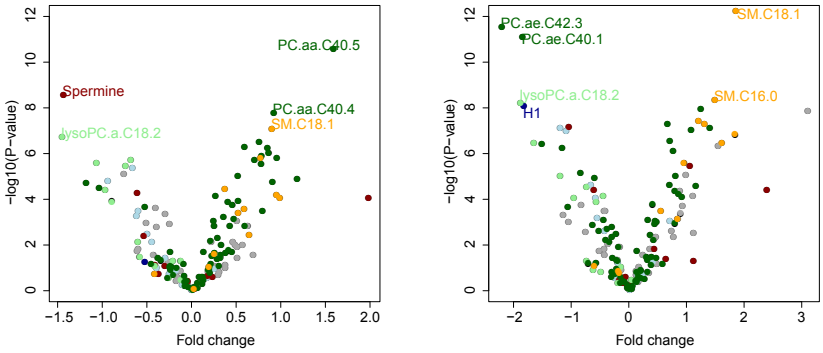
When we compared the gene expression profiles of each weight loss intervention group to the diet-switch control group, we found an elevated number of differentially expressed genes thus indicating that the weight loss itself, independent of the treatment, had a great impact on hepatic transcript levels. We observed 3,821 differentially expressed genes between Ex4 and H>C (Figure 48b and Table A29), and 1,692 genes altering between CR and H>C (Figure 48c and Table A30). From these genes, 2,159 and 952 genes were up- and 1,662 and 740 genes were down-regulated for Ex4 vs. H>C and CR vs. H>C, respectively. Interestingly, *Ces3b* was also down-regulated in the livers of exendin-4 treated mice compared to the diet-switch mice, thus confirming an exendin-4 specific effect. This effect could also be observed by the high number of genes that are deregulated in the same direction when comparing Ex4 to either CR or H>C (Figures 48a and 48b).

10 days of CR changed the transcript profile regulating hepatic lipid metabolism

The performed GO term overrepresentation analysis of the 50 genes with the lowest adjusted p-values for each comparison showed that 22 of the genes differing between CR and H>C did belong to lipid metabolic processes (FDR: 3.38e-14). Genes differing for Ex4 vs. H>C as well as for Ex4 vs. CR were mainly involved in chromatin organisation (FDR: 0.006 and 0.0011, respectively). This indicates, that ten days of CR already changed the transcript profile regulating hepatic lipid metabolism, despite no obvious changes in liver TAG levels compared to the diet-switch control. In contrast, lipid metabolism in livers of exendin-4 treated mice was not changed but rather further impaired compared to the still obese diet-switch control, which was also mirrored on the metabolite level by elevated triacylglyceride levels in livers of exendin-4 treated mice.

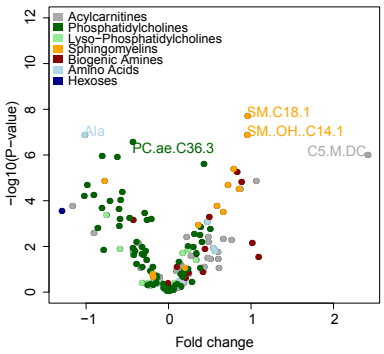
Following the differential gene expression analysis, differential metabolome analysis was performed. In livers of exendin-4 treated mice it identified 70 metabolites to be altered compared to CR and 92 differed when compared to H>C. Two thirds of the metabolites (46 of the 70) that differed in their concentrations between Ex4 compared to CR belonged to one of the three measured phospholipid classes: Phosphatidylcholines (PC), Sphingomyelins (SM), and Lyso-Phosphatidylcholines (LPC) (Figure 49a and Table A31).

A similar pattern of deregulated metabolite classes was seen for the comparison of Ex4 to the diet-switch (Figure 49b and Table A32). Surprisingly, the comparison of CR to the diet-switch control also revealed that two thirds (44 of the 68) of the metabolites belong to the three measured phospholipid classes (Figure 49c and Table A33). It might, however, just reflect the narrow portfolio of metabolites measured in our targeted approach.



(a) Ex4 vs. CR

(b) Ex4 vs. H>C



(c) CR vs. H>C

Figure 49: Volcano plots of deregulated liver metabolites comparing **(a)** Ex4 treatment to CR, **(b)** Ex4 treatment to H>C, and **(c)** CR to H>C.

A clear regulation pattern for the deregulated phospholipid levels could be observed

We observed a clear regulation pattern for the deregulated phospholipid levels in the livers of exendin-4 treated mice compared to CR mice. Metabolites that belonged to the PC and SM classes were up-regulated upon exendin-4 treatment compared to CR, while down-regulated metabolites mainly belonged to the LPC, AC, and AA classes. As discussed in Section 3.3.1.1, in the state of hepatosteatosis and steatohepatitis in mice, however, PC levels are decreased compared to histologically healthy livers. This could indicate that despite having high triglyceride levels in the liver the exendin-4 treated mice do not suffer from hepatosteatosis. This conclusion would be in line with the transient manner of the triglyceride level elevation and supports the hypothesis of an acute buffering capacity of the liver during acute, pharmacological-induced weight loss.

Within the metabolite class of biogenic amines, we found a quite prominent down-regulation of the naturally occurring polyamine spermine upon exendin-4 treatment compared to CR. Spermine is, together with its precursor spermidine, an essential regulator of several

cellular processes, such as DNA stability, cellular growth, differentiation, and apoptosis and might also be involved in autophagy [4]. Recent reports have not only shown an inter-relationship between autophagy and the process of lipid breakdown but also that the inhibition of macroautophagy, which is one of the three forms of autophagy, leads to an increase in cellular triglyceride content [7, 51, 208]. Therefore, it would be interesting to investigate whether the transient increase of liver triglycerides could be explained by an alteration in autophagal processes.

Overall, genetic and metabolic profiling in livers of mice that underwent weight loss interventions revealed massive changes in transcript and metabolite levels compared to their diet-switch control that can be attributed mostly to the loss in body weight. However, the defined changes that were observed between the two treatment groups point towards a differential regulation pattern in liver metabolism, which could play an important role for whole body metabolism and for adaptive mechanisms after the period of acute weight loss.

3.3.2.2 *Correlation pattern of metabolites revealed weight loss intervention-dependent changes in the hepatic metabolism*

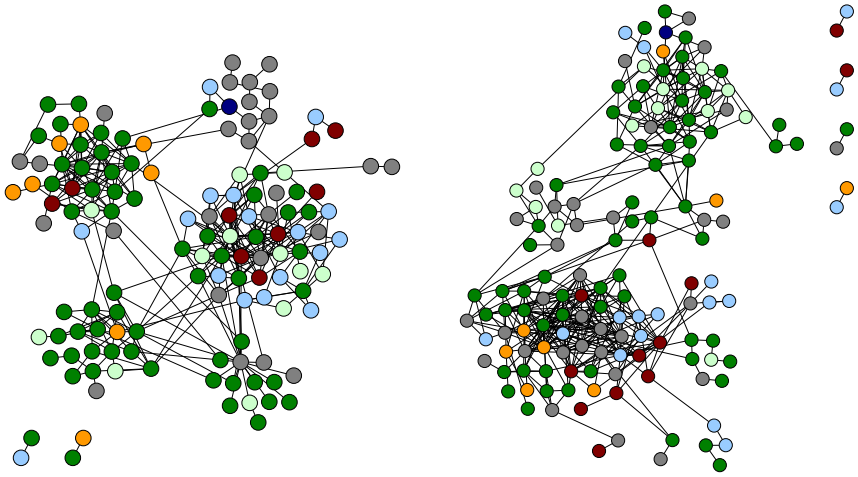
In order to identify differences in the genetic control of hepatic metabolic networks during weight loss, the newly developed CoNI approach was applied to the Ex4 and CR dataset. We mainly concentrated on the detection of LRGs, those genes that were significantly enriched in a local sub-graph within the CR or Ex4 network.

To investigate whether the method of weight loss intervention altered the interaction between liver metabolites, we calculated for each metabolite pair the Pearson correlation coefficient under Ex4 and CR treatment, respectively. By allowing a non-adjusted p-value$<$0.05 for the correlation coefficient significance, 1,536 and 2,643 metabolite pairs were selected for Ex4 and CR, respectively. 341 metabolite pairs correlated significantly in both sets, indicating that the administered interventions substantially altered the hepatic metabolism.

3.3.2.3 *Networks obtained with CoNI reflected the differences in the metabolic phenotype of Ex4 treated mice compared to calorie restriction*

To assess the genetic impact on the hepatic metabolite network formed under each weight loss intervention, we integrated the gene expression profiles and the metabolite concentration profiles applying our developed approach CoNI (Section 2.7.1).

In the first step, we obtained two independent networks for CR and Ex4 (Figure 50). The networks were constructed of 389 metabolite –



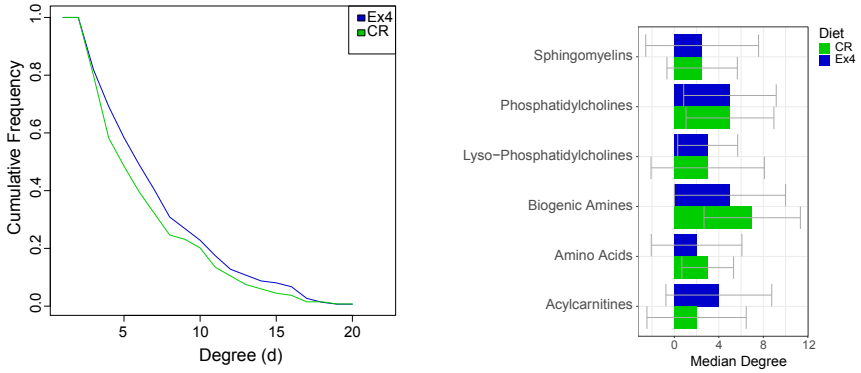
(a) CR Network (b) Ex4 Network

Figure 50: Networks obtained with CoNI for **(a)** CR and **(b)** Ex4. Nodes represent metabolites and are coloured according to their metabolite class. Edges are build from genes that influence the respective metabolite pair.

gene triplets for the CR and 565 metabolite – gene triplets for the Ex4 set, respectively. From the 181 measured metabolites 134 were connected in the CR and 149 metabolites in Ex4 network, respectively. Of the metabolites that could be connected, 114 were found in both networks. When the resulting metabolite pairs were compared between the CR and Ex4 dataset, 13 metabolite pairs were found to be connected in both diet networks. In the CR network, 319 metabolite pairs were exclusively connected, whereas 409 metabolite pairs were connected specifically in the Ex4 network. It is worth to note here that the number of metabolite pairs connected in both weight loss intervention networks was more than five times lower than in the Chow and HFD datasets, where 67 metabolite pairs were connected in both networks (Section 3.3.1.3).

When the node degree distributions of both networks were compared, similar degrees for both networks were observed. Both graphs followed a power law distribution (Figure 51a). The comparison of the node degrees for the specific metabolite classes showed that the Ex4 network had higher degrees for the acylcarnitines (Figure 51b). In contrast, the biogenic amines and amino acids showed higher degrees in the CR network. The three phospholipid classes showed similar node degrees.

A more detailed analysis of the inferred networks revealed that the networks were organised in communities but in contrast to the Chow and HFD networks, they did not show a clear separation between



(a) Degree distribution (b) Node degrees per class

Figure 51: Node degree comparison. **(a)** Degree distribution and **(b)** node degrees shown for each metabolite class of the CR and Ex4 networks.

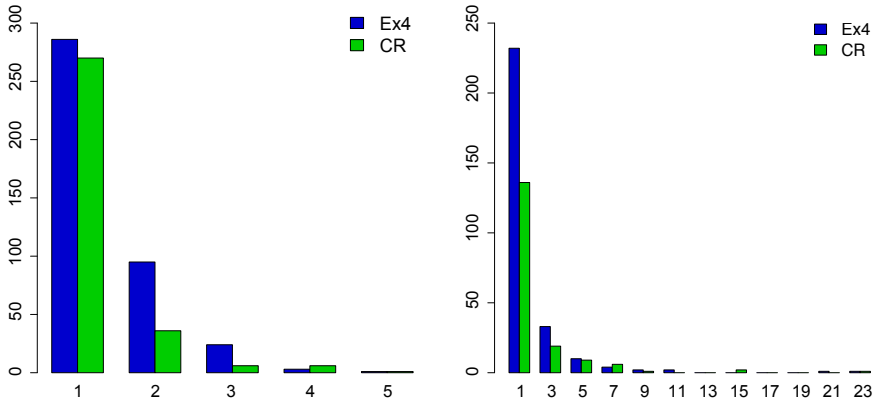
the metabolite classes. This might be a technical consequence as the graphs consisted of less metabolite – gene triplets than the Chow and HFD graphs.

Next, we had a closer look at the genes connecting the metabolite pairs. In contrast to the metabolites, the genes in both networks showed greater heterogeneity. Only nine genes were present in edges of both the CR and the Ex4 network (*Eno1, Hsd17b12, Acsf2, Mrps12, Rhoc, Abhd2, Grhpr, Acat1, Ifi47*). Of these genes present in both networks, only *Rhoc* was differentially expressed between Ex4 and CR. Within the CR network 175 individual genes were identified, of which 15 were differentially expressed between the two weight loss intervention groups Ex4 and CR (Figure A8). In contrast, almost twice as many individual genes (n=286) were found in the Ex4 network, of which five were differentially expressed in Ex4 compared to CR (Figure A9).

When the number of genes that map to a single edge was counted, we observed that most edges consisted of a single gene (69.23% in CR, 50.53% in Ex4). The maximum number of genes per edge was five genes in both networks (Figure 52a). The distribution of genes over the edges (Figure 52b) showed that the distribution is similar for both graphs, where less genes control more of the network.

Metabolism of lipids was enriched for the Ex4 network genes

To further classify the genes identified in each of the two networks, a functional enrichment analysis using Reactome pathways was performed. For the CR network we did not identify informative categories for the individual genes. However, the individual genes of the Ex4 network were identified to be enriched in the Reactome cate-



(a) Genes per edge (b) Edges per gene

Figure 52: Gene distribution in CR and Ex4 networks. **(a)** shows the number of genes that build up an edge. **(b)** shows the number of edges that could be found for one gene.

gories "Metabolism of lipids" and "Phospholipid metabolism" (Tables A34 and A35). This finding confirmed that our CoNI approach is reflecting the metabolic phenotype of the Ex4 group since the mice showed elevated TAG levels in the liver.

3.3.2.4 *Local regulator genes impact metabolic sub-networks within the Ex4 and CR networks*

Here, the primary motivation for integrating the genetic data into the metabolic network was to identify genes that control hepatic metabolite levels during the weight loss interventions.

The network-based approach reflected the intervention-dependent changes in hepatic metabolism

The CoNI approach helped to identify 17 LRGs in the CR and 24 LRGs in the Ex4 network, respectively. None of the LRGs was identified in both networks showing that the network-based approach reflected the intervention-dependent changes in hepatic metabolism. From these 41 identified LRGs, two genes were differentially expressed between Ex4 and CR mice. In the CR network *Magix* and in the Ex4 network *Cyba* were identified as differentially expressed LRGs, with *Magix* being down-regulated and *Cyba* being up-regulated upon exendin-4 treatment compared to CR (Figures A8 and A9).

The function of *Magix* (MAGI family member, X-linked) has not been annotated, yet, and the gene had not been described in the context of fasting or obesity so far. Interestingly, the Mouse ENCODE transcriptome project found that *Magix* is not expressed in the developing liver but highly expressed in the adult liver thus pointing towards a potential role in liver metabolism [259]. *Cyba* (Cytochrome b-245, alpha polypeptide) is a NADPH oxidase complex gene, which is also referred to as $p22^{phox}$ [62]. Abreu et al. [2] suggested that the el-

evated expression of $p22^{phox}$ during a hypercholesterolemic diet could be involved in the activation of the NADPH oxidase enzyme complex, which could favour the generation of reactive oxygen species (ROS), thus further escalating the oxidative stress in the cell. Since it was proposed that NADPH oxidases play a role in chronic liver diseases [115, 162], the increased expression of *Cyba* in the livers of exendin-4 treated mice along with their high concentration of triglycerides could indicate an induction of oxidative stress. However, on the long run GLP-1 analogs, such as exendin-4, have been shown to reduce oxidative stress by inhibiting NADPH oxidases at least in the kidneys [100].

Additional to the analysis of the differentially expressed genes, the Type 2 Diabetes Knowledge Portal [112] was queried to determine LRGs of the CR and the Ex4 set that contain SNPs associated with obesity and related disease markers in humans. For the complete set of 41 LRGs, we could identify obesity-associated SNPs for three LRGs of the CR set (Table A36) and for seven LRGs of the Ex4 set (Table A37). The LRG associated most with the human SNPs of interest was nuclear receptor binding factor 2 (*Nrbf2*), strongly associated with several obesity and type 2 diabetes related markers, such as BMI, HDL cholesterol and Triglycerides. *Nrbf2* is involved in the regulation of autophagy, but the exact function of the gene remains elusive [150, 151, 264].

The identification of an autophagy-involved gene in this context is quite interesting, as a possible alteration in autophagic behaviour had already been suggested by the down-regulation of the autophagy-involved metabolite Spermine upon exendin-4 treatment compared to calorie restriction. Autophagy is the process of degrading cytosolic materials. In the liver, autophagy contributes to essential functions, such as glycogenolysis, gluconeogenesis, and β-oxidation and functions as a defence mechanism against NAFLD [233]. It was reported that lipotoxic effects, such as insulin resistance and oxidative stress, which are induced by elevated triglyceride and free fatty acid levels in NAFLD, suppress autophagic processes [233], which would be in line with our findings in livers of exendin-4 treated mice.

In a final step, we analysed the metabolites that were directly regulated by the LRGs within the isolated sub-networks. For all selected genes, the regulated metabolites showed substantial differences between CR and Ex4 (Table A38) indicating that these genes were indeed influencing the corresponding metabolic sub-network in the respective phenotype setting.

Overall, the identification of several genes associated with obesity, fasting and autophagy in the livers of the two weight loss intervention groups indicated that the developed method is a valuable tool to select genes of interest in the context of the observed metabolic environment. However, in future experiments the exact function of the identified LRGs in the context of acute hepatic steatosis during exendin-4 treatment needs to be investigated.

4

SUMMARY AND CONCLUSION

4.1 TIME-MATCHED ANALYSIS OF DNA ADDUCT FORMATION AND EARLY GENE EXPRESSION IN METHYLAZOXYMETHANOL ACETATE TREATED EKER RATS

The concept currently accepted in the field states that the quantification of DNA adducts can only serve as a biomarker for internal exposure but not as a marker for DNA damage relevant for cancer or tumourigenesis [137, 177]. This concept is also supported by our data.

DNA adduct formation is considered to be crucial for carcinogenesis induced by genotoxic substances. But there are several factors, such as mutagenesis in critical genes and cell proliferation, that are additionally necessary for the translation of DNA adducts into cancer-causing mutations. Predicting the effect of genotoxic carcinogens on the organism should therefore be based on the detection of distinct biomarkers capturing the multistage mechanism of carcinogenesis.

Exposure to low dose MAMAc for two weeks resulted in a time-dependent accumulation of pro-mutagenic DNA adducts in the Eker rat kidney cortex. This was not reflected by time-matched expression changes of genes involved in cyto-protective mechanisms, such as DNA repair, cell cycle arrest or apoptosis. Instead, genes involved in tumour-related MAPK, FoxO, and TGFβ pathways were induced. Continuous MAMAc exposure for six months leads to a mild but significant increase in pre-neoplastic and neoplastic lesions. This hints to the DNA repair not being efficient enough to protect the cell against the increasing numbers of DNA adducts and an elevated risk of these DNA adducts to manifest into tumour-causing mutations.

We suggest that combining the quantification of DNA adducts with time-matched gene expression analysis improves the predictive value for carcinogenesis induced by a genotoxicant.

4.2 EFFECTS OF WEIGHT LOSS INTERVENTIONS ON HYPOTHALAMIC GENE EXPRESSION

Obese individuals show high levels of leptin correlating with increased fat mass while being "leptin resistant", meaning that the anorectic and metabolic effects of leptin are absent. Accordingly, therapy with recombinant leptin is inefficient in decreasing body weight of diet-induced obese mice and obese humans. Molecular underpinnings for the insensitivity towards leptin action are not entirely understood, yet. The aim of the present study was to investigate the superior restoration of leptin sensitivity by pharmacologically mediated weight loss compared to calorie restriction. Therefore, hypothalamic gene expression profiles of five diet groups were analysed – two weight stable control groups, one diet-switch control group, and two weight loss intervention groups.

First, we aimed to identify leptin-dependent changes of the translatome by investigating mRNA bound to ribosomes in neurons activated by leptin treatment. Unfortunately, the employed antibody did not bind the target protein reliably, leading to many intron-mapping reads and a high unreliability in the obtained dataset. Therefore, in the second approach the mRNA expression data of whole tissue hypothalamus samples were analysed and several genes known to be regulated by fasting were identified as differentially regulated in calorie-restricted mice compared to exendin-4 treated mice.

We were also able to show that some of these fasting-induced genes are expressed in AgRP and POMC neurons, which belong to the most important neuronal subpopulations involved in the regulation of food intake and energy expenditure, indicating that our analysis of the whole hypothalamus gene expression pattern is a good starting point to define genes of interest in the context of fasting.

4.3 ESTIMATING GENETIC IMPACT ON METABOLIC NETWORKS

Metabolites as endpoints of molecular processes are an important readout of cellular signaling and metabolic processes in a tissue of interest at a defined time-point. However, the genetic control of metabolic pathways, their dynamics, and dependence on different diets and disease states remains rather unknown. Here, we presented a newly developed statistical method for correlation-based network integration (CoNI) of generic character that can be applied to multiple experimental settings.

Here, we used the novel method to integrate metabolome and transcriptome datasets to unravel previously hidden regulatory genes that exert major changes to hepatic metabolite levels. We applied the new method CoNI to two fascinating liver phenotypes. First, we integrated liver metabolomics and transcriptomics of chow- and HFD-fed mice, and identified several genes that could play an important role in the development of liver steatosis in diet-induced obesity. Selected candidate genes were validated by siRNA mediated knockdown in vitro and by transcriptional profiling in human liver biopsies, showing that these genes are indeed involved in the regulation of the constructed metabolic networks. Next, we applied the CoNI method to liver datasets obtained from two different weight loss interventions to identify genes that might be involved in the development of the acute hepatic steatosis observed by exendin-4 treatment compared to CR-induced weight loss. Two genes with so far unknown function in the regulation of liver metabolism were identified along with one gene involved in oxidative stress. These candidate genes await further characterisation and validation but they could be potential new targets for the treatment of acute hepatic steatosis.

Overall, the CoNI method allowed us to identify genes regulating metabolic networks in livers of obese mice or of mice undergoing weight loss interventions that would not have been detected by analysing solely the transcriptome dataset. The two validation approaches demonstrated that our fully data-driven versatile method can be used as a flexible and solid tool for multiple omics data integration and interpretation.

5 OUTLOOK

5.1 TIME-MATCHED ANALYSIS OF DNA ADDUCT FORMATION AND EARLY GENE EXPRESSION IN METHYLAZOXYMETHANOL ACETATE TREATED EKER RATS

In further experiments, it would be interesting to additionally quantify the mRNA levels of kidney cortices of Eker rats treated with MAMAc chronically for three to six month and their controls. This would show if an increased DNA adduct accumulation leads to an activation of DNA damage repair genes. It would also be interesting to test if a higher dose of MAMAc could induce stronger transcriptional changes than the ones observed after the first day of MAMAc treatment and if a higher dose could induce Ogg1 as observed by others either on mRNA or on protein level.

5.2 EFFECTS OF WEIGHT LOSS INTERVENTIONS ON HYPOTHALAMIC GENE EXPRESSION

Further analyses are needed to gain a deeper insight into the molecular mechanisms of the observed fasting-induced genes that were differentially expressed between the two weight loss groups. Understanding their role in hypothalamic leptin resistance and leptin resensitisation will help to define them as potential targets for obesity treatment. To overcome the limitations of the possibly hidden neuron-specific effects of the weight loss interventions, single-cell sequencing instead of whole tissue RNA-Sequencing could be performed in our diet setting.

5.3 ESTIMATING GENETIC IMPACT ON METABOLIC NETWORKS

In a future analysis it would be very interesting to compare the CoNI networks of Ex4-treated and HFD-fed mice to gain a better understanding of the weight loss induced increase in hepatic lipid storage. With CoNI we could elucidate, which genes and metabolites drive the differences in lipid metabolism in these two settings. By identifying the gene-metabolite interactions driving the chronic versus acute lipid storage, we could unravel novel genes involved in the aetiology and reversal of hepatic steatosis. Our work will thus help to delineate the molecular underpinnings of hepatic steatosis and advance our search for novel, druggable targets.

A newly developed method such as CoNI can always be improved. Currently, zero- and first-order correlations are used to infer the networks of interest, however recently researchers in the field suggested to additionally use higher order correlations in order to reduce the number of false positive inferred network edges. Another suggestion for improvement is to use a more sophisticated method to infer the local regulator genes. Therefore, one could use nearest neighbour algorithms instead of selecting the next two adjacent nodes to define the sub-graph for binomial testing.

The biggest asset of CoNI is that the method is not restricted to be used with transcriptome and metabolome datasets. It can be applied to any types of data that interact, such as proteome and metabolome data, or lipidome and proteome or transcriptome datasets, to name only a few. In contrast to many other approaches our method is fully unsupervised and independent from additional prior knowledge. Its straightforward but flexible setting allows for multiple applications, yet the produced networks may convey complex interactions but are easy to interpret. Overall, the method holds great potential and could be applied widely to solve research questions in the life science field and beyond.

A

APPENDIX

A.1 SUPPLEMENTARY TABLES - PART 1

Table A1: Differentially expressed genes in Eker control rats altered over the sampled time period.

ARRAY ID	DAY 3 VS DAY 1	DAY 7 VS DAY 3	DAY 14 VS DAY 7	AVERAGE EXPRESSION	P-VALUE	ADJUSTED P-VALUE	GENE SYMBOL
1387874_at	-1.58254441	0.15477784	-0.089046160	3.962288	9.910849e-09	0.0001578104	*Dbp*
1386981_at	0.55123695	0.23285023	-0.196684360	6.633349	3.690957e-08	0.0002938556	*Slc16a1*
1390430_at	-1.15480677	0.14347506	0.279337088	4.481322	1.736862e-07	0.0009218687	*Nr1d2*
1368511_at	-0.82964762	0.06112083	-0.096923987	3.487948	5.709765e-07	0.0017658306	*Bhlhe41*
1370510_a_at	0.84165994	0.05151621	-0.221614758	4.062811	6.323307e-07	0.0017658306	*Arntl*
1370816_at	-1.70367781	0.40576752	0.302570740	3.800409	6.653886e-07	0.0017658306	*Nr1d1*
1377407_at	0.12161954	-0.48986204	-0.016501745	4.175413	1.025823e-05	0.0217476716	*Acsm5*
1370541_at	-1.16554705	0.15757585	0.316446702	4.168719	1.109277e-05	0.0217476716	*Nr1d2*
1373043_at	0.43304157	-0.29205910	0.026974837	5.228216	1.229222e-05	0.0217476716	*Sdf2l1*
1372001_at	0.08514516	-0.15604194	-0.254101435	4.404160	1.546960e-05	0.0230393671	*Fam96b*
1373526_at	0.02362696	-0.36520402	-0.297851006	4.196020	1.658932e-05	0.0230393671	*Atg5*
1390592_at	-0.02505859	-0.40258280	0.169960340	4.611800	1.736309e-05	0.0230393671	NA
1387146_a_at	-0.18662373	-0.12054372	-0.082185825	4.148919	2.183763e-05	0.0251029547	*Ednrb*
1371202_a_at	-0.09094477	-0.30095323	0.196770191	5.345673	2.207130e-05	0.0251029547	*Nfib*
1387041_at	0.31535451	0.07735984	0.010541916	5.143865	2.644404e-05	0.0274863665	*Ubqln1*
1374585_at	0.24595674	0.18191751	-0.113824844	5.493505	3.198468e-05	0.0274863665	*Echdc1*
1372452_at	0.64832656	0.11253516	-0.109554926	4.512428	3.233919e-05	0.0274863665	NA
1389355_at	0.35598214	-0.05021318	-0.060883681	4.918948	3.277802e-05	0.0274863665	*Ier5*
1372151_at	-0.06364361	0.37894543	0.014122327	3.662133	3.496133e-05	0.0274863665	*Tcerg1*
1383160_at	0.62450218	-0.04330715	0.048439979	5.099766	3.549169e-05	0.0274863665	*Chordc1*
1369971_a_at	-0.56249889	0.09082476	0.044863860	4.300693	3.625031e-05	0.0274863665	*Hnrnpd*
1370174_at	-0.18186744	-0.21854631	-0.078942299	3.720235	4.118148e-05	0.0298060295	*Ppp1r15a*
1372810_at	-0.38763046	0.00840807	0.157007535	4.858478	4.574578e-05	0.0307010787	*Hnrnpdl*
1373158_at	-0.46488190	0.15331491	0.025881926	4.269187	4.627431e-05	0.0307010787	*Gpr146*
1371310_s_at	0.39338017	-0.01820374	-0.094340960	6.021709	5.206520e-05	0.0331613657	*Serpinh1*
1367919_at	0.04883862	-0.66449388	0.400321722	3.346442	5.884534e-05	0.0360382451	*Nup210*
1370283_at	0.56286208	-0.18084780	-0.024865945	6.447950	7.143060e-05	0.0421255341	*Hspa5*
1376518_at	0.43447812	-0.15428209	0.004062335	5.066327	7.418710e-05	0.0421886145	*Palm3*
1370411_at	-0.97898571	1.03864678	-0.003437599	3.946266	7.756465e-05	0.0425883436	*Trpc1*
1370381_at	0.33519109	0.22669395	-0.049853961	5.145883	8.432654e-05	0.0430429623	*Pnrc1*
1380854_at	0.03301048	0.32182391	-0.033223947	5.447217	8.630849e-05	0.0430429623	*Vegfb*
1371132_a_at	-0.41741323	-0.28982917	0.052586158	4.234037	8.650222e-05	0.0430429623	*Ank3*
1390027_at	-0.23098803	-0.10861270	0.072813193	5.341700	9.774847e-05	0.0461109851	*Usp8*
1369926_at	0.03228855	-0.31997363	0.102582296	8.494349	9.845968e-05	0.0461109851	*Gpx3*
1380466_at	-0.12082219	0.77545102	-0.320862929	2.918114	1.087430e-04	0.0494718641	NA

Table A2: Differentially expressed genes in Eker rats treated with MAMAc altered over the sampled time period.

ARRAY ID	DAY 3 VS DAY 1	DAY 7 VS DAY 3	DAY 14 VS DAY 7	AVERAGE EXPRESSION	P-VALUE	ADJUSTED P-VALUE	GENE SYMBOL
1370541_at	-1.21507486	-0.23929222	-0.16177233	3.7250703	4.921253e-12	7.836111e-08	*Nr1d2*
1370816_at	-1.43239943	0.04222790	-0.15189020	3.1388813	8.865305e-08	7.058113e-04	*Nr1d1*
1390430_at	-1.01694965	-0.36046743	-0.08782959	3.8004326	1.031580e-06	5.475285e-03	*Nr1d2*
1374625_at	-0.46250200	0.12783384	-0.11794345	6.0900713	1.402006e-06	5.557025e-03	*Hes6*
1370245_at	-0.35511128	-0.09524171	-0.14332263	6.5318151	1.744968e-06	5.557025e-03	*Ctsl*
1387874_at	-0.96508789	-0.01245769	0.20371119	3.8260615	4.542814e-06	1.205587e-02	*Dbp*
1373309_at	-0.90716791	-0.12718137	-0.06702598	4.7548697	5.412570e-06	1.231205e-02	*Tmem86a*
1368488_at	0.88186296	-0.19276158	0.16966216	4.1323764	6.481672e-06	1.290096e-02	*Nfil3*
1372911_at	-0.52840392	-0.24752108	0.11297671	5.3899284	7.684729e-06	1.359599e-02	*Mthfr*
1389355_at	0.26621215	0.20148786	-0.28398546	5.0457483	1.019493e-05	1.481846e-02	*Ier5*
1398819_at	0.37511365	-0.08899514	0.03674650	6.5239087	1.073314e-05	1.481846e-02	*Dnaja1*
1388722_at	0.45948823	-0.03271151	0.03183270	5.0735160	1.116759e-05	1.481846e-02	*Dnajb1*
1368992_a_at	-0.58905681	-0.05716769	0.27510468	5.7349572	1.241085e-05	1.520138e-02	*Srsf5*
1367771_at	0.54341284	0.01545127	0.06290372	5.5408550	2.059624e-05	2.342528e-02	*Tsc22d3*
1383160_at	0.49162006	-0.12003771	0.02431758	5.1405705	2.547648e-05	2.704413e-02	*Chordc1*
1389514_at	0.31344973	1.41180211	-1.05189912	0.6628551	3.466012e-05	3.449331e-02	*Lingo1*
1370610_at	-0.26291513	-0.33158429	0.58069897	6.0377897	3.981318e-05	3.557691e-02	*Slc34a1*
1376135_at	-0.37171332	-0.21766027	0.38231389	4.4135759	4.062696e-05	3.557691e-02	*Dars2*
1372823_at	-0.08807031	-0.48951666	0.45908666	3.7366102	4.333652e-05	3.557691e-02	*Hnrnpu*
1388010_at	0.02867063	-0.18576523	1.10897219	2.0240366	4.468619e-05	3.557691e-02	*Parpbp*
1387513_at	-0.21167755	-0.31821092	0.01065350	3.0425292	5.543694e-05	4.203440e-02	*Cyth3*
1368542_at	1.72643342	0.07225291	-0.12957708	2.5678892	6.546719e-05	4.738337e-02	*Zfp423*

Table A3: 18 KEGG pathways enriched by genes altered after the first day of MAMAc exposure.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	GENE ID
rno05212	Pancreatic cancer	6/42	65/7887	1.030077e-06	0.0001648124	*Tgfbr2/Smad3/Tgfbr1/Smad4/Egfr/Mapk9*
rno05210	Colorectal cancer	5/42	64/7887	2.026426e-05	0.0016211406	*Tgfbr2/Smad3/Tgfbr1/Smad4/Mapk9*
rno05200	Pathways in cancer	10/42	400/7887	3.431822e-05	0.0016467467	*Tgfbr2/Gnai3/Smad3/Tgfbr1/Hsp90ab1/Smad4/Nfkb2/Prkacb/Egfr/Mapk9*
rno04520	Adherens junction	5/42	74/7887	4.116867e-05	0.0016467467	*Tgfbr2/Smad3/Tgfbr1/Smad4/Egfr*
rno04010	MAPK signaling pathway	8/42	259/7887	5.432370e-05	0.0017383584	*Tgfbr2/Tgfbr1/Pak2/Nfkb2/Prkacb/Egfr/Mapk9/Cacng4*
rno04068	FoxO signaling pathway	6/42	136/7887	7.412887e-05	0.0019767697	*Tgfbr2/Smad3/Tgfbr1/Smad4/Egfr/Mapk9*
rno04933	AGE-RAGE signaling pathway in diabetic complications	5/42	104/7887	2.089716e-04	0.0047749974	*Tgfbr2/Smad3/Tgfbr1/Smad4/Mapk9*
rno05142	Chagas disease (American trypanosomiasis)	5/42	107/7887	2.387499e-04	0.0047749974	*Tgfbr2/Gnai3/Smad3/Tgfbr1/Mapk9*
rno04971	Gastric acid secretion	4/42	74/7887	6.103180e-04	0.0108048546	*Gnai3/Kcnj15/Kcnj1/Prkacb*
rno05220	Chronic myeloid leukemia	4/42	76/7887	6.753034e-04	0.0108048546	*Tgfbr2/Smad3/Tgfbr1/Smad4*
rno04350	TGF-beta signaling pathway	4/42	86/7887	1.075586e-03	0.0156448934	*Tgfbr2/Smad3/Tgfbr1/Smad4*
rno04914	Progesterone-mediated oocyte maturation	4/42	90/7887	1.274468e-03	0.0169929016	*Gnai3/Hsp90ab1/Prkacb/Mapk9*
rno04915	Estrogen signaling pathway	4/42	96/7887	1.619159e-03	0.0199281077	*Gnai3/Hsp90ab1/Prkacb/Egfr*
rno04724	Glutamatergic synapse	4/42	115/7887	3.132658e-03	0.0358018074	*Gnai3/Shank3/Prkacb/Slc38a2*
rno05166	HTLV-I infection	6/42	293/7887	4.242244e-03	0.0448883947	*Tgfbr2/Smad3/Tgfbr1/Smad4/Nfkb2/Prkacb*
rno04510	Focal adhesion	5/42	206/7887	4.488839e-03	0.0448883947	*Cav2/Pak2/Itga1/Egfr/Mapk9*
rno05230	Central carbon metabolism in cancer	3/42	65/7887	4.877423e-03	0.0459051600	*Pdha1/Egfr/Pfkl*
rno04380	Osteoclast differentiation	4/42	134/7887	5.409953e-03	0.0480884687	*Tgfbr2/Tgfbr1/Nfkb2/Mapk9*

Table A4: Enrichment statistics on DNA repair pathways. Up-regulation of genes is indicated in red and down-regulation in blue.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	GENE ID
GO:0043065	positive regulation of apoptotic process	8/69	572/17324	0.001913542	0.028174386	*Smad3*/*Tgfbr1*/*Slc11a2*/*Pak2*/*Itga1*/*Ppp1r15a*/*Ip6k2*/*Mapk9*
GO:0043066	negative regulation of apoptotic process	7/69	861/17324	0.05482791	0.15751587	*Smad3*/*Kcnj1*/*Tgfbr1*/*Hsp90ab1*/*Cebpb*/*Pak2*/*Egfr*
GO:0045786	negative regulation of cell cycle	4/69	363/17324	0.056700575	0.159997354	*Smad3*/*Ctdsp2*/*Prkacb*/*Egfr*
GO:0045787	positive regulation of cell cycle	2/69	300/17324	0.336258954	0.391340374	*Egfr*/*Rdx*
GO:0051726	regulation of cell cycle	8/69	891/17324	0.024806476	0.104748855	*Smad3*/*Cav2*/*Pak2*/*Ctdsp2*/*Prkacb*/*Egfr*/*Rdx*/*Rnf4*

A.2 SUPPLEMENTARY TABLES - PART 2

Table A5: Normalised protein abundances extracted with the four evaluated antibodies. Abundances of proteins, for which only one sample showed a spectral count larger than 0 and proteins that were not measured in either the IgG control or all 4 antibody samples are not shown (Section 2.5.2).

PROTEIN	PEPTIDE COUNT	UNIQUE PEPTIDES	IGG	PBSAB	PBSAB3P	COMAB	COMAB3P
Rps6	1	1	171	10932	4633	4900	8370
Rpl13	4	4	2726	73799	33776	53771	113463
Add1	2	2	650	14172	12798	1262	99
Tubb5	30	3	91538	8642670	8666394	568941	192282
Sf3a1	4	4	3958	103822	83532	137577	9413
Puf60	5	5	490358	382299	385255	118594	704173
Camk2b	3	2	7245	103504	89018	323	3930
Hdgfrp2	10	10	1158	27498	10817	677524	915940
Rps11	2	2	586	53772	10737	44605	85136
Rpl8	3	3	1167	75759	28449	20600	80696
Grm3	2	2	128	68647	70758	1567	376
Rps5	2	2	1603	57803	25497	5906	30884
Slc1a2	1	1	138	28907	28864	3850	102
Slc1a3	2	2	203	126024	118263	5912	1089
Krt14	13	2	57199	16722	12522	37672	27540
Krt13	9	2	21787	20106	1445	27821	17510
Hnrnpa0	1	1	127	29157	30468	1049	1658
Rps18	4	4	5638	109286	32526	39127	107640
Rpl30	1	1	231599	237816	227585	93833	188571
Snap47	3	3	87300	137414	84904	586533	114801
Slc25a11	3	3	172	139465	123189	8531	298
Hspa8	17	14	112411	820669	695373	30863	241906
Cd5l	6	6	115094	124169	123103	13911	180945
Slc25a5	5	3	6711	304552	373175	9817	13178
Krt42	8	2	381841	161405	127995	740660	214403
Rpl19	2	2	41256	92030	42521	76292	126007
Hist1h2bc	2	2	3981	59051	36203	3515	7068
Prpf8	9	9	4269	198143	151962	11150	1867
Dync1h1	5	5	535	44066	39009	7824	1419
Brsk2	2	2	182	22822	23176	271	506
Ywhah	3	2	5529	22746	24250	222	2398
Cand1	3	3	1510	89829	87837	9135	3313
Snrpd3	2	2	7603	136268	122264	14654	8633
Hcn2	13	13	16488	1133451	1085675	42520	26598
Gnb2l1	4	4	23028	183208	82559	355114	300488
Eftud2	5	5	761	162522	125529	6384	351
Pnn	3	3	1422	17077	14818	2689	2749
Erh	2	2	6912	123936	124642	406	17661
Nefm	7	6	104353	87382	91395	24860	5654
Nefl	6	5	46460	28673	27142	65051	1462
Sucla2	2	2	208	22101	22330	10441	270
Sub1	1	1	171123	173293	176353	127582	217333
Ywhaz	5	3	17875	302930	303759	3604	17058
Dnm1l	3	3	1083	76388	65690	16509	622
Fyttd1	3	3	8474	100706	77235	23911	18654
Krt5	16	12	1167768	939367	650285	3061975	1141326
Krt2	8	5	3373373	3113707	2150204	14645665	3193835
Krt6b	10	1	234222	38824	57030	234125	118366
Krt1	9	5	5852041	5704066	4050988	37481288	5935134
Krt79	5	2	3006820	2027603	1384647	3506245	2419241
C3	2	2	25436	33231	33843	16	64864

Normalised protein abundances extracted with the four evaluated antibodies.

PROTEIN	PEPTIDE COUNT	UNIQUE PEPTIDES	IGG	PBSAB	PBSAB3P	COMAB	COMAB3P
Mapre2	2	2	355	184542	121877	10	37320
Rps14	3	3	9211	240093	103022	181921	268079
Fth1	7	7	5156	260008	257678	749	3993
Sf3b2	2	2	7979	71807	63347	9397	27818
Rps26	2	2	1515	93750	27882	19979	87177
Eif4a3	6	6	15688	173600	132562	15879	17189
Sf3b1	10	10	13399	247454	214304	7387	13103
Hspa5	11	8	19447	433368	390207	109191	34142
Kcnip3	2	2	22517	87132	72001	25083	19085
Raly	7	7	16111	299897	234374	36166	9880
Rbm39	2	2	28254	30552	26786	7031	39468
Fcgr1	6	6	467007	570118	617778	163642	1005440
Zranb2	3	3	49639	44140	34457	54137	75974
Atad3a	2	2	7996	16995	17822	5123	7243
Alb	4	4	125134	65233	125823	212888	338923
Arl6ip4	3	3	474247	383692	341769	462749	692207
Rpl6	4	4	1119	62929	24627	158392	80404
Ndufa4	3	3	6461	284886	333873	7259	13485
Prss1	2	1	88085898	46742238	59577187	530046180	96815553
2210010C04Rik	4	4	1020252	422881	597601	12333020	910140
Rps3	7	7	19971	443453	142669	120010	484092
Rps4x	5	5	1412	162714	44299	14609	120435
Slc25a4	6	4	13271	1022350	1070471	7562	19239
Tmprss13	1	1	106218	42365	60616	1450314	86955
Rpl4	3	3	5268	108167	46812	86906	159396
Map4	4	4	596	50913	47331	3181	579
Rpsa	4	4	9893	346355	132600	147351	380897
Krt75	9	2	261753	241082	143407	616496	241036
C1qc	2	2	292532	88174	78934	51761	37309
H2afz	1	1	4822	31965	22674	9860	7359
Sv2a	2	2	1105	36007	39467	4002	463
Snrpd2	2	2	540	93679	80109	701	2290
Dhx9	14	14	19062	645605	583634	110159	12918
Arglu1	2	2	58527	56499	60441	24131	85212
Brsk1	3	3	424	59408	48831	8195	324
Atp6v1h	3	3	73610	153641	144392	12927	123569
Ina	7	6	263859	106382	101689	29860	12188
Rpl21	2	2	730	22652	9951	9714	39834
Tubb4b	33	1	288389	1337358	1342820	22762	297388
Eef1a1	3	3	25430	250451	264141	143675	30842
Alg2	5	5	5349	176571	231536	16622	2072
Bclaf1	4	4	21876	115062	101825	8702	8348
Sf3b6	1	1	1469	54289	50734	19046	7291
Rbm8a	3	3	440	145302	108873	10161	2363
Hacd3	3	3	1450	114850	116379	799	127
Sf3b3	5	5	19857	136295	96730	1876	3793
Sptan1	9	9	264590	126549	122574	3146	109775
Rpl10a	1	1	3329	53354	25223	47	62356
C1qa	2	2	216238	40622	40203	1716	299
Hnrnph2	3	3	2796	246938	221166	24609	2570
Ywhag	5	4	11694	245918	241741	538	7761
Mbp	4	4	595904	3305362	2729441	375958	771736
Dsg1c	1	1	16820	6677	4837	84814	8367
Luc7l2	2	2	24807	19031	17512	10823	32130
Pura	8	8	13902	1058492	801018	10406	58355
Sfxn3	2	2	2855	26859	34860	30951	3507
Tubb2a	36	3	77793	6771981	6443899	1050783	150612
Dhx30	1	1	266180	178610	227100	36219	478204
Srsf11	1	1	50968	34292	29560	9775	78431

Normalised protein abundances extracted with the four evaluated antibodies.

PROTEIN	PEPTIDE COUNT	UNIQUE PEPTIDES	IGG	PBSAB	PBSAB3P	COMAB	COMAB3P
Dnm3	5	3	5213	60535	66541	4863	535
Gfap	19	17	3034361	118003	95747	440933	128954
Phgdh	4	4	712	103720	99966	7580	163
Purg	1	1	51	51465	26684	168	1824
Krt73	6	1	30934	36229	3416	6548	31764
Khdrbs1	1	1	200	4862	4679	66	354
Hspa12a	3	3	1413	85263	85699	33566	4548
C4b	4	4	153090	151624	169482	2605	246676
Tubb3	27	5	55561	2139088	2013484	33360	51741
Tubb4a	32	7	82059	6188544	5674744	267010	95357
Nf1	6	6	5174	86248	78573	60358	29021
Actg1	9	9	1431630	871828	801262	264522	180884
Try10	4	3	5535403	5868385	7470256	28996780	4784142
Rpl18	3	3	847	121988	52126	47736	156923
Rps7	2	2	1784	61883	28366	28033	101531
Slc25a3	2	2	41335	36369	35983	50870	22145
Dnm1	6	4	8934	212877	242279	31452	13470
Kcnd3	4	4	11450	119004	127396	1403	3026
Tuba4a	21	4	74074	5680273	5978744	122768	142072
Rpl17	2	2	32474	54908	19150	119481	67521
Gm10036	2	2	3207	120194	43004	16413	147888
Rps25	2	2	25330	137533	47183	61004	126992
Rps17	1	1	2563	28936	12014	23949	30878
Thrap3	7	7	2946	97752	90883	23319	4167
Kcnd2	8	8	18637	372070	337493	20710	14279
Rplp0	4	4	2285	216453	99481	79514	248152
Igj	1	1	51066	57786	55211	20054	118642
Kcnip4	2	2	11065	104078	91107	27956	18313
Krt77	6	2	303001	291042	202795	614378	333069
Srrm2	3	3	38951	37117	30649	893	55335
Rpl7a-ps3	4	4	2542	124122	44551	180274	129989
Ssb	7	7	13	165270	98259	1670	1118
Srsf2	2	2	50264	43165	37469	128587	59668
Rpl27	2	2	260	56321	21525	24085	75140
Sugp2	7	7	8270	276272	209206	102435	10328
Gm9755	8	8	14183	313809	253540	85044	21866
Rps2	3	3	2246	242230	85733	39954	254906
Tuba1a	22	1	8232	113624	119044	4216	4481
Mb21d2	18	18	5158753	7063647	6178381	4583115	14830306
Krt76	5	1	2177455	1972762	1233207	7438566	1832451
Gh	2	2	661	18364	14898	3838	229
Nsf	14	14	24432	633103	652703	73106	55930
Cnp	8	8	42959	426759	420362	12526	30804
Krt10	10	4	8550436	6074820	5536775	33678307	7271377
Rpl23	2	2	967	31564	9164	26043	32163
Snrnp200	13	12	5024	259174	202262	21846	7828
Rpl23a	2	2	7387	141067	36432	191661	152253
Rps8	3	3	2529	228233	53693	258453	209935
Camk2a	5	4	35369	263671	242438	14107	14744
Rnh1	7	7	5671453	4848460	4555582	4174485	5008795
Rps15a	1	1	214	34990	11454	13668	46814
Rps19	5	5	34906	337164	151805	1753369	374915
Hnrnpc	13	12	47787	3286420	2713011	80659	41918
Rpl9	3	3	30604	177773	64422	320746	216608
Rps28	2	2	1684	141429	56076	4549	224383
Rps10	1	1	1150	68818	20506	13295	58207
Dmxl2	4	4	1617	81519	73626	2502	3843
Dsp	5	4	77006	58899	61944	389542	88662
Myo5a	2	2	7300	14659	10640	2098	144

Normalised protein abundances extracted with the four evaluated antibodies.

PROTEIN	PEPTIDE COUNT	UNIQUE PEPTIDES	IGG	PBSAB	PBSAB3P	COMAB	COMAB3P
Rpl12	3	3	1694	126998	49120	1073	151817
Rps20	1	1	7345	87954	31469	30113	101218
Slc25a12	15	15	29725	964738	921151	125333	36925
Pcmt1	4	4	233239	245468	255965	36448	452421
Matr3	22	22	114179	2431779	1970956	171552	59785
Krt78	3	3	4290233	4720040	3059204	7336781	4257667
Snrpe	1	1	3379	48301	43579	21001	5458
Rpl14	1	1	986	58138	17270	22122	76643

Table A6: Number of differentially expressed genes for pairwise diet/intervention comparisons.

COMPARISON	TOTAL # OF DE GENES	# UP-REGULATED GENES	# DOWN-REGULATED GENES
HFD vs Chow	10	8	2
H>C vs HFD	1319	468	851
H>C vs Chow	501	138	318
CR vs Chow	1728	828	900
CR vs HFD	2691	1342	1349
CR vs H>C	119	102	17
Ex4 vs Chow	2142	1113	1029
Ex4 vs HFD	3079	1617	1462
Ex4 vs H>C	230	132	98
Ex4 vs CR	29	2	27

Table A7: Log2 fold changes of neuropeptides compared between weight loss intervention and control groups. Adjusted p-values of the corresponding test are given in brackets

GENE	MEAN EXPRESSION	EX4VSCHOW	EX4VSHFD	EX4VSHC	CRVSCHOW	CRVSHFD	CRVSHC
Npy	581.85	0.449 (0.004)	1.034 (1.46E-12)	0.790 (3.54E-19)	0.619 (1.86E-04)	1.204 (8.59E-20)	0.960 (5.24E-31)
Agrp	145.06	1.072 (8.41E-05)	1.691 (3.46E-11)	1.512 (1.99E-23)	1.094 (1.47E-04)	1.713 (9.07E-13)	1.534 (3.00E-26)
Bdnf	628.26	-0.105 (0.477)	0.040 (0.813)	-0.051 (0.678)	-0.010 (0.959)	0.134 (0.316)	0.043 (0.815)
Cartpt	2136.75	0.132 (0.504)	-0.186 (0.305)	-0.248 (0.046)	0.058 (0.797)	-0.259 (0.115)	-0.321 (0.008)
Glp1r	76.81	0.603 (0.016)	0.366 (0.162)	0.167 (0.47)	0.423 (0.098)	0.185 (0.527)	-0.013 (0.979)
Grm1	1946.30	0.087 (0.405)	0.140 (0.137)	0.030 (0.745)	0.099 (0.320)	0.152 (0.091)	0.042 (0.720)
Lepr	100.26	0.029 (0.938)	-0.090 (0.775)	0.008 (0.980)	0.006 (0.987)	-0.113 (0.711)	-0.015 (0.975)
Mc3r	203.49	0.245 (0.105)	0.177 (0.258)	-0.004 (0.986)	0.169 (0.283)	0.101 (0.557)	-0.080 (0.666)
Pomc	1215.88	-0.793 (0.226)	-1.670 (0.003)	-1.167 (0.007)	-0.438 (0.543)	-1.315 (0.021)	-0.812 (0.149)
Trmt1	893.88	-0.177 (0.196)	-0.208 (0.109)	-0.079 (0.490)	-0.091 (0.553)	-0.122 (0.385)	0.007 (0.976)

Table A8: Log2 fold changes of genes differentially expressed between Ex4 and CR. Adjusted p-values of the corresponding test are given in brackets

GENE	MEAN EXPRESSION	EX4VSCR	EX4VSCHOW	EX4VSHFD	EX4VSHC	CRVSCHOW	CRVSHFD	CRVSHC
Daam2	2100.31	-0.226 (0.04)	-0.258 (0.04)	-0.256 (0.03)	-0.009 (0.95)	-0.032 (0.86)	-0.03 (0.86)	0.217 (0.03)
Mertk	276.27	-0.318 (4.80E-02)	-0.143 (0.51)	-0.061 (0.80)	0.152 (0.32)	0.175 (0.39)	0.257 (0.16)	0.47 (1.90E-05)
Hif3a	209.79	-0.500 (0.03)	-0.451 (0.11)	-0.251 (0.41)	0.383 (0.06)	0.049 (0.90)	0.249 (0.40)	0.883 (6.00E-09)
Cdh19	169.82	-0.434 (0.03)	-0.541 (0.02)	-0.501 (0.03)	-0.052 (0.84)	-0.107 (0.73)	-0.067 (0.83)	0.382 (0.046)
Prodh	336.67	-0.284 (4.00E-03)	-0.072 (0.67)	-0.055 (0.75)	0.084 (0.48)	0.212 (0.12)	0.229 (0.08)	0.369 (3.79E-06)
Zbtb16	277.99	-0.648 (1.00E-03)	0.137 (0.71)	0.482 (0.10)	0.159 (0.54)	0.785 (4.00E-03)	1.129 (2.37E-05)	0.806 (1.84E-06)
Adipor2	2041.16	-0.228 (0.01)	0.024 (0.88)	0.101 (0.43)	0.081 (0.41)	0.252 (0.02)	0.328 (1.00E-03)	0.308 (4.31E-06)
Cpm	318.24	-0.284 (0.03)	-0.068 (0.74)	0.053 (0.80)	0.004 (0.98)	0.217 (0.17)	0.337 (0.02)	0.289 (0.01)
Anln	1214.09	-0.291 (0.02)	-0.086 (0.65)	0.040 (0.85)	0.067 (0.64)	0.205 (0.19)	0.331 (0.02)	0.358 (2.00E-04)
Arhgef10	1328.44	-0.226 (8.00E-03)	0.013 (0.94)	0.082 (0.53)	0.018 (0.89)	0.239 (0.02)	0.307 (2.00E-03)	0.243 (7.00E-04)
Sgk1	2762.78	-0.403 (3.80E-05)	0.002 (0.99)	0.268 (0.08)	0.297 (6.00E-03)	0.405 (6.00E-03)	0.671 (4.05E-06)	0.7 (2.84E-18)
Arl4d	356.10	-0.537 (3.00E-04)	-0.260 (0.28)	0.056 (0.85)	0.191 (0.30)	0.277 (0.23)	0.593 (3.00E-03)	0.728 (5.05E-09)
Arrdc2	366.24	-0.381 (2.00E-03)	-0.104 (0.63)	0.152 (0.45)	0.398 (1.00E-03)	0.276 (0.12)	0.533 (9.00E-04)	0.778 (2.28E-17)
Plin4	248.46	-1.254 (2.64E-07)	-0.249 (0.63)	0.179 (0.74)	1.063 (2.00E-04)	1.005 (0.01)	1.433 (2.00E-04)	2.317 (8.58E-28)
Sult1a1	115.22	-0.554 (1.00E-03)	0.134 (0.67)	0.220 (0.44)	0.594 (7.00E-04)	0.688 (4.00E-03)	0.774 (6.00E-04)	1.147 (1.72E-18)
Agxt2l1	1352.61	-0.389 (0.03)	0.047 (0.88)	0.233 (0.31)	0.227 (0.18)	0.436 (0.03)	0.622 (1.00E-03)	0.616 (5.78E-07)
Upp2	589.10	0.208 (0.02)	0.562 (2.35E-08)	0.436 (1.36E-05)	0.347 (1.66E-07)	0.354 (8.00E-04)	0.228 (0.02)	0.140 (.17)
Irf7	53.44	0.526 (9.00E-03)	0.315 (0.26)	0.100 (0.76)	0.146 (0.51)	-0.210 (0.49)	-0.426 (0.09)	-0.379 (0.10)
Hr	1038.00	-0.557 (2.82E-06)	-0.313 (0.12)	0.039 (0.89)	0.215 (0.16)	0.244 (0.23)	0.596 (7.00E-04)	0.772 (1.18E-13)
Cdkn1a	346.56	-0.632 (8.61E-09)	-0.325 (0.09)	-0.112 (0.63)	0.228 (0.13)	0.306 (0.11)	0.520 (3.00E-03)	0.860 (7.09E-18)
Mfsd2a	339.07	-0.614 (9.61E-06)	-0.484 (0.03)	-0.375 (0.09)	-0.037 (0.89)	0.129 (0.64)	0.238 (0.31)	0.577 (1.82E-05)
Frmpd1	421.90	-0.276 (0.02)	0.029 (0.89)	0.177 (0.24)	0.068 (0.61)	0.306 (0.03)	0.454 (5.00E-04)	0.345 (1.00E-04)
Tekt4	64.36	-0.487 (0.03)	-0.584 (0.02)	-0.298 (0.29)	-0.042 (0.9)	-0.097 (0.77)	0.188 (0.52)	0.445 (0.03)
Ddit4	875.81	-0.430 (1.00E-03)	-0.144 (0.51)	-0.082 (0.73)	-0.033 (0.88)	0.286 (0.13)	0.348 (4.90E-02)	0.397 (1.00E-03)
Plekhf1	175.78	-0.424 (0.02)	-0.283 (0.22)	-0.223 (0.35)	-0.032 (0.90)	0.141 (0.59)	0.201 (0.39)	0.392 (0.02)
Spsb1	470.35	-0.229 (0.01)	-0.003 (0.99)	0.080 (0.56)	0.074 (0.47)	0.226 (0.04)	0.308 (3.00E-03)	0.303 (1.41E-05)
Pnpla2	636.17	-0.234 (1.00E-03)	0.047 (0.72)	0.111 (0.31)	0.028 (0.80)	0.281 (4.00E-03)	0.345 (2.00E-04)	0.262 (2.75E-05)
Gpr17	797.95	-0.280 (1.00E-03)	-0.389 (9.00E-04)	-0.420 (2.00E-04)	-0.431 (2.81E-09)	-0.109 (0.42)	-0.140 (0.26)	-0.151 (0.19)
Bcl6	553.84	-0.212 (0.01)	0.152 (0.17)	0.107 (0.36)	-0.024 (0.84)	0.365 (4.00E-04)	0.320 (9.00E-04)	0.189 (0.02)

Table A9: Log2 fold changes of genes differentially expressed between Ex4 and CR between fed and fasted state in AgRP neurones.

GENE	MEAN EXPRESSION	LOG2-FOLDCHANGE	P-VALUE	ADJUSTED P-VALUE
Sgk1	1630.00	2.497	7.12E-30	9.45E-27
Cdkn1a	4245.28	6.349	1.54E-16	4.75E-14
Zbtb16	1575.52	1.859	2.35E-06	6.92E-05
Upp2	195.35	-2.289	1.16E-05	2.69E-04
Spsb1	389.14	1.994	3.13E-05	6.34E-04
Ddit4	615.83	1.128	3.41E-03	0.0275
Agxt2l1	36.87	3.312	0.0086	0.0548
Arhgef10	144.86	-1.096	0.0183	0.0963
Frmpd1	96.72	-0.996	0.0191	0.0986
Arrdc2	25.09	-1.527	0.1191	0.3261
Cpm	59.72	-1.448	0.1253	0.3361
Prodh	46.47	-1.460	0.1488	0.3727
Sult1a1	33.27	1.038	0.2542	0.5090
Mertk	22.42	1.644	0.2579	0.5134
Bcl6	522.96	0.446	0.2760	0.5336
Adipor2	604.63	0.322	0.2997	0.5595
Pnpla2	629.03	0.208	0.4144	0.6663
Tekt4	15.05	1.058	0.4270	0.6774
Arl4d	0.59	-2.235	0.4667	NA
Anln	28.24	1.149	0.4704	0.7102
Irf7	56.93	-0.440	0.5822	0.7906
Gpr17	0.67	-1.683	0.5838	NA
Plin4	47.61	-0.443	0.6081	0.8082
Hif3a	20.48	0.562	0.6690	0.8465
Hr	199.48	0.047	0.9171	0.9711
Mfsd2a	5.97	0.170	0.9174	0.9713
Daam2	70.16	-2.694	NA	NA

Table A10: Log2 fold changes of genes differentially expressed between Ex4 and CR between fed and fasted state in POMC neurones.

GENE	MEAN EXPRESSION	LOG2-FOLDCHANGE	P-VALUE	ADJUSTED P-VALUE
Zbtb16	829.01	2.652	5.42E-20	2.15E-16
Bcl6	365.73	1.518	2.59E-05	3.93E-03
Cpm	26.40	-3.798	1.46E-04	0.0127
Sult1a1	35.75	-3.847	8.09E-04	0.0333
Hr	175.77	-2.129	9.13E-04	0.0355
Pnpla2	390.58	0.956	0.0153	0.1882
Spsb1	19.77	2.547	0.0296	0.2593
Anln	39.74	1.957	0.0652	0.3753
Mfsd2a	5.18	-3.174	0.0901	NA
Ddit4	417.16	0.467	0.1741	0.5812
Tekt4	47.38	-1.260	0.2663	0.6848
Arrdc2	23.01	-1.305	0.2974	0.7126
Arl4d	0.71	3.194	0.2998	NA
Arhgef10	201.01	0.385	0.3244	0.7304
Plin4	40.92	0.586	0.3721	0.7669
Gpr17	1.52	-2.331	0.4438	NA
Cdkn1a	312.81	-0.372	0.4589	0.8227
Sgk1	321.94	0.390	0.4881	0.8397
Frmpd1	35.71	-0.333	0.6127	0.8925
Adipor2	628.73	0.122	0.6298	0.8976
Irf7	69.66	-0.186	0.7580	0.9426
Agxt2l1	10.27	0.671	0.7953	0.9535
Plekhf1	0.13	-0.755	0.8113	NA
Prodh	29.92	-0.205	0.8369	0.9652
Upp2	227.93	0.055	0.8969	0.9764
Daam2	41.88	-0.046	0.9655	0.9914
Hif3a	23.70	0.054	0.9665	0.9914

A.3 SUPPLEMENTARY TABLES - PART 3

Table A11: Human primer sequences for real-time PCR.

HUMAN GENE	UPSTREAM PRIMER	DOWNSTREAM PRIMER
RPS13	5′-ccccacttggttgaagttga-3′	5′-acaccatgtgaatctctcagga-3′
ARHGAP24	5′-tgtcttgagctcccagcaa-3′	5′-tgacaaagcctccttgcttc-3′
CD82	5′-gccgacaagagcagtttcat-3′	5′-gacataggcccccatccta-3′
GK	5′-ttgattcatggcttatttggagt-3′	5′-tctgtacagtggacacctccat-3′
INHBE	5′-tcagctttgctactgtcacagac-3′	5′-cgaggagtggacaggtgaa-3′
MYC	5′-tgctccatgaggagacacc-3′	5′-cttttccacagaaacaacatcg-3′
RAPGEF4	5′-ttttatgccaaatacccagctt-3′	5′-tgaaggctgtgcgtggta-3′
SMIM13	5′-ctgactctgcttgtgttcgtg-3′	5′-agatgccatacaaaataccaacc-3′
DDX3X	5′-gctggcctagacctgaactc-3′	5′-gcttctcggttccttaaatgag-3′
COBLL1	5′-ccgagtcacctagtgccagt-3′	5′-ttcattatgtgcagagttattttcct-3′
APPL2	5′-caagcagtgactcccattacaa-3′	5′-tcattttcattttccatctctgaa-3′
TAP1	5′-ctcagggctatgacacagagg-3′	5′-acacggtttccggatcaat-3′

Table A12: Phenotype monitoring of chow- and HFD-fed mice, giving the mean and standard deviation for each measured feature.

	CHOW N=10	HFD N=8	P-VALUE
Body weight at the end of the study (g)	33.49 ± 1.56	49.2 ± 4.49	4.40E-05
Liver triglyceride levels (µg/mg)	4.13 ± 0.9	9.4 ± 3.11	4.50E-04
Plasma triglyceride levels (mg/dl)	44.09 ± 6.84	54.49 ± 8.43	1.60E-02
Plasma cholesterol levels (mg/dl)	45.74 ± 7.84	121.24 ± 18.79	4.10E-04
Plasma insulin levels (mg/dl)	0.02 ± 0.01	0.07 ± 0.06	6.20E-04

Table A13: Variance explained by phenotypic features for metabolites in the Chow/HFD setting.

METABOLITE	LMG DIET	P-VALUE DIET	ADJ. P-VALUE DIET	LMG BWE	P-VALUE BWE	ADJ. P-VALUE BWE	LMG CHOLESTEROL	P-VALUE CHOLESTEROL	ADJ. P-VALUE CHOLESTEROL	LMG TAG	P-VALUE TAG	ADJ. P-VALUE TAG
PC.aa.C34.3	0.2397331	4.31E-09	7.36E-07	0.3637867	9.65E-05	0.01678879	0.16846980	0.36934000	1	0.1719371	0.527081000	1
C14.2	0.3824772	2.47E-06	3.71E-04	0.2007075	0.162514343	1	0.12291326	0.60786798	1	0.1279165	0.917985552	1
C16.2	0.3318389	1.56E-05	2.20E-03	0.1846086	0.314900457	1	0.11465882	0.64218349	1	0.1493048	0.487416036	1
C18.2	0.3901557	8.86E-07	1.38E-04	0.2048238	0.127650983	1	0.11937203	0.25109748	1	0.1458316	0.626830635	1
Asp	0.3194756	4.65E-05	6.37E-03	0.1740213	0.319939103	1	0.15424877	0.43150864	1	0.1015451	0.249832364	1
Putrescine	0.2607054	8.35E-05	1.12E-02	0.1827139	0.927184516	1	0.13266913	0.86303834	1	0.1328082	0.836380917	1
t4.OH.Pro	0.2222667	9.56E-05	1.27E-02	0.1873886	0.651458492	1	0.18168860	0.33759564	1	0.1195831	0.778944422	1
lysoPC.a.C14.0	0.4411156	6.86E-07	1.08E-04	0.1937812	0.007522638	1	0.10759604	0.52111911	1	0.1314394	0.475029116	1
lysoPC.a.C16.1	0.3650641	2.23E-04	2.85E-02	0.1694339	0.014151719	1	0.07554781	0.85732311	1	0.1202797	0.219040861	1
lysoPC.a.C17.0	0.2843330	7.39E-07	1.16E-04	0.2681873	0.266186706	1	0.16977473	0.90677111	1	0.1395947	0.221561671	1
PC.aa.C30.0	0.3830651	1.01E-10	1.75E-08	0.2283272	0.046594043	1	0.15540033	0.88868113	1	0.1976983	0.033203621	1
PC.aa.C32.0	0.2912734	7.54E-06	1.07E-03	0.2104443	0.962157674	1	0.18477355	0.34688448	1	0.1199915	0.163641780	1
PC.aa.C32.1	0.4133186	2.34E-10	4.05E-08	0.2256981	0.010562309	1	0.15119989	0.89820360	1	0.1695124	0.252286374	1
PC.aa.C32.2	0.4722346	2.62E-08	4.38E-06	0.2063656	0.000296728	0.05133389	0.12077814	0.72825299	1	0.1269030	0.419608320	1
PC.aa.C34.1	0.3339641	1.75E-07	2.87E-05	0.2084898	0.350744410	1	0.16112915	0.53555909	1	0.1849410	0.206710021	1
PC.aa.C34.4	0.3238726	2.41E-04	3.06E-02	0.2208350	0.002071953	0.34601607	0.12234363	0.05627055	1	0.1197969	0.071057232	1
PC.aa.C36.1	0.2779443	3.98E-09	6.85E-07	0.2400864	0.190393044	1	0.18593593	0.35753964	1	0.2344774	0.028962682	1
PC.aa.C36.6	0.4267069	8.48E-09	1.44E-06	0.2108815	0.005021467	0.82352064	0.13720851	0.94982906	1	0.1578511	0.252646848	1
PC.aa.C38.0	0.3984595	6.83E-06	9.84E-04	0.2055150	0.159174987	1	0.10484757	0.24105572	1	0.1159605	0.036213180	1
PC.ae.C36.2	0.3025715	3.20E-08	5.31E-06	0.2776466	0.179721930	1	0.17348706	0.92401730	1	0.1594081	0.491738266	1
PC.ae.C36.3	0.2955483	1.00E-08	1.70E-06	0.2916777	0.062584684	1	0.17742403	0.89255367	1	0.1633029	0.432452505	1
PC.ae.C38.2	0.2812057	2.96E-07	4.79E-05	0.2745673	0.172731138	1	0.18237148	0.69453668	1	0.1426644	0.174620697	1
PC.ae.C38.3	0.3273298	2.99E-06	4.45E-04	0.2712003	0.895886184	1	0.11377866	0.06438646	1	0.1421146	0.004806265	0.8122588
PC.ae.C38.4	0.3914228	2.46E-06	3.71E-04	0.2048159	0.158422713	1	0.12936476	0.85110585	1	0.1143467	0.104516738	1
PC.ae.C38.5	0.3741524	1.88E-04	2.43E-02	0.1631596	0.106978910	1	0.07704193	0.35315828	1	0.1292784	0.018098797	1
PC.ae.C40.4	0.3831085	3.30E-06	4.85E-04	0.1895061	0.090916404	1	0.14840700	0.48343384	1	0.1240288	0.017201300	1
PC.ae.C40.5	0.3575933	1.88E-06	2.88E-04	0.2449014	0.585152874	1	0.11388839	0.18098502	1	0.1450956	0.003649595	0.6240808
PC.ae.C40.6	0.3499597	5.93E-06	8.59E-04	0.1946199	0.277241374	1	0.15430081	0.55441853	1	0.1131498	0.316907781	1
SM.C18.0	0.3067348	1.05E-06	1.62E-04	0.2074867	0.733184245	1	0.12568437	0.26411555	1	0.2180816	0.078555633	1
SM.C24.1	0.3720558	1.93E-05	2.70E-03	0.1903311	0.206921589	1	0.11456750	0.75704235	1	0.1067450	0.126299312	1

Table A14: Variance explained by phenotypic features for genes in the Chow/HFD setting.

GENE	LMG DIET	P-VALUE DIET	ADJ. P-VALUE DIET	LMG BWE	P-VALUE BWE	ADJ. P-VALUE BWE	LMG CHOLESTEROL	P-VALUE CHOLESTEROL	ADJ. P-VALUE CHOLESTEROL	LMG TAG	P-VALUE TAG	ADJ. P-VALUE TAG
Slc22a27	0.191592	7.26E-07	0.007363921	0.3375418	3.24E-06	0.03290757	0.2594746	0.006819713	1	0.1310797	0.5356962	1
Ces2a	0.2984175	1.55E-10	1.57E-06	0.2446594	0.23623412	1	0.22972	0.014173468	1	0.1893612	0.3155422	1
Aqp4	0.2939618	3.69E-07	3.74E-03	0.2111855	0.92927098	1	0.1767754	0.369768717	1	0.1928836	0.21706407	1
Arsg	0.2830303	5.50E-07	5.58E-03	0.2545769	0.31897918	1	0.1814344	0.623370419	1	0.147151	0.48756077	1
Apoa4	0.2593114	2.62E-06	2.66E-02	0.17866	0.45013905	1	0.2260362	0.02976046	1	0.1816289	0.13759269	1
Bdh1	0.238317	1.41E-06	1.43E-02	0.2244287	0.22931477	1	0.188653	0.337926236	1	0.198287	0.27475172	1
Psmb9	0.3266071	3.82E-06	3.86E-02	0.2553741	0.89627629	1	0.1259163	0.054600349	1	0.1224802	0.30086504	1
Rpl12	0.294529	1.21E-06	1.22E-02	0.2111916	0.94932714	1	0.1967805	0.227186316	1	0.1466623	0.97267174	1
Rps20	0.3705123	3.33E-06	3.37E-02	0.1928776	0.14012128	1	0.1469854	0.59900382	1	0.1173836	0.63966834	1
Gm9396	0.2351322	7.46E-07	7.56E-03	0.1944879	0.61129928	1	0.2184573	0.055137265	1	0.2217089	0.0628943	1
Cyp2b9	0.2893306	8.79E-12	8.93E-08	0.2606278	0.01602672	1	0.238335	0.003740839	1	0.1875259	0.51246048	1
Bphl	0.3345739	1.19E-06	1.21E-02	0.2515636	0.97053376	1	0.1392637	0.539787079	1	0.1293883	0.0567737	1
Khk	0.2693985	1.45E-06	1.47E-02	0.2517797	0.28296918	1	0.180004	0.615489725	1	0.1444608	0.52380387	1
Sema5b	0.2508726	2.18E-06	2.21E-02	0.221625	0.38652311	1	0.2087802	0.199961959	1	0.1560804	0.90781099	1
H6ob	0.3214939	7.71E-08	7.83E-04	0.2274201	0.88938497	1	0.1361042	0.177087394	1	0.2184438	0.06696902	1
Gm3934	0.2830693	1.77E-06	1.80E-02	0.1915237	0.62109772	1	0.1713278	0.285974181	1	0.1990761	0.1297356	1
Lgals8	0.249661	4.88E-06	4.94E-02	0.1780928	0.68922352	1	0.2059458	0.075994268	1	0.1926876	0.1308367	1
Dhrs1	0.2634845	1.39E-06	1.41E-02	0.2220795	0.48265451	1	0.1643399	0.702249364	1	0.1966663	0.27333796	1
Gstp1	0.2650882	2.89E-06	2.92E-02	0.1849026	0.68327679	1	0.1697475	0.255584377	1	0.216774	0.07101221	1

Table A15: Top 50 significantly differentially expressed genes between HFD and Chow in the liver samples.

GENE	LOG2FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Cyp2b9	10.276	8.112	17.619	4.02E-14	4.08E-10
Ces2a	-2.311	13.168	-14.084	3.21E-12	1.63E-08
H60b	-3.360	8.709	-10.833	4.31E-10	1.46E-06
Gpc1	3.269	7.083	10.340	9.92E-10	2.28E-06
Aqp4	2.678	6.201	10.270	1.12E-09	2.28E-06
Arsg	2.232	7.862	9.912	2.10E-09	3.55E-06
Sema5b	3.065	8.191	8.988	1.14E-08	1.65E-05
Fabp2	2.508	12.854	8.916	1.30E-08	1.65E-05
Ugt2b38	-2.188	9.550	-8.764	1.74E-08	1.96E-05
Apobec1	-1.720	11.600	-8.618	2.30E-08	2.22E-05
Tmem86a	1.820	7.716	8.597	2.40E-08	2.22E-05
Tsc22d1	3.812	13.052	8.516	2.81E-08	2.38E-05
Gm3934	-1.872	8.128	-8.312	4.19E-08	3.27E-05
Mettl20	-1.676	11.002	-8.238	4.85E-08	3.52E-05
Sult1c2	2.448	6.539	8.099	6.40E-08	4.33E-05
Vnn1	2.784	12.699	7.820	1.12E-07	7.13E-05
Gstp1	-1.850	16.556	-7.738	1.33E-07	7.94E-05
Bdh1	1.228	14.408	7.681	1.49E-07	8.35E-05
Aqp8	3.238	12.104	7.660	1.56E-07	8.35E-05
9030619P08Rik	2.740	7.409	7.627	1.67E-07	8.48E-05
Tjp3	1.577	6.720	7.604	1.75E-07	8.48E-05
Apoa4	2.973	17.802	7.513	2.11E-07	9.76E-05
Psmb9	1.726	8.180	7.461	2.36E-07	0.000104215
Lect2	1.514	14.095	7.278	3.46E-07	0.000137849
Rpl12	-0.942	15.692	-7.262	3.58E-07	0.000137849
Siae	1.218	12.551	7.241	3.75E-07	0.000137849
Ppp1r3c	2.440	9.062	7.236	3.78E-07	0.000137849
Asns	-3.977	7.845	-7.234	3.80E-07	0.000137849
Khk	0.945	16.295	7.169	4.36E-07	0.00015287
Slc35f2	1.839	5.804	7.064	5.46E-07	0.000184742
Serpina1e	-2.913	17.451	-6.981	6.52E-07	0.00021361
Cyp4a12b	-2.106	13.548	-6.901	7.75E-07	0.000243371
Gstp2	-1.999	13.559	-6.892	7.91E-07	0.000243371
Acmsd	-2.433	8.542	-6.862	8.43E-07	0.000251796
Ces2e	1.400	11.616	6.825	9.15E-07	0.000265089
Gm9396	-0.998	10.891	-6.812	9.39E-07	0.000265089
Slco2a1	-1.493	9.449	-6.781	1.01E-06	0.000274249
Igfbp2	-1.655	15.882	-6.772	1.03E-06	0.000274249
Cd207	-1.826	4.932	-6.720	1.15E-06	0.000299239
Cd36	2.908	10.195	6.702	1.19E-06	0.000303454
Camk2b	2.108	6.484	6.680	1.25E-06	0.000310727
Svil	1.171	8.685	6.647	1.35E-06	0.000326076
Fetub	-1.150	13.762	-6.598	1.50E-06	0.000354628
Apoc2	1.071	16.571	6.585	1.54E-06	0.000356573
Esrrg	1.495	4.756	6.573	1.59E-06	0.000357992
Selenbp2	-2.356	11.250	-6.532	1.74E-06	0.00038325
Spp2	1.124	9.089	6.522	1.78E-06	0.000383718
Slc17a4	1.899	7.699	6.489	1.91E-06	0.000404326
Tuba8	-1.578	6.232	-6.461	2.03E-06	0.000418911
Leap2	1.458	8.787	6.455	2.06E-06	0.000418911
Ebpl	1.049	11.518	6.385	2.41E-06	0.000475711

Table A16: Top 30 GO terms enriched by differentially expressed genes. Terms that had overlapping enriched gene sets were summarised.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
GO:0022613	ribonucleoprotein complex biogenesis	50/972	395/23577	1.87e-12	4.49e-09	50	*Noc4l/Gm17430/Rrn3/Nop2/Rplp0/Noct/Celf1/Ddx3x/Fbl/Grb7/Rpsa/Ncl/Rpl10a/ Rpl6/Rpl7a/Rpl7/Rps14/Rps16/Rps17/Rps18/Rps24/Rps5/Rps6/Rps7/Rps8/Utp18/Eif5/ Ranbp9/Gnl2/Rpl12/Wdr3/Rpl3/Lsm2/Gnl3/Rps28/Eif3d/Nop58/Ftsj3/Ddx21/Rpl35/ Gtf2h5/Rps21/Nop56/Gemin6/Rpl38/Srsf6/Rpf1/Wdr43/Eif4b*
GO:0044283	small molecule biosynthetic process	54/972	467/23577	8.81e-12	1.41e-08	54	*Gck/Gk/Cth/Cyb5r3/Asl/Acadm/Apoa4/Apoc2/Atf4/Car5a/Cyp7b1/Dhcr7/Egr1/ Fdft1/Fgf1/Gamt/Gnmt/Got1/Gsto1/Hmgcl/Hnf4a/Hyal1/Mvk/Pck1/Pcx/Pklr/ Pltp/Por/Mvd/Ptk2b/Rbp1/Rdh16/Acsm3/Scd2/Shmt1/Srebf1/Hmgcs1/Slc45a3/ Tcf4/Gls2/Cps1/Decr2/Acmsd/Gulo/Asns/Aass/Slc35b4/0610007P14Rik/Moxd1/ Bhmt2/Dgat2/Elovl5/Oxsm/Fads1/Adck4*
GO:0019216	regulation of lipid metabolic process	40/972	306/23577	1.17e-10	1.13e-07	40	*Adora1/Apoa4/Apobec1/Apoc2/Cd36/Cd81/Cidea/Cyp17a1/Dhcr7/Egr1/Fgf1/ Gk/Hnf4a/Id2/Cyr61/Il1a/Ldlr/Pcx/Pdgfa/Por/Ptk2b/Rora/Sorbs1/Srebf1/ Slc45a3/Tcf4/Thrsp/Sf1/Abcd2/Mettl20/Nr1d2/Irs2/Acsl4/Fgf21/Smpd3/Dgat2/ Elovl5/Crtc3/Tysnd1/Gpld1*
GO:0006631	fatty acid metabolic process	43/972	355/23577	2.77e-10	2.22e-07	43	*Gpat4/Acat1/Acadm/Acads/Abcd1/Apoa4/Apoc2/Cd36/Cyp2b10/Cyp2b9/Cyp2c38/ Cyp2c39/Fabp2/Gk/Hadh/Hnf4a/Hpgd/Lipa/Por/Abhd16a/Acsm3/Scd2/Srebf1/ Slc45a3/Cyp2c70/Cyp2c44/Acad9/Daglb/Cyp2a22/Decr2/Abcd2/Mettl20/Irs2/ Acsl4/Acadsb/Dgat2/Elovl5/Oxsm/Cyp2d40/Tysnd1/Cyp2c55/Fads1/Aacs*
GO:0005996	monosaccharide metabolic process	33/972	244/23577	2.07e-09	1.11e-06	33	*H6pd/Gck/Gk/Acadm/Atf4/Car5a/Gnmt/Gpld1/Gsto1/Hnf4a/Khk/Man2a1/Myc/ Pck1/Pcx/Enpp1/Rora/Sorbs1/Slc45a3/Tcf4/Ppp1r3b/Gulo/Galm/Irs2/Slc23a2/ Slc23a1/Slc35b4/Midn/Pgm2/Dgat2/Ugt2b1/Gale/Fggy*
GO:0002181	cytoplasmic translation	16/972	62/23577	2.83e-09	1.36e-06	16	*Rbm4/Rpl6/Rpl7a/Rpl7/Rpl9/Rps7/Eif5/Rpl27a/Rpl8/Rps3/Eif3d/Rplp1/Rpl35/Rpl15/ Rplp2/Rps20/Eif4b*
GO:0016042	lipid catabolic process	33/972	269/23577	2.40e-08	8.90e-06	33	*Hsd3b7/Acat1/Acadm/Acads/Adora1/Abcd1/Apoa4/Apoc2/Cidea/Cyp26a1/ Cyp46a1/Gpld1/Hadh/Hexb/Ldlr/Lipa/Lipg/Bco2/Neu1/Abhd16a/Cps1/Daglb/ Cyp26b1/Pnpla7/Abcd2/Mettl20/Irs2/Fgf21/Smpdl3a/Plbd1/Crtc3/Tysnd1/Gpcpd1*
GO:0015711	organic anion transport	40/972	369/23577	2.79e-08	9.59e-06	40	*Adora1/Abcd1/Apoa4/Apoc2/Aqp8/Atp9a/Slc7a2/Cd36/Slc25a1/Fabp2/Gja1/Il1a/ Ldlr/Slc22a27/Myc/Pctp/Pltp/Slc10a1/Slc16a7/Slc23a2/Slc23a1/Slc2a2/Slc36a1/ Slc17a8/Slc16a11/Slc25a21/Slc38a7/Slc22a26/Slco2a1/Slco1b2/Irs2/Acsl4/*

Top 30 GO terms enriched by differentially expressed genes. Terms that had overlapping enriched gene sets were summarised.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
GO:0006790	sulfur compound metabolic process	33/972	272/23577	3.15e-08	1.01e-05	33	*Slc35b4/Prelid1/Slc38a2/Gipc1/Slc25a22/Slc16a13/Tmem30a/Slc35b2 Gpat4/Blmh/Cth/Acat1/Cs/Gamt/Gnmt/Gpc1/Gstp2/Gstp1/Gstt1/Gstt2/Gsto1/ Hmgcl/Hnf4a/Mvk/Enpp1/Mvd/Acsm3/Tcf4/Cps1/Mpst/Hs3st3b1/Bhmt2/ Ethe1/Acadsb/Dgat2/Sult1c2/Oxsm/Sulf2/Arsg/Gstk1/Glce*
GO:0008202	steroid metabolic process	33/972	273/23577	3.44e-08	1.04e-05	33	*Hsd3b7/Cyb5r3/Akr1c20/Apoa4/Cyp17a1/Cyp26a1/Cyp2b10/Cyp46a1/Cyp7b1/ Dhcr7/Egr1/Fdft1/Fgf1/Fgfr1/Il1a/Ldlr/Lipa/Mvk/Pctp/Por/Med1/Mvd/Rora/ Srebf1/Srebf2/Hmgcs1/Vldlr/Sf1/Cyp26b1/0610007P14Rik/Dgat2/Ebpl/Sdr42e1*
GO:0044282	small molecule catabolic process	31/972	258/23577	1.02e-07	2.46e-05	31	*Hsd3b7/Blmh/Acat1/Acadm/Acads/Abcd1/Cyp26a1/Cyp46a1/Gk/Got1/Hadh/ Hyal1/Khk/Abhd16a/Shmt1/Ido2/Gls2/Ppat/Cyp26b1/Aadat/Abcd2/Aass/Galm/ Mettl20/Irs2/Nagk/Qprt/Afmid/Tysnd1/Pgm2/Gale*
GO:1903829	positive regulation of cellular protein localization	38/972	381/23577	5.23e-07	9.69e-05	38	*Bmp4/Bmpr1a/Cct2/Cct5/Plk3/Csnk2a2/Egfr/Hnmt/Ube2j2/Gas6/Gnb2l1/Hcls1/ Sqstm1/Med1/Rbpms/Rpl28/Sorbs1/Srebf2/Stx4a/Dtx3l/Tcf4/Gls2/Kat7/Tgfbr1/ Zpr1/Ddx58/Leprot/Mrgpre/Epb41/Gnl3/Ablim3/Plagl2/Pdcd10/Rtn4/1810011O10Rik/ Tmem30a/Ipo5/Sulf2*
GO:0044262	cellular carbohydrate metabolic process	29/972	251/23577	5.99e-07	0.11e-03	29	*Gck/Gk/Phka2/Acadm/Gnmt/Got1/Gpld1/Hnf4a/Khk/Pck1/Enpp1/Ptk2b/Rora/ Sorbs1/Stat3/Slc45a3/Tcf4/Ppp1r3b/Irs2/Ppp1r3c/Slc35b4/Midn/Dgat2/Mogat1/ Pgm2/Stk40/Gbe1/Fggy*
GO:0006766	vitamin metabolic process	13/972	62/23577	1.17e-06	0.16e-03	13	*Cbr1/Cyp26a1/Fgfr1/Gsto1/Pltp/Rbp1/Shmt1/Vnn1/Cyp26b1/Vnn3/Gulo/Slc23a2/ Slc23a1/Mmab*
GO:0034660	ncRNA metabolic process	40/972	427/23577	1.32e-06	0.16e-03	40	*Noc4l/Rrn3/Eprs/Nop2/Tars/Ang/Ell/Fbl/Rpsa/Ncl/Rpl10a/Rpl7a/Rpl7/Rps14/Rps16/ Rps17/Rps24/Rps6/Rps7/Rps8/Smarca4/Spin1/Utp18/Rpp38/Snapc4/Wdr3/Rps28/ Nop58/Ftsj3/Ddx21/Rpl35/Rnf113a2/Gtf2h5/Rps21/Nop56/Yars2/Nars/Rpf1/Wdr43*
GO:0010876	lipid localization	34/972	334/23577	1.30e-06	0.16e-03	34	*Abcd1/Apoa4/Apoc2/Aqp8/Atp9a/Anxa2/Cd36/Cidea/Fabp2/Hexb/Il1a/Ldlr/Lipg/ Lrp6/Mest/Pctp/Enpp1/Pltp/Ptch1/Slc10a1/Srebf2/Spns2/Vldlr/Slco2a1/Slco1b2/ Irs2/Acsl4/Apom/Prelid1/Dgat2/Tmem30a/Osbpl3/Hbp1/Spin1*
GO:0051262	protein tetramerization	21/972	159/23577	2.53e-06	0.25e-03	21	*Cth/Acadm/Acads/Asl/Anxa2/Cryz/Gnmt/Hist2h3c1/Hmgcl/S100a10/Shmt1/Ppat/ Cpsf7/Hist1h3h/Hist1h3i/Txnrd1/Samhd1/Dctpp1/Acot13/Crtc3/Hist2h3c2*

Top 30 GO terms enriched by differentially expressed genes. Terms that had overlapping enriched gene sets were summarised.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
GO:0006979	response to oxidative stress	36/972	376/23577	2.68-06	0.26e-03	36	*Foxp1/Ppp1r15b/Vimp/Apoa4/Aqp1/Atf4/Atox1/Atp2a2/Atrn/Bmp4/Casp6/Cd36/Plk3/Egfr/Stx2/Gnb2l1/Gstp1/Hyal1/Ppargc1b/Prkca/Ptk2b/Rbpms/Ripk1/Stx4a/Ucp2/Vnn1/Mapk3/Rps3/Msrb3/Txnrd1/Fut8/Rcan2/Pdcd10/1600014C10Rik/Neil1/Sesn3*
GO:0048732	gland development	39/972	429/23577	3.69e-06	0.35e-03	39	*Gpat4/Asl/Acadm/Bmp4/Bmpr1a/Btrc/Capn1/Cebpb/Cobl/Cyp7b1/Egfr/Fgfr1/Foxa3/Id2/Lmo4/Lrp6/Man2a1/Cited2/Neu1/Nme1/Notch1/Pdgfa/Med1/Prlr/Ptch1/Rps6ka1/Smarca4/Tcf4/Tgfbr1/Lims2/Mpst/Mapk3/Irs2/Irf6/Adrm1/Rtn4/Sulf2/Aacs/Gja1*
GO:0061008	hepaticobiliary system development	17/972	113/23577	3.81e-06	0.35e-03	17	*Asl/Acadm/Bmp4/Cebpb/Cobl/Foxa3/Man2a1/Cited2/Notch1/Med1/Rps6ka1/Smarca4/Lims2/Mpst/Nipbl/Sulf2/Aacs*
GO:0016051	carbohydrate biosynthetic process	22/972	177/23577	4.07e-06	0.37e-03	22	*Gck/Gk/Acadm/Atf4/Car5a/Gnmt/Got1/Gsto1/Hnf4a/Pck1/Pcx/Enpp1/Ptk2b/Sorbs1/Tcf4/Ppp1r3b/Gulo/Irs2/Ppp1r3c/Slc35b4/Dgat2/Pgm2/Gbe1*
GO:0097006	regulation of plasma lipoprotein particle levels	12/972	59/23577	4.21e-06	0.38e-03	12	*Apoa4/Apoc2/Anxa2/Cd36/Gpld1/Ldlr/Lipg/Pltp/Vldlr/Plagl2/Apom/Dgat2*
GO:0033157	regulation of intracellular protein transport	35/972	376/23577	6.87e-06	0.57e-03	35	*Atpif1/Bmp4/Bmpr1a/Cd36/Cdkn1a/Plk3/Csnk2a2/Ddx5/Egfr/Hnmt/Ube2j2/Gas6/Hcls1/Hnf4a/Med1/Ptpn1/Rbpms/Rpl28/Srebf1/Srebf2/Tcf4/Gls2/Tgfbr1/Zpr1/Ddx58/Leprot/Mrgpre/Ablim3/Plagl2/Pdcd10/1810011O10Rik/Tmem30a/Ipo5/Sulf2/Svip*
GO:1901615	organic hydroxy compound metabolic process	38/972	426/23577	7.60e-06	0.48e-03	38	*Hsd3b7/Cyb5r3/Apoa4/Cyp26a1/Cyp46a1/Cyp7b1/Ddc/Dhcr7/Fdft1/Fgf1/Fgfr1/Gk/Got1/Ldlr/Lipa/Mvk/Pck1/Pctp/Pltp/Por/Med1/Mvd/Ptk2b/Rbp1/Srebf1/Srebf2/Hmgcs1/Vldlr/Cyp26b1/Spp2/Gde1/0610007P14Rik/Moxd1/Dgat2/Ebpl/Mogat1/Rapgef2*
GO:0006520	cellular amino acid metabolic process	27/972	256/23577	8.18e-06	0.63e-03	27	*Blmh/Eprs/Cth/Asl/Acat1/Tars/Ddc/Gamt/Gnmt/Got1/Hnf4a/P4ha2/Shmt1/Ido2/Gls2/Cps1/Ppat/Aadat/Mpst/Asns/Aass/Txnrd1/Nit2/Bhmt2/Yars2/Nars/Afmid*
GO:0010608	posttranscriptional regulation of gene expression	39/972	452/23577	1.26e-05	0.86e-03	39	*Eprs/Ppp1r15b/Ang/Apobec1/Bcl3/Bmp4/Noct/Ddx3x/Ddx5/Egfr/Enc1/Gnb2l1/Grb7/Matr3/Mvk/Ncl/Neu1/Prkca/Ptbp1/Ptk2b/Rara/Rbm4/Shmt1/Stat3/Eif5/Epb41l5/Gigyf2/Pabpc4/Mapk3/Ctif/Rps3/Eif3d/Rplp1/Rmnd1/Rbm4b/Epb41l5/Rpl38/Cnot8/Taco1/Pum1*

Top 30 GO terms enriched by differentially expressed genes. Terms that had overlapping enriched gene sets were summarised.

PATHWAY ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
GO:0034248	regulation of cellular amide metabolic process	37/972	420/23577	1.32e-05	0.89e-03	37	*Eprs/Ppp1r15b/Ang/Bcl3/Bmp4/Plk3/Ddx3x/Ddx5/Egfr/Enc1/Gnb2l1/Grb7/Cyr61/Mvk/Ncl/Neu1/Prkca/Ptbp1/Ptk2b/Rara/Rbm4/Shmt1/Stat3/Eif5/Gigyf2/Mapk3/Ctif/Rps3/Eif3d/Rplp1/Smpd3/Rmnd1/Rbm4b/Rpl38/Cnot8/Taco1/Pum1*
GO:0006805	xenobiotic metabolic process	10/972	46/23577	1.434e-05	0.94e-03	10	*Cryz/Cyp26a1/Cyp2b10/Cyp46a1/Gstt1/Gsto1/Hnf4a/Rora/Cyp26b1/Ugt2b1*
GO:0061614	pri-miRNA transcription from RNA polymerase II promoter	9/972	38/23577	1.84e-05	0.11e-02	9	*Bmp4/Bmpr1a/Ddx5/Smarca4/Stat3/Tgfbr1/Yy1/Ets1/Gnl3*
GO:0051186	cofactor metabolic process	32/972	348/23577	2.14e-05	0.13e-02	32	*H6pd/Gpat4/Gck/Gk/Acat1/Atpif1/Cpox/Cs/Gamt/Gnmt/Hmgcl/Hnf4a/Khk/Alad/Mvk/Myc/Pklr/Mvd/Acsm3/Shmt1/Stat3/Vnn1/Vnn3/Txnrd1/Bhmt2/Acadsb/Qprt/Dgat2/Oxsm/Ppat/Gale/Adck4*

Table A17: Significantly altered metabolites between HFD and Chow in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(18:1;9E-enoyl/14:0) aa Phosphatidylcholine	PC aa C32:1	14.30866265	3.49E-10	6.12E-08	1.635841011	0.11432522	Phosphatidylcholines
C(12:0/18:0) aa Phosphatidylcholine	PC aa C30:0	14.91015834	9.05E-10	7.92E-08	2.049536066	0.137459041	Phosphatidylcholines
C(18:1;9Z-enoyl/e18:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:2	-12.04744491	1.96E-09	9.02E-08	-1.03751766	0.086119311	Phosphatidylcholines
C(20:3;8Z,11Z,14Z-trienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:3	-12.21111465	2.06E-09	9.02E-08	-1.233483754	0.101013199	Phosphatidylcholines
C(18:1;9Z,12Z-dienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C34:2	-10.4876851	1.66E-08	4.89E-07	-1.370791411	0.13070486	Phosphatidylcholines
Acetyl-L-ornithine	Ac-Orn	-10.69090152	1.68E-08	4.89E-07	-1.643630232	0.153741032	Biogenic Amines
C(16:0/18:1;9E-enoyl) aa Phosphatidylcholine	PC aa C34:1	10.49049553	2.02E-08	5.05E-07	0.673766131	0.06422634	Phosphatidylcholines
C17:0 a Lysophosphatidylcholine	lysoPC a C17:0	-9.06977874	1.05E-07	2.20E-06	-1.101082181	0.121401217	Lyso-Phosphatidylcholines
C(16:0/18:2;2E,4E-dienoyl) aa L-Phosphatidylcholine	PC aa C34:2	-9.150299365	1.20E-07	2.20E-06	-0.466270614	0.05095687	Phosphatidylcholines
C(18:2;9Z,12Z-dienoyl/18:4;9E,11E,13E,15E-tetraenoyl) aa L-Phosphatidylcholine	PC aa C36:6	9.963411588	1.25E-07	2.20E-06	1.258088089	0.126270814	Phosphatidylcholines
Octadecadienylcarnitine	C18:2	-8.513168851	2.56E-07	4.08E-06	-2.498208596	0.293452255	Acylcarnitines
Tetradecadienylcarnitine	C14:2	-8.435779611	3.38E-07	4.93E-06	-1.982172982	0.234972116	Acylcarnitines
C(16:0/20:1;11Z-enoyl) aa Phosphatidylcholine	PC aa C36:1	11.07174403	4.17E-07	5.61E-06	1.211814171	0.109451065	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C38:2	PC ae C38:2	-8.9447655	6.37E-07	7.97E-06	-0.889849579	0.099482717	Phosphatidylcholines
C(18:2;9Z,12Z-dienoyl/16:1;7Z-enoyl) aa Phosphatidylcholine	PC aa C34:3	-8.02139501	1.35E-06	1.57E-05	-1.006705112	0.125502498	Phosphatidylcholines
Hexadecadienylcarnitine	C16:2	-7.318894289	2.57E-06	2.81E-05	-2.260155226	0.308811022	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C40:2	PC ae C40:2	-7.079451734	2.98E-06	3.07E-05	-0.666891038	0.094200944	Phosphatidylcholines
C(22:5;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C40:6	-7.276790717	3.57E-06	3.47E-05	-0.628409335	0.086358033	Phosphatidylcholines
C(16:1;9E-enoyl/16:1;9E-enoyl) aa Phosphatidylcholine	PC aa C32:2	7.42427024	5.59E-06	5.15E-05	1.325192627	0.178494665	Phosphatidylcholines
C18:0 Sphingomyelin	SM C18:0	7.640123298	9.13E-06	7.99E-05	0.945148447	0.123708533	Sphingomyelins
C(18:1;6Z,9Z,12Z-trienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C34:3	-6.303549989	1.09E-05	9.12E-05	-0.845655366	0.134155415	Phosphatidylcholines
C(11:0/25:0) aa Phosphatidylcholine	PC aa C36:0	-6.468471769	1.49E-05	0.000118797	-0.791770667	0.122404595	Phosphatidylcholines
Hydroxyproline	OH-Pro	-6.07394404	1.65E-05	0.0001254	-0.603386623	0.099340168	Biogenic Amines
C(20:1;13Z-enoyl/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C38:3	7.16932514	2.24E-05	0.000163602	1.078395705	0.150418022	Phosphatidylcholines
C(22:5;7Z,10Z,13Z,16Z,19Z-pentaenoyl/18:1;11Z-enoyl) aa L-Phosphatidylcholine	PC aa C40:6	6.737660503	3.01E-05	0.00020419	0.820985476	0.121850229	Phosphatidylcholines
C(18:0/14:0) aa Phosphatidylcholine	PC aa C32:0	-6.74539856	3.03E-05	0.00020419	-0.830776826	0.123162007	Phosphatidylcholines
C(20:4;5E,8E,11E,14E-tetraenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C38:4	-6.950486049	3.43E-05	0.000217298	-0.814096275	0.117127963	Phosphatidylcholines
C(20:1;9Z-enoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:1	-5.802389172	3.48E-05	0.000217298	-0.557957801	0.09616001	Phosphatidylcholines

Significantly altered metabolites between HFD and Chow in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Sphingomyelin with acyl residue sum C20:2	SM C20:2	-6.581077486	3.65E-05	0.000220103	-1.269922259	0.192965705	Sphingomyelins
Putrescine	Putrescine	5.96957651	6.18E-05	0.000360755	1.616966542	0.270867881	Biogenic Amines
Phosphatidylcholine with acyl-alkyl residue sum C40:4	PC ae C40:4	-5.972509986	7.60E-05	0.000428831	-0.582167426	0.0974745	Phosphatidylcholines
Aspartic acid	Asp	-5.722766937	8.46E-05	0.00046273	-0.697933973	0.121957434	Amino Acids
C18:1;9Z,12Z-dienoyl a Lysophosphatidylcholine	lysoPC a C18:2	-5.71578709	9.49E-05	0.000503457	-1.008726926	0.176480843	Lyso-Phosphatidylcholines
C18:1;11Z-enoyl Sphingomyelin	SM C18:1	5.200054855	9.84E-05	0.000506359	0.887607177	0.170691887	Sphingomyelins
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-pentaenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C38:6	-5.111585542	0.00010475	0.000523751	-0.424770206	0.083099501	Phosphatidylcholines
C14:0 a Lysophosphatidylcholine	lysoPC a C14:0	6.659219104	0.000108863	0.000529195	0.361476147	0.054282062	Lyso-Phosphatidylcholines
C(22:5;4Z,7Z,10Z,13Z,16Z-pentaenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C40:5	-5.826766562	0.000119343	0.000564461	-0.514331731	0.088270523	Phosphatidylcholines
C(12:0/26:0) aa Phosphatidylcholine	PC aa C38:0	-5.745133729	0.000165461	0.00076199	-0.530674524	0.092369394	Phosphatidylcholines
C24:1 Sphingomyelin	SM C24:1	-5.813020619	0.000177788	0.000797769	-0.825868849	0.142072238	Sphingomyelins
Phosphatidylcholine with acyl-alkyl residue sum C38:3	PC ae C38:3	-5.509829851	0.000188432	0.00082439	-0.542330086	0.098429552	Phosphatidylcholines
Decadienylcarnitine	C10:2	5.088121671	0.000218641	0.000933224	0.311707276	0.061261758	Acylcarnitines
Hydroxytetradecadienylcarnitine	C14:2-OH	-4.829884199	0.000241704	0.001007099	-0.820455033	0.169870539	Acylcarnitines
C(22:4;7Z,10Z,13Z,16Z,19Z-tetraenoyl/18:1;9Z-enoyl) aa Phosphatidylcholine	PC aa C40:5	5.465435511	0.000301708	0.00122788	0.843748765	0.154379054	Phosphatidylcholines
Serotonin	Serotonin	-4.580557281	0.000315976	0.001256724	-0.782373943	0.170803222	Biogenic Amines
C(16:0/e18:0) ae D-Phosphatidylcholine	PC ae C34:0	-4.704307546	0.000353601	0.001360752	-0.448168688	0.095267727	Phosphatidylcholines
Hydroxysphingomyelin with acyl residue sum C22:1	SM (OH) C22:1	-4.546397284	0.000357683	0.001360752	-0.674724004	0.1484085	Sphingomyelins
Hexenoylcarnitine	C6:1	4.493579457	0.000373443	0.001390479	0.598642787	0.13322181	Acylcarnitines
Hydroxysphingomyelin with acyl residue sum C16:1	SM (OH) C16:1	-4.375631183	0.000473072	0.001724741	-0.663386691	0.151609371	Sphingomyelins
C(20:1;11Z-enoyl/20:0) aa L-Phosphatidylcholine	PC aa C40:1	-4.406882795	0.000592115	0.002114697	-0.345164401	0.078323935	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C40:1	PC ae C40:1	-4.18512885	0.000846017	0.00296106	-0.41345793	0.098792163	Phosphatidylcholines
Hydroxysphingomyelin with acyl residue sum C24:1	SM (OH) C24:1	-4.182484677	0.000897959	0.003081231	-0.555571808	0.132832957	Sphingomyelins
C(22:0/e20:0) ae L-Phosphatidylcholine	PC ae C42:0	4.480543288	0.000969359	0.003262266	0.34256801	0.076456802	Phosphatidylcholines
C(18:0/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C36:2	4.652146115	0.001002208	0.003309177	0.383683466	0.082474509	Phosphatidylcholines
Glutamine	Gln	4.06982264	0.001039155	0.003367633	0.313354714	0.076994685	Amino Acids
C16:1;9Z-enoyl a Lysophosphatidylcholine	lysoPC a C16:1	4.126528723	0.001145627	0.003645176	0.898164399	0.217656161	Lyso-Phosphatidylcholines
C(21:1;12Z-enoyl/21:1;12Z-enoyl) aa Phosphatidylcholine	PC aa C42:2	-3.966069548	0.001183369	0.003698027	-0.575692991	0.145154538	Phosphatidylcholines

Significantly altered metabolites between HFD and Chow in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Butenylcarnitine	C4:1	-4.216019793	0.001227714	0.003769297	-0.402490487	0.095466935	Acylcarnitines
C(20:4;5E,8E,11E,14E-tetraenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:4	-4.087347232	0.001543422	0.004656878	-0.427507713	0.104592952	Phosphatidylcholines
C18:1;9E-enoyl a Lysophosphatidylcholine	lysoPC a C18:1	4.286672817	0.001859507	0.005515487	0.715671708	0.166952725	Lyso-Phosphatidylcholines
2-Methylbutyrylcarnitine	C5	3.911475813	0.002016114	0.005880331	0.760849608	0.194517273	Acylcarnitines
Glutaconylcarnitine	C5:1-DC	4.034792344	0.00210586	0.006041403	0.687227981	0.170325489	Acylcarnitines
C(18:0;12Z-enoyl/22:4;7Z,10Z,13Z,16Z-tetraenoyl) aa Phosphatidylcholine	PC aa C40:4	4.105081862	0.002243315	0.006331937	0.490070825	0.119381499	Phosphatidylcholines
Pimelylcarnitine	C7-DC	-3.631687795	0.002411377	0.006698268	-0.990235673	0.272665419	Acylcarnitines
C(18:1;9Z-enoyl/e14:0) ae L-Phosphatidylcholine	PC ae C32:1	3.54955285	0.002749436	0.007462906	0.448253682	0.126284549	Phosphatidylcholines
C(16:1;9Z-enoyl/e14:0) ae L-Phosphatidylcholine	PC ae C30:1	-3.708774049	0.002771936	0.007462906	-1.113488649	0.300230921	Phosphatidylcholines
Hydroxytetradecenoylcarnitine	C14:1-OH	-3.538281888	0.002906404	0.007635125	-0.766372362	0.21659449	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C42:2	PC ae C42:2	3.513573088	0.002923162	0.007635125	0.515241583	0.146643195	Phosphatidylcholines
Spermine	Spermine	-3.501661517	0.002973632	0.007652731	-0.491506822	0.140363887	Biogenic Amines
C(16:0/18:4;9Z,11E,13E,15Z-tetraenoyl) aa L-Phosphatidylcholine	PC aa C34:4	3.495071843	0.003103006	0.007869942	0.470858299	0.134720635	Phosphatidylcholines
Hydroxyhexadecadienylcarnitine	C16:2-OH	-3.476691199	0.003521703	0.008804258	-0.682542932	0.196319688	Acylcarnitines
lysoPC a C20:3	lysoPC a C20:3	3.893178042	0.003688499	0.00909137	0.70277155	0.180513591	Lyso-Phosphatidylcholines
Hydroxysphingomyelin with acyl residue sum C22:2	SM (OH) C22:2	-3.496881639	0.003917667	0.009522108	-0.474129905	0.135586489	Sphingomyelins
C(18:1;6Z-enoyl/24:0) aa L-Phosphatidylcholine	PC aa C42:1	-3.408185802	0.004161338	0.009975811	-0.584566365	0.171518338	Phosphatidylcholines
C(20:4;(5Z,8Z,11Z,14Z-tetraenoyl/e18:1;9Z-en) ae L-Phosphatidylcholine	PC ae C38:5	-3.831394156	0.004265876	0.01008822	-0.442305347	0.115442403	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C44:6	PC ae C44:6	-3.381274509	0.004376262	0.010211277	-0.308334837	0.09118894	Phosphatidylcholines
Dodecanoylcarnitine	C12	3.621698249	0.004591002	0.010571385	1.052186109	0.290522853	Acylcarnitines
Propionylcarnitine	C3	3.225697496	0.005393048	0.012256927	0.873466208	0.270783671	Acylcarnitines
C(16:0/26:0) aa L-Phosphatidylcholine	PC aa C42:0	-3.178008138	0.005842929	0.013109137	-0.403063134	0.126828855	Phosphatidylcholines
C(16:0/e14:0) ae Phosphatidylcholine (PC Ae C30:0)	PC ae C30:0	3.226628073	0.005952734	0.013186166	0.579631224	0.179639925	Phosphatidylcholines
Aldohexose	H1	3.179754358	0.006027961	0.013186166	0.595193673	0.187182281	Hexoses
Dodecenoylcarnitine	C12:1	-3.063020165	0.00838659	0.018119176	-0.495078739	0.161630911	Acylcarnitines
Glycine	Gly	-2.830439712	0.012068774	0.02575653	-0.392547353	0.138687763	Amino Acids
C(20:3;5Z,8Z,11Z-trienoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	PC aa C38:4	3.116888933	0.012258279	0.025845769	0.499241034	0.160172866	Phosphatidylcholines
Hydroxyhexadecenoylcarnitine	C16:1-OH	-2.800511087	0.012944217	0.026967119	-0.899745292	0.321278961	Acylcarnitines

Significantly altered metabolites between HFD and Chow in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Alanine	Ala	2.768647786	0.014875106	0.030625218	0.288156961	0.104078591	Amino Acids
Phosphatidylcholine with acyl-alkyl residue sum C30:2	PC ae C30:2	-2.77972129	0.018533605	0.037713731	-0.396587618	0.14267172	Phosphatidylcholines
C(10:0/16:0) aa Phosphatidylcholine	PC aa C26:0	-2.627196328	0.020667764	0.041573089	-0.622816473	0.237065067	Phosphatidylcholines
Hexanoylcarnitine (Fumarylcarnitine)	C6 (C4:1-DC)	2.574285087	0.021012376	0.041785975	0.531064267	0.206295825	Acylcarnitines
C(20:0/e18:0) ae L-Phosphatidylcholine	PC ae C38:0	2.538697342	0.021898434	0.043058718	0.267700123	0.105447829	Phosphatidylcholines
Myristoleylcarnitine (Tetradecenoylcarnitine)	C14:1	-2.487546338	0.024325198	0.047298996	-0.968173648	0.389208287	Acylcarnitines
C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	PC aa C38:5	2.567230784	0.025765113	0.049548294	0.362494069	0.141200421	Phosphatidylcholines

Table A18: Predictive potential of liver metabolites under Chow and HFD. Only metabolites are shown that were significant for at least one of plasma Cholesterol, Insulin, or TAG levels.

METABOLITE	P-VALUE CHOLESTEROL CHOW	ADJ. P-VALUE CHOLESTEROL CHOW	P-VALUE CHOLESTEROL HFD	ADJ. P-VALUE CHOLESTEROL HFD	P-VALUE INSULIN CHOW	ADJ. P-VALUE INSULIN CHOW	P-VALUE INSULIN HFD	ADJ. P-VALUE INSULIN HFD	P-VALUE TAG CHOW	ADJ. P-VALUE TAG CHOW	P-VALUE TAG HFD	ADJ. P-VALUE TAG HFD	METABOLON CLASS
Asn	0.911994114	1	0.101759262	1	0.411918124	1	0.009818572	1	0.480716541	1	0.35446621	1	Amino Acids
C10	0.854410205	1	0.400112292	1	0.238000155	1	0.006879049	1	0.254020658	1	0.319883145	1	Acylcarnitines
C12	0.751144335	1	0.340191824	1	0.255508452	1	0.003942599	0.66235656	0.347216597	1	0.579985278	1	Acylcarnitines
C14	0.48417276	1	0.204513776	1	0.384812051	1	0.010539734	1	0.362260907	1	0.694101599	1	Acylcarnitines
C14.2.OH	0.337292874	1	0.514904041	1	0.563022103	1	0.003272801	0.556376116	0.318526037	1	0.468515419	1	Acylcarnitines
Ile	0.591156898	1	0.11433655	1	0.689260705	1	0.011301566	1	0.348959479	1	0.149610849	1	Amino Acids
PC.aa.C30.0	0.451584975	1	0.318219775	1	0.903084779	1	0.007618332	1	0.585584451	1	0.919721661	1	Phosphatidylcholines
PC.aa.C30.2	0.518588186	1	0.646856413	1	0.550718289	1	0.005424338	0.90586452	0.850726102	1	0.684805555	1	Phosphatidylcholines
PC.aa.C42.2	0.557276182	1	0.699838857	1	0.450782606	1	0.011439119	1	0.716218534	1	0.755201656	1	Phosphatidylcholines
PC.ae.C30.0	0.654817899	1	0.690010183	1	0.621912107	1	0.003672629	0.62067435	0.71488709	1	0.191483907	1	Phosphatidylcholines
PC.ae.C32.1	0.631270347	1	0.986749901	1	0.526978964	1	0.001456876	0.252039497	0.754639991	1	0.982411266	1	Phosphatidylcholines
PC.ae.C36.0	0.901923703	1	0.364969504	1	0.458248493	1	0.010656327	1	0.795761033	1	0.968450207	1	Phosphatidylcholines
Phe	0.548217508	1	0.162330625	1	0.827752553	1	0.003153427	0.539236098	0.262329618	1	0.19363353	1	Amino Acids
Pro	0.700773988	1	0.103065193	1	0.546724499	1	0.002894156	0.497794751	0.368814639	1	0.472794182	1	Amino Acids
Thr	0.99812224	1	0.177779704	1	0.458711955	1	0.001293661	0.225097085	0.558354374	1	0.543179564	1	Amino Acids
Trp	0.4612156	1	0.203961116	1	0.882486841	1	0.001130763	0.197883548	0.384615888	1	0.321460962	1	Amino Acids
lysoPC.a.C17.0	0.453764884	1	0.696689284	1	0.183931667	1	0.006732498	1	0.730805384	1	0.49244394	1	Lyso-Phosphatidylcholines
C3.OH	0.004939188	0.864357963	0.334941562	1	0.767368896	1	0.5416634	1	0.347651109	1	0.388735991	1	Acylcarnitines
PC.ae.C44.5	0.895475752	1	0.699877356	1	0.406369788	1	0.998818923	1	0.979497158	1	0.007997177	1	Phosphatidylcholines
SM.C20.2	0.488898187	1	0.754897739	1	0.603701175	1	0.797515251	1	0.901199188	1	0.006287184	1	Sphingomyelins
SM.C24.1	0.303389332	1	0.933367869	1	0.407842386	1	0.436079254	1	0.355682348	1	0.002724067	0.476711799	Sphingomyelins

Table A19: Comparison of metabolites present in the communities between Chow and HFD networks.

COMMUNITY CHOW	COMMUNITY HFD	TOTAL	ELEMENTS
1	1	6	PC.ae.C36.4, PC.ae.C40.5, PC.aa.C34.1, PC.ae.C40.6, SM.C24.0, SM.C18.1
1	2	3	PC.ae.C34.2, Carnosine, SM.C20.2
1	3	5	SM.C18.0, SM..OH..C24.1, PC.ae.C36.3, PC.aa.C32.0, PC.ae.C38.2
1	6	6	PC.ae.C36.2, PC.ae.C34.0, PC.ae.C38.6, PC.ae.C36.5, PC.ae.C34.1, PC.aa.C34.2
2	1	6	PC.ae.C42.3, PC.aa.C38.0, PC.ae.C38.3, PC.ae.C42.2, PC.ae.C38.0, PC.ae.C38.5
2	2	10	C5.DC..C6.OH., PC.aa.C36.6, PC.aa.C34.3, PC.aa.C34.4, C5.M.DC, PC.aa.C32.1, PC.ae.C36.0, PC.aa.C36.5, PC.ae.C40.2, PC.aa.C32.2
2	3	3	lysoPC.a.C16.1, PC.aa.C32.3, lysoPC.a.C18.1
2	6	2	PC.ae.C34.3, PC.ae.C36.1
2	7	1	His
3	1	7	PC.ae.C44.5, PC.aa.C36.0, PC.ae.C44.4, PC.aa.C42.4, PC.aa.C42.5, PC.ae.C44.6, PC.aa.C40.2
3	2	1	PC.ae.C42.4
3	3	5	PC.aa.C30.0, t4.OH.Pro, lysoPC.a.C14.0, PC.ae.C38.1, PC.aa.C38.1
3	5	2	PC.aa.C42.0, PC.aa.C42.1
3	6	19	lysoPC.a.C26.0, PC.aa.C40.1, PC.aa.C42.2, PC.aa.C30.2, PC.ae.C32.1, PC.ae.C32.2, PC.aa.C26.0, PC.ae.C30.0, lysoPC.a.C28.1, PC.ae.C30.2, PC.aa.C28.1, PC.ae.C30.1, lysoPC.a.C24.0, PC.aa.C24.0, PC.ae.C42.5, lysoPC.a.C28.0, PC.ae.C44.3, SM.C16.1, lysoPC.a.C26.1
4	1	9	SM.C16.0, PC.aa.C36.1, PC.aa.C40.5, PC.aa.C38.5, PC.aa.C40.6, PC.aa.C38.3, PC.aa.C36.2, PC.aa.C36.4, PC.aa.C38.6
4	3	1	Taurine
4	4	1	C18.1
4	5	1	SM.C24.1
4	6	2	SM..OH..C16.1, SM..OH..C22.1
5	1	3	C10.2, PC.aa.C36.3, PC.ae.C42.0
5	3	2	Spermidine, H1
5	4	13	C14.2, C14.2.OH, C16.1.OH, C16, C14.1, C16.1, C12.1, C16.2.OH, C18.1.OH, C18.2, C14, C14.1.OH, C16.2
6	2	2	C0, Ac.Orn
6	3	12	lysoPC.a.C16.0, Lys, Asn, Gly, Met, Orn, lysoPC.a.C18.2, Trp, lysoPC.a.C20.4, Spermine, Thr, C3
6	4	1	lysoPC.a.C17.0
6	5	1	Asp

Table A20: KEGG pathways and GO terms enriched by genes present in edges of the Chow network.

ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
mmu04141	Protein processing in endoplasmic reticulum	7/59	163/8370	1.36E-04	0.018769087	7	*Sec63/Stt3a/Hsp90b1/Ssr3/Derl1/Sec24d/Tram1*
mmu03060	Protein export	3/59	28/8370	9.62E-04	0.066351375	3	*Sec63/Srp54a/Sec11a*
GO:0048193	Golgi vesicle transport	12/158	257/23239	2.05E-07	4.91E-04	12	*Kif13a/Rab1a/Vamp3/Mppe1/Klhl20/Golga5/Tmed2/Sec24d/Ddhd2/Rab1b/Scfd1/Scyl1*
GO:0006613	cotranslational protein targeting to membrane	4/158	18/23239	5.84E-06	0.004675	4	*Sec63/Srp54a/Ssr3/Tram1*
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	4/158	17/23239	4.57E-06	0.004675	4	*Sec63/Srp54a/Ssr3/Tram1*
GO:0006888	ER to Golgi vesicle-mediated transport	7/158	124/23239	2.27E-05	0.013642464	7	*Rab1a/Mppe1/Tmed2/Sec24d/Ddhd2/Rab1b/Scfd1*
GO:0045047	protein targeting to ER	4/158	32/23239	6.38E-05	0.030607807	4	*Sec63/Srp54a/Ssr3/Tram1*
GO:0072599	establishment of protein localization to endoplasmic reticulum	4/158	36/23239	1.02E-04	0.04090721	4	*Sec63/Srp54a/Ssr3/Tram1*

Table A21: KEGG pathways and GO terms pathways enriched by genes present in edges of the HFD network.

ID	DESCRIPTION	GENERATIO	BGRATIO	P-VALUE	ADJUSTED P-VALUE	COUNT	GENE ID
mmu00561	Glycerolipid metabolism	6/143	61/8370	0.000574079	0.09826275	6	*Gpat4/Dgkz/Gk/Agpat1/Lpin2/Agpat2*
mmu05010	Alzheimer disease	10/143	175/8370	0.000802145	0.09826275	10	*Atp2a2/Atp5b/Bid/Cox5a/Ndufa2/Mapk3/Bace2/Ndufb5/Ndufb9/Ern1*
mmu00190	Oxidative phosphorylation	8/143	134/8370	0.002034651	0.166163147	8	*Atp5b/Cox5a/ND3/Ndufa2/Ndufb5/Ndufb9/Atp6v1g2/Ppa1*
mmu04010	MAPK signaling pathway	12/143	294/8370	0.004409988	0.180074491	12	*Mapkapk3/Erbb3/Fgfr4/Jun/Jund/Mknk2/Myc/Pdgfa/Rac1/Rasa1/ Map3k2/Mapk3*
mmu04072	Phospholipase D signaling pathway	8/143	149/8370	0.003925327	0.180074491	8	*Dgkz/Adcy6/Pdgfa/Pip5k1c/Mapk3/Agpat1/Rapgef4/Agpat2*
mmu04932	Non-alcoholic fatty liver disease (NAFLD)	8/143	151/8370	0.004255706	0.180074491	8	*Bid/Cox5a/Jun/Ndufa2/Rac1/Ndufb5/Ndufb9/Ern1*
mmu04520	Adherens junction	5/143	72/8370	0.007590962	0.206642843	5	*Ctnna1/Pvrl2/Rac1/Tcf7/Mapk3*
mmu05231	Choline metabolism in cancer	6/143	99/8370	0.00683156	0.206642843	6	*Dgkz/Jun/Pdgfa/Pip5k1c/Rac1/Mapk3*
mmu05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	5/143	72/8370	0.007590962	0.206642843	5	*Atp2a2/Ctnna1/Dsc2/Itga3/Tcf7*
GO:0043543	protein acylation	13/304	226/23239	9.59E-06	0.032238887	13	*Trrap/Taf9/Nmt2/Abhd17a/Mapk3/Brd7/Park7/Naa40/Zdhhc4/Setd5/Iws1/ Naa60/Brd8*
GO:0045786	negative regulation of cell cycle	18/304	445/23239	2.60E-05	0.043727346	18	*Trrap/Gadd45gip1/Dgkz/Dlg1/Gata6/Hpgd/Recql5/Myc/Pmp22/Ppp2r3a /Zbtb17/Topbp1/Ern1/Brd7/Angel2/Cep192/Tom1l1*
GO:0006473	protein acetylation	10/304	183/23239	0.000156369	0.049557471	10	*Trrap/Taf9/Mapk3/Brd7/Park7/Naa40/Setd5/Iws1/Naa60/Brd8*
GO:0006475	internal protein amino acid acetylation	9/304	145/23239	0.000127746	0.049557471	9	*Trrap/Taf9/Mapk3/Brd7/Naa40/Setd5/Iws1/Naa60/Brd8*
GO:0009108	coenzyme biosynthetic process	10/304	178/23239	0.000124607	0.049557471	10	*Gk/Mocs2/Myc/Nfs1/Naprt/Zbtb20/Adck3/Coq10b/Mocos*
GO:0016573	histone acetylation	9/304	137/23239	8.27E-05	0.049557471	9	*Trrap/Taf9/Mapk3/Brd7/Naa40/Setd5/Iws1/Naa60/Brd8*
GO:0018393	internal peptidyl-lysine acetylation	9/304	142/23239	0.000108904	0.049557471	9	*Trrap/Taf9/Mapk3/Brd7/Naa40/Setd5/Iws1/Naa60/Brd8*
GO:0018394	peptidyl-lysine acetylation	9/304	153/23239	0.00019174	0.049557471	9	*Trrap/Taf9/Mapk3/Brd7/Naa40/Setd5/Iws1/Naa60/Brd8*
GO:0034976	response to endoplasmic reticulum stress	11/304	213/23239	0.000122138	0.049557471	11	*Hspa13/Bid/Jun/Pmp22/Ern1/Park7/Ube4b/Ufm1/Tmx1/Erp44*
GO:0045981	positive regulation of nucleotide metabolic process	5/304	41/23239	0.000188835	0.049557471	5	*Gadd45gip1/Gk/Myc/Zbtb20/Park7*
GO:0090407	organophosphate biosynthetic process	17/304	473/23239	0.000182041	0.049557471	17	*Gpat4/Gk/Dgkz/Adcy6/Atp5b/Mocs2/Myc/Nfs1/Pdgfa/Pip5k1c/Naprt/ Agpat1/Zbtb20/Agpat2/Mocos*
GO:1900544	positive regulation of purine nucleotide metabolic process	5/304	41/23239	0.000188835	0.049557471	5	*Gadd45gip1/Gk/Myc/Zbtb20/Park7*
GO:1903580	positive regulation of ATP metabolic process	5/304	35/23239	8.72E-05	0.049557471	5	*Gadd45gip1/Gk/Myc/Zbtb20/Park7*

Table A22: SNPs associated with LRGs of Chow graph.

	BMI	Waist-hip ratio	Waist-hip ratio adj BMI	Waist circumference	Hip circumference	Hip circumference adj BMI	Triglycerides	Cholesterol	HDL cholesterol	LDL cholesterol
Pcsk7	X						X	X	X	X
Eif6		X	X		X	X				
Fndc3b				X						

Table A23: SNPs associated with LRGs of HFD graph.

	BMI	Waist-hip ratio	Waist-hip ratio adj BMI	Waist circumference	Hip circumference	Triglycerides	Cholesterol	HDL cholesterol	LDL cholesterol	Fasting insulin	Type 2 diabetes	Type 2 diabetes adj BMI	eGFR-creat (serum creatinine)	Heart rate	Incremental insulin at 30 min OGTT	HOMA-B	HbA1c	Creatinine Systolic blood pressure	Diastolic blood pressure	Coronary artery disease	
Cobll1	X	X	X	X	X	X		X	X	X	X	X									
Appl2	X	X		X							X		X	X	X	X	X				
Topbp1	X			X							X							X	X	X	
Pbrm1	X	X	X	X				X													
Lgr4	X			X	X																
Tmed1							X		X												X
Rpl17							X	X													
Atp9b	X												X								
Rac1	X							X													
Tmem205							X		X												
Hddc3											X	X									
Inhbe						X		X													
Dgkz								X													
Cep192	X																				
Tbcc	X																				
Setd5	X																				
CoX5a	X																				

Table A24: Metabolites influenced by the 10 LRGs selected for the Chow and HFD networks.

SHORT NAME	MOLECULE	METABOLON CLASS	FOLD CHANGE	P-VALUE	ADJUSTED P-VALUE	LRG
PC.aa.C34.2	C(16:0/18:2;2E,4E-dienoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	-0.466270614	1.20E-07	2.20E-06	*Ddx3x*
PC.ae.C36.3	C(20:3;8Z,11Z,14Z-trienoyl/e16:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-1.233483754	2.06E-09	9.02E-08	*Ddx3x*
PC.ae.C36.2	C(18:1;9Z-enoyl/e18:1;1Z-en) ae L-Phosphatidylcholine	Phosphatidylcholines	-1.03751766	1.96E-09	9.02E-08	*Ddx3x*
PC.ae.C34.2	C(18:1;9Z,12Z-dienoyl/e16:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-1.370791411	1.66E-08	4.89E-07	*Ddx3x*
SM.C20.2	Sphingomyelin with acyl residue sum C20:2	Sphingomyelins	-1.269922259	3.65E-05	0.000220103	*Ddx3x*
PC.ae.C36.1	C(20:1;9Z-enoyl/e16:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-0.557957801	3.48E-05	0.000217298	*Ddx3x*
PC.aa.C38.0	C(12:0/26:0) aa Phosphatidylcholine	Phosphatidylcholines	-0.530674524	0.000165461	0.00076199	*Arhgap24, Gk, Cobll1*
PC.aa.C40.2	C(20:1;11E-enoyl/20:1;11E-enoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	-0.038473213	0.710597926	0.7272201	*Arhgap24*
PC.aa.C42.5	C(22:1;7Z,10Z,13Z,16Z,19Z-pentanoyl/20:0) aa L-Phosphatidylcholine	Phosphatidylcholines	0.012277716	0.890667707	0.890667707	*Arhgap24, Cd82, Gk*
SM.C24.0	C24:0 Sphingomyelin	Sphingomyelins	0.167049525	0.157145082	0.221777333	*Arhgap24*
PC.aa.C42.6	C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/20:0) aa Phosphatidylcholine	Phosphatidylcholines	0.137708651	0.046976764	0.083039734	*Arhgap24*
PC.ae.C42.0	C(22:0/e20:0) ae L-Phosphatidylcholine	Phosphatidylcholines	0.34256801	0.000969359	0.003262266	*Arhgap24, Gk, Cobll1*
PC.aa.C40.4	C(18:0;12Z-enoyl/22:4;7Z,10Z,13Z,16Z-tetraenoyl) aa Phosphatidylcholine	Phosphatidylcholines	0.490070825	0.002243315	0.006331937	*Arhgap24, Gk, Cobll1*
PC.aa.C36.2	C(18:0/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.383683466	0.001002208	0.003309177	*Arhgap24, Gk*
PC.ae.C42.2	Phosphatidylcholine with acyl-alkyl residue sum C42:2	Phosphatidylcholines	0.515241583	0.002923162	0.007635125	*Arhgap24*
PC.aa.C36.3	C(18:3;6Z,9Z,12Z-trienoyl/18:0) aa Phosphatidylcholine	Phosphatidylcholines	0.243808025	0.036635409	0.067486279	*Arhgap24*
PC.ae.C42.1	Phosphatidylcholine with acyl-alkyl residue sum C42:1	Phosphatidylcholines	-0.176489614	0.128192533	0.186947444	*Cd82, Appl2, Cobll1*
PC.ae.C40.3	Phosphatidylcholine with acyl-alkyl residue sum C40:3	Phosphatidylcholines	-0.173194963	0.094790077	0.149969482	*Cd82, Appl2*
PC.aa.C42.4	Phosphatidylcholine with diacyl residue sum C42:4	Phosphatidylcholines	0.089705388	0.358261969	0.417972297	*Cd82*
PC.ae.C40.5	C(22:5;4Z,7Z,10Z,13Z,16Z-pentaenoyl/e18:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-0.514331731	0.000119343	0.000564461	*Cd82, Appl2, Cobll1*
PC.ae.C40.4	Phosphatidylcholine with acyl-alkyl residue sum C40:4	Phosphatidylcholines	-0.582167426	7.60E-05	0.000428831	*Cd82, Appl2, Cobll1*
SM.C18.1	C18:1;11Z-enoyl Sphingomyelin	Sphingomyelins	0.887607177	9.84E-05	0.000506359	*Cd82, Appl2*
PC.aa.C40.5	C(22:4;7Z,10Z,13Z,16Z,19Z-tetraenoyl/18:1;9Z-enoyl) aa Phosphatidylcholine	Phosphatidylcholines	0.843748765	0.000301708	0.00122788	*Cd82, Appl2, Cobll1*
PC.aa.C38.5	C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.362494069	0.025765113	0.049548294	*Cd82, Appl2, Cobll1*
PC.aa.C38.1	C(18:1;11E-enoyl/20:0) aa Phosphatidylcholine	Phosphatidylcholines	-0.267706204	0.086789162	0.140630587	*Rapgef4, Smim13*
t4.OH.Pro	Hydroxyproline	Biogenic Amines	-0.603386623	1.65E-05	0.0001254	*Rapgef4*
lysoPC.a.C16.0	C16:0 a Lysophosphatidylcholine	Lyso-Phosphatidylcholines	-0.228280356	0.081355221	0.133057605	*Rapgef4*
Val	Valine	Amino Acids	-0.367560007	0.097661643	0.151245907	*Rapgef4, Smim13*

Metabolites influenced by the 10 LRGs selected for the Chow and HFD networks.

SHORT NAME	MOLECULE	METABOLON CLASS	FOLD CHANGE	P-VALUE	ADJUSTED P-VALUE	LRG
Phe	Phenylalanine	Amino Acids	-0.208019722	0.254189294	0.317736618	*Rapgef4*
Trp	Tryptophan	Amino Acids	-0.162756332	0.45204104	0.503867402	*Rapgef4*
Thr	Threonine	Amino Acids	-0.222649868	0.118520475	0.174294816	*Rapgef4, Smim13*
Pro	Proline	Amino Acids	-0.245618252	0.276543266	0.343227458	*Rapgef4, Smim13*
Asn	Asparagine	Amino Acids	-0.132257174	0.36666474	0.42494258	*Rapgef4, Smim13*
C3	Propionylcarnitine	Acylcarnitines	0.873466208	0.005393048	0.012256927	*Rapgef4*
H1	Aldohexose	Hexoses	0.595193673	0.006027961	0.013186166	*Rapgef4*
alpha.AAA	2-Aminoadipic acid	Biogenic Amines	-0.396150042	0.224209802	0.28850526	*Smim13*
C16.1.OH	Hydroxyhexadecenoylcarnitine	Acylcarnitines	-0.899745292	0.012944217	0.026967119	*Inhbe*
C16.1	Palmitoleylcarnitine (Hexadecenoylcarnitine)	Acylcarnitines	-0.48754569	0.245294856	0.311062318	*Inhbe*
C16.OH	Hydroxyhexadecanoylcarnitine	Acylcarnitines	-0.35882423	0.355525595	0.417563618	*Inhbe*
C14.2	Tetradecadienylcarnitine	Acylcarnitines	-1.982172982	3.38E-07	4.93E-06	*Inhbe*
C14	Myristylcarnitine (Tetradecanoylcarnitine)	Acylcarnitines	0.638058621	0.09320799	0.149645855	*Inhbe*
C12	Dodecanoylcarnitine	Acylcarnitines	1.052186109	0.004591002	0.010571385	*Inhbe*
C10	Decanoylcarnitine	Acylcarnitines	0.121683069	0.402401585	0.457846443	*Inhbe*
PC.aa.C36.4	C(16:0/20:4;Z,8E,11E,14E,17E-tetraenoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	-0.14112726	0.158462029	0.22184684	*Myc, Cobll1*
PC.ae.C38.5	C(20:4;(5Z,8Z,11Z,14Z-tetraenoyl/e18:1;9Z-en) ae L-Phosphatidylcholine	Phosphatidylcholines	-0.442305347	0.004265876	0.01008822	*Myc, Gk, Cobll1*
PC.aa.C38.6	C(16:0/22:6;4E,7E,10E,13E,16E,19E-hexaenoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.133453182	0.133347382	0.192857783	*Myc, Gk, Cobll1*
C10.2	Decadienylcarnitine	Acylcarnitines	0.311707276	0.000218641	0.000933224	*Myc*
PC.aa.C40.6	C(22:5;7Z,10Z,13Z,16Z,19Z-pentaenoyl/18:1;11Z-enoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.820985476	3.01E-05	0.00020419	*Myc*
PC.aa.C38.4	C(20:3;5Z,8Z,11Z-trienoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.499241034	0.012258279	0.025845769	*Myc*
PC.aa.C38.3	C(20:1;13Z-enoyl/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	1.078395705	2.24E-05	0.000163602	*Gk, Cobll1*
PC.aa.C36.1	C(16:0/20:1;11Z-enoyl) aa Phosphatidylcholine	Phosphatidylcholines	1.211814171	4.17E-07	5.61E-06	Gk
PC.ae.C36.4	C(20:4;5E,8E,11E,14E-tetraenoyl/e16:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-0.427507713	0.001543422	0.004656878	*Gk, Cobll1*
PC.ae.C38.4	C(20:4;5E,8E,11E,14E-tetraenoyl/e18:0) ae L-Phosphatidylcholine	Phosphatidylcholines	-0.814096275	3.43E-05	0.000217298	*Cobll1*
SM..OH..C22.2	Hydroxysphingomyelin with acyl residue sum C22:2	Sphingomyelins	-0.474129905	0.003917667	0.009522108	*Appl2*
PC.ae.C38.3	Phosphatidylcholine with acyl-alkyl residue sum C38:3	Phosphatidylcholines	-0.542330086	0.000188432	0.00082439	*Appl2*

Table A25: Pairwise metabolite correlations of control and siRNA knockdown samples for candidate genes. The last column shows the p-values obtained with the Steiger test, testing the difference between the two correlation coefficients.

METABOLITE 1	METABOLITE 2	GENE	CORRELATION CONTROL	CORRELATION KD	CONTROL VS. KD
PC.aa.C40.5	PC.ae.C40.5	*Appl2*	0.573	0.977	0.0528
PC.aa.C38.5	PC.ae.C40.5	*Appl2*	0.763	0.970	0.1797
PC.ae.C38.3	PC.ae.C38.4	*Appl2*	0.665	0.970	0.1147
PC.ae.C38.3	PC.ae.C40.5	*Appl2*	0.327	0.953	0.0614
PC.ae.C38.4	PC.ae.C40.5	*Appl2*	0.721	0.942	0.3008
PC.ae.C38.3	SM.C18.1	*Appl2*	0.780	0.780	0.9998
PC.ae.C38.4	PC.ae.C40.3	*Appl2*	0.564	0.694	0.7912
PC.aa.C38.5	PC.aa.C38.6	*Cobbl1*	0.995	0.999	0.2545
PC.aa.C40.4	PC.ae.C38.4	*Cobbl1*	0.959	0.995	0.2565
PC.aa.C38.6	PC.ae.C40.5	*Cobbl1*	0.793	0.994	0.0444
PC.aa.C36.4	PC.ae.C38.5	*Cobbl1*	0.864	0.993	0.0931
PC.ae.C38.4	PC.ae.C38.5	*Cobbl1*	0.796	0.993	0.0589
PC.aa.C38.3	PC.ae.C38.4	*Cobbl1*	0.542	0.983	0.0541
PC.aa.C38.5	PC.ae.C38.4	*Cobbl1*	0.936	0.980	0.5168
PC.aa.C40.5	PC.ae.C38.4	*Cobbl1*	0.906	0.978	0.4121
PC.aa.C38.0	PC.ae.C40.5	*Cobbl1*	0.576	0.978	0.0814
PC.aa.C38.6	PC.ae.C38.4	*Cobbl1*	0.951	0.976	0.6809
PC.aa.C36.4	PC.aa.C38.5	*Cobbl1*	0.976	0.964	0.8124
PC.aa.C36.4	PC.aa.C38.6	*Cobbl1*	0.993	0.958	0.3408
PC.aa.C38.5	PC.ae.C38.5	*Cobbl1*	0.933	0.951	0.8628
PC.aa.C38.0	PC.ae.C38.4	*Cobbl1*	0.689	0.934	0.3544
PC.aa.C38.5	PC.ae.C36.4	*Cobbl1*	0.963	0.918	0.6496
PC.aa.C36.4	PC.aa.C38.0	*Cobbl1*	0.836	0.915	0.7033
PC.aa.C38.6	PC.aa.C42.5	*Gk*	0.657	0.852	0.5585
PC.aa.C36.2	PC.aa.C38.6	*Gk*	0.164	-0.835	0.0931
PC.aa.C38.3	PC.aa.C38.6	*Gk*	0.362	-0.810	0.0649
PC.aa.C38.6	PC.aa.C40.4	*Gk*	0.942	0.735	0.3186
PC.aa.C38.0	PC.aa.C42.5	*Gk*	0.343	0.657	0.5982
PC.aa.C38.3	PC.ae.C38.5	*Gk*	-0.027	-0.589	0.4272
PC.aa.C40.4	PC.ae.C38.5	*Gk*	0.715	0.540	0.7205
PC.aa.C36.1	PC.aa.C38.0	*Gk*	0.037	-0.442	0.5303
PC.aa.C36.1	PC.aa.C38.6	*Gk*	0.234	-0.297	0.5044
PC.aa.C38.3	PC.ae.C36.4	*Gk*	0.405	-0.264	0.3914
PC.aa.C36.1	PC.ae.C38.5	*Gk*	-0.097	-0.178	0.9198
Phe	H1	*Rapgef4*	0.770	0.846	0.7852
Pro	lysoPC.a.C16.0	*Rapgef4*	-0.760	-0.813	0.8640
Trp	lysoPC.a.C16.0	*Rapgef4*	-0.028	-0.706	0.2973
Phe	t4.OH.Pro	*Rapgef4*	0.733	0.700	0.9352
t4.OH.Pro	lysoPC.a.C16.0	*Rapgef4*	0.656	0.692	0.9340
Pro	PC.aa.C38.1	*Rapgef4*	-0.799	-0.581	0.5959
Pro	t4.OH.Pro	*Rapgef4*	-0.453	-0.513	0.9228
Phe	lysoPC.a.C16.0	*Rapgef4*	0.388	0.513	0.8475
Phe	PC.aa.C38.1	*Rapgef4*	0.238	0.508	0.6976
lysoPC.a.C16.0	PC.aa.C38.1	*Rapgef4*	0.981	0.317	0.0148
t4.OH.Pro	PC.aa.C38.1	*Rapgef4*	0.506	0.292	0.7529
Trp	PC.aa.C38.1	*Rapgef4*	-0.180	0.239	0.6016
Val	lysoPC.a.C16.0	*Rapgef4*	0.012	-0.053	0.9356

Table A26: Associations between hepatic gene expression levels and metabolic traits: BMI (n=170), TAG (n=170), and HOMA-IR (n=77) from human volunteers.

HUMAN GENE	BMI P-VALUE	BMI R^2	TAG P-VALUE	TAG R^2	HOMA-IR P-VALUE	HOMA-IR R^2
RAPGEF4	0.7570	0.0006	0.7010	0.0009	0.2330	0.0189
ARHGAP24	0.0934	0.0167	0.0889	0.0171	0.8610	0.0004
GK	0.0034	0.0500	0.0887	0.0171	0.1010	0.0355
COBLL1	0.7430	0.0006	0.1570	0.0119	0.7300	0.0016
INHBE	0.0001	0.0842	< 0,0001	0.0990	0.0078	0.0906
CD82	0.8000	0.0004	0.1030	-0.0157	0.6410	0.0029
TAP1	0.0078	0.0413	0.1023	0.0158	0.1440	0.0282
SMIM13	0.0810	0.0180	0.0169	0.0335	0.1820	0.0237
DDX3X	0.0696	0.0009	0.9670	< 0,0010	0.6420	0.0029
MYC	0.0092	0.0397	0.7100	0.0008	0.2290	0.0190
APPL2	0.3890	0.0044	0.6100	0.0015	0.9700	< 0,0010

Table A27: Phenotype monitoring of Ex4 treated, CR and H>C mice, giving the mean and standard deviation for each measured feature.

	H>C	CR	EX4	P-VALUE
Body weight at the end of the study (g)	47.26 ± 5.05	37.16 ± 6.40	39.37 ± 2.60	1.5E-03
Liver triglyceride levels (µg/mg)	7.73 ± 3.03	9.15 ± 3.67	17.76 ± 6.46	1.4E-03
Plasma triglyceride levels (mg/dl)	62.38 ± 12.85	43.02 ± 14.24	47.12 ± 10.80	1.8E-02
Plasma cholesterol levels (mg/dl)	67.15 ± 8.00	54.65 ± 13.62	60.51 ± 25.02	1.3E-01
Plasma insulin levels (mg/dl)	5.23 ± 3.74	0.31 ± 0.19	0.30 ± 0.11	2E-04

Table A28: 50 genes with the smallest adjusted p-values for the comparison of liver expression profiles of exendin-4 treated mice compared to calorie restricted mice.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Mki67	4.822	5.85	23.49	3.45E-20	3.54E-16
Hist2h3b	3.473	9.08	20.67	1.09E-18	1.12E-14
Hist1h2ao	4.207	10.76	18.87	1.22E-17	1.25E-13
Hist1h2ap	4.207	10.76	18.87	1.22E-17	1.25E-13
D17H6S56E-5	6.189	7.57	18.59	1.81E-17	1.85E-13
Hist2h3c2	3.963	9.03	18.02	4.13E-17	4.23E-13
Hist2h3c1	3.963	9.03	18.02	4.13E-17	4.23E-13
Hist1h2af	4.108	8.05	16.49	4.17E-16	4.27E-12
Hist1h2ai	3.285	11.12	16.26	6.02E-16	6.17E-12
Gstm2	4.004	11.05	16.04	8.51E-16	8.72E-12
Hist1h2ag	3.741	10.70	14.98	4.89E-15	5.01E-11
Chpt1	-2.846	14.51	-14.83	6.31E-15	6.47E-11
Hist1h2ah	3.723	10.65	14.53	1.06E-14	1.08E-10
Ces3b	-5.415	12.38	-14.49	1.14E-14	1.17E-10
Cenpm	3.071	5.18	14.31	1.56E-14	1.60E-10
Hist1h2ab	3.018	7.93	14.25	1.72E-14	1.76E-10
Hist1h3h	2.968	10.34	14.24	1.75E-14	1.79E-10
Hist1h3i	2.985	10.90	13.80	3.87E-14	3.97E-10
Inmt	-5.833	12.46	-13.51	6.58E-14	6.75E-10
Hist1h1a	3.112	6.19	13.47	7.01E-14	7.19E-10
Hist1h2ae	3.174	7.00	13.19	1.18E-13	1.21E-09
Gstm3	5.013	11.59	13.00	1.68E-13	1.72E-09

Top 50 DE genes in the liver Ex4 vs. CR.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
BC021614	3.151	9.56	12.81	2.40E-13	2.47E-09
Hist1h2ad	3.295	10.20	12.63	3.42E-13	3.50E-09
Hist1h3g	2.930	7.76	12.61	3.54E-13	3.63E-09
Wfdc21	-4.440	14.19	-12.54	4.05E-13	4.16E-09
Pigr	-2.396	16.91	-12.07	1.01E-12	1.04E-08
Cyp2c54	-5.258	11.56	-12.06	1.04E-12	1.07E-08
Hist1h2ak	3.289	7.28	12.01	1.15E-12	1.18E-08
Hist1h2an	3.152	11.13	11.94	1.32E-12	1.35E-08
Mcm5	3.554	5.51	11.85	1.57E-12	1.61E-08
Car1	-3.350	6.83	-11.67	2.27E-12	2.32E-08
Serpina1d	-1.920	17.38	-11.65	2.39E-12	2.45E-08
Hist1h3e	2.486	11.12	11.52	3.11E-12	3.19E-08
Srxn1	3.577	10.96	11.32	4.63E-12	4.75E-08
Gm6682	1.859	12.87	11.18	6.23E-12	6.39E-08
Palmd	1.895	12.48	11.17	6.37E-12	6.54E-08
Scnn1a	-3.477	9.18	-11.10	7.36E-12	7.54E-08
Acadsb	-1.601	13.26	-11.08	7.75E-12	7.95E-08
Slc27a5	-2.700	14.58	-11.06	8.00E-12	8.20E-08
Serpina3m	-1.960	18.18	-10.97	9.69E-12	9.93E-08
Nox4	-2.428	9.09	-10.94	1.02E-11	1.05E-07
Mcm2	2.633	5.85	10.80	1.40E-11	1.43E-07
Rorc	-2.491	9.69	-10.77	1.50E-11	1.53E-07
H2afx	1.764	8.66	10.76	1.53E-11	1.57E-07
Csad	-2.972	13.99	-10.69	1.75E-11	1.79E-07
Tuba1b	1.871	11.10	10.68	1.81E-11	1.85E-07
Mcm6	3.640	5.90	10.63	1.99E-11	2.04E-07
Cyp2d9	-2.714	17.88	-10.61	2.11E-11	2.16E-07
Slc48a1	2.193	12.69	10.60	2.15E-11	2.20E-07

Table A29: 50 genes with the smallest adjusted p-values for the comparison of liver expression profiles of exendin-4 treated mice compared to diet switch H>C mice.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Hist2h3b	3.815	9.08	22.20	1.60E-19	1.05E-15
Mki67	4.617	5.85	22.00	2.04E-19	1.05E-15
Hist2h3c2	4.409	9.03	19.61	4.45E-18	1.14E-14
Hist2h3c1	4.409	9.03	19.61	4.45E-18	1.14E-14
Cyp4a31	5.271	10.15	18.90	1.17E-17	2.05E-14
D17H6S56E-5	6.428	7.57	18.89	1.20E-17	2.05E-14
Chpt1	-3.573	14.51	-18.21	3.15E-17	4.62E-14
Ces3b	-6.668	12.38	-17.45	9.62E-17	1.23E-13
App	3.499	12.99	16.93	2.12E-16	2.42E-13
Hist1h2ao	3.679	10.76	16.14	7.23E-16	6.74E-13
Hist1h2ap	3.679	10.76	16.14	7.23E-16	6.74E-13
Slc15a5	-4.164	6.92	-15.96	9.70E-16	8.29E-13
Acadsb	-2.291	13.26	-15.50	2.07E-15	1.63E-12
Hist1h1a	3.560	6.19	15.07	4.19E-15	3.07E-12
Hist1h2ai	3.020	11.12	14.62	9.08E-15	6.21E-12
Trhde	-2.837	4.97	-14.42	1.28E-14	8.17E-12
Haao	-3.341	13.99	-14.29	1.61E-14	8.74E-12
Hist1h2af	3.641	8.05	14.29	1.60E-14	8.74E-12
Acot29	4.943	11.78	14.29	1.62E-14	8.74E-12
Inmt	-6.273	12.46	-14.20	1.88E-14	9.63E-12
Wfdc21	-5.103	14.19	-14.09	2.28E-14	1.11E-11
Hist1h2ab	3.000	7.93	13.86	3.48E-14	1.62E-11
Cenpm	3.011	5.18	13.72	4.47E-14	1.99E-11
H2afx	2.295	8.66	13.69	4.70E-14	2.01E-11

Top 50 DE genes in the liver Ex4 vs. H>C.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Scnn1a	-4.348	9.18	-13.58	5.78E-14	2.37E-11
Car1	-3.948	6.83	-13.45	7.27E-14	2.87E-11
Srxn1	4.335	10.96	13.42	7.68E-14	2.92E-11
Hist1h3h	2.844	10.34	13.35	8.77E-14	3.21E-11
Ido2	-3.175	10.16	-13.26	1.03E-13	3.65E-11
Hist1h2ag	3.377	10.70	13.22	1.11E-13	3.79E-11
Hist1h2ah	3.437	10.65	13.12	1.34E-13	4.45E-11
Ghr9	-2.121	16.88	-13.07	1.49E-13	4.76E-11
Rplp1	1.503	16.92	13.03	1.59E-13	4.93E-11
Ces1e	-3.943	11.54	-12.95	1.84E-13	5.55E-11
Rps12	1.922	13.14	12.80	2.44E-13	7.14E-11
Myc	4.264	8.71	12.67	3.13E-13	8.92E-11
Slc27a59	-3.144	14.58	-12.60	3.61E-13	1.00E-10
Serpina1d	-2.122	17.38	-12.58	3.72E-13	1.00E-10
AI506816	4.909	7.29	12.46	4.73E-13	1.24E-10
Ttc39c	-3.409	13.90	-12.30	6.45E-13	1.61E-10
Nt5dc2	2.773	9.10	12.31	6.31E-13	1.61E-10
Hist1h3i	2.713	10.90	12.27	6.91E-13	1.69E-10
Serpinb6a	2.559	7.57	12.25	7.08E-13	1.69E-10
Hnrnpa1	1.698	10.08	12.21	7.66E-13	1.79E-10
Mcm2	3.029	5.85	12.15	8.71E-13	1.98E-10
Cyp2d9	-3.166	17.88	-12.10	9.58E-13	2.14E-10
Hist1h2ae	2.966	7.00	12.06	1.05E-12	2.29E-10
Serpina3m	-2.169	18.18	-11.88	1.50E-12	3.01E-10
Slc38a3	-1.486	17.33	-11.88	1.47E-12	3.01E-10
Hist1h2ak	3.329	7.28	11.89	1.47E-12	3.01E-10

Table A30: 50 genes with the smallest adjusted p-values for the comparison of liver expression profiles of calorie restricted mice compared to diet switch H>C mice.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Acly	-3.197	12.09	-13.03	1.59E-13	1.63E-09
Fasn	-4.433	11.06	-12.65	3.25E-13	1.66E-09
Dpy19l3	-3.006	7.46	-12.35	5.88E-13	2.01E-09
Anks4b	-2.180	8.41	-12.02	1.11E-12	2.81E-09
Cyp4a31	3.031	10.15	11.92	1.37E-12	2.81E-09
Dhcr7	-3.341	9.96	-11.62	2.54E-12	3.73E-09
Cyp39a1	3.834	13.70	11.67	2.29E-12	3.73E-09
Trhde	-2.050	4.97	-11.43	3.70E-12	4.74E-09
Acot2	3.574	11.78	11.33	4.57E-12	5.21E-09
Mt1	5.277	14.64	10.94	1.04E-11	1.06E-08
Elovl6	-3.233	12.15	-10.60	2.16E-11	2.01E-08
Mid1ip1	-2.644	9.67	-10.38	3.43E-11	2.70E-08
Pcsk4	-2.224	5.71	-10.41	3.24E-11	2.70E-08
Fdps	-3.186	8.76	-10.17	5.53E-11	3.15E-08
Acss2	-3.042	12.29	-10.22	4.95E-11	3.15E-08
Rorc	2.199	9.69	10.19	5.22E-11	3.15E-08
Cln6	2.347	8.49	10.26	4.52E-11	3.15E-08
Pdk4	3.388	8.38	10.17	5.51E-11	3.15E-08
Nsdhl	-4.233	12.04	-10.03	7.44E-11	4.01E-08
Aqp8	-4.455	10.88	-9.92	9.60E-11	4.38E-08
Insig1	-2.855	12.81	-9.94	9.19E-11	4.38E-08
Gpam	-2.061	14.55	-9.96	8.77E-11	4.38E-08
Zfp36l1	1.414	13.32	9.91	9.83E-11	4.38E-08
Thrsp	-5.906	12.30	-9.76	1.38E-10	5.67E-08
Igfbp1	5.862	10.38	9.77	1.34E-10	5.67E-08
Plbd2	1.379	11.25	9.62	1.90E-10	7.48E-08

Top 50 DE genes in the liver CR vs. H>C.

GENE	LOG2 FC	AVERAGE EXPRESSION	T	P-VALUE	ADJUSTED P-VALUE
Slc15a5	-2.274	6.92	-9.56	2.18E-10	8.09E-08
Mt2	6.084	13.36	9.55	2.21E-10	8.09E-08
Nt5dc2	1.954	9.10	9.52	2.40E-10	8.49E-08
Clpx	2.332	11.93	9.46	2.72E-10	9.30E-08
Cyp4a10	2.737	16.72	9.43	2.95E-10	9.75E-08
Mvd	-2.386	5.61	-9.37	3.35E-10	1.07E-07
Cyp2u1	-2.992	12.01	-9.36	3.48E-10	1.08E-07
Slc23a1	-2.171	11.94	-9.34	3.59E-10	1.08E-07
Nr1d1	-4.957	10.99	-9.32	3.82E-10	1.12E-07
Gstm7	-1.726	10.73	-9.26	4.37E-10	1.25E-07
Nfil3	2.024	7.93	9.19	5.12E-10	1.42E-07
Ldlr	-1.955	10.58	-9.15	5.68E-10	1.53E-07
Mmp14	1.426	8.44	8.99	8.28E-10	2.18E-07
Msmo1	-2.779	9.04	-8.94	9.15E-10	2.23E-07
Tbcel	-1.536	8.27	-8.95	9.08E-10	2.23E-07
Acot1	3.246	11.32	8.94	9.13E-10	2.23E-07
Pcsk9	-2.482	6.85	-8.86	1.12E-09	2.66E-07
Echdc1	-2.088	11.58	-8.82	1.22E-09	2.84E-07
Mmab	-1.387	8.02	-8.70	1.66E-09	3.77E-07
Tsc22d1	-3.127	13.40	-8.64	1.92E-09	4.20E-07
Cda	2.187	10.03	8.63	1.93E-09	4.20E-07
Slc38a3	-0.977	17.33	-8.56	2.27E-09	4.75E-07
Tspan4	1.688	11.25	8.57	2.26E-09	4.75E-07
Myc	2.614	8.71	8.52	2.53E-09	5.19E-07

Table A31: Significantly altered metabolites between Ex4 and CR in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(22:4;7Z,10Z,13Z,16Z,19Z-tetraenoyl/18:1;9Z-enoyl) aa Phosphatidylcholine	PC aa C40:5	17.21	3.03E-11	5.30E-09	1.58	0.09	Phosphatidylcholines
Spermine	Spermine	-12.95	2.98E-09	2.61E-07	-1.44	0.11	Biogenic Amines
C(18:0;12Z-enoyl/22:4;7Z,10Z,13Z,16Z-tetraenoyl) aa Phosphatidylcholine	PC aa C40:4	11.06	1.80E-08	1.05E-06	0.91	0.08	Phosphatidylcholines
C18:1;11Z-enoyl Sphingomyelin	SM C18:1	10.46	8.82E-08	3.86E-06	0.89	0.09	Sphingomyelins
C18:1;9Z,12Z-dienoyl a Lysophosphatidylcholine	lysoPC a C18:2	-9.19	2.12E-07	7.43E-06	-1.46	0.16	Lyso-Phosphatidylcholines
C(22:5;7Z,10Z,13Z,16Z,19Z-pentaenoyl/18:1;11Z-enoyl) aa L-Phosphatidylcholine	PC aa C40:6	8.72	3.46E-07	1.01E-05	0.75	0.09	Phosphatidylcholines
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/20:0) aa Phosphatidylcholine	PC aa C42:6	8.53	5.78E-07	1.35E-05	0.59	0.07	Phosphatidylcholines
C(18:1;9Z-enoyl/e14:0) ae L-Phosphatidylcholine	PC ae C32:1	8.25	6.19E-07	1.35E-05	0.84	0.10	Phosphatidylcholines
C(20:4;5E,8E,11E,14E-tetraenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:4	7.92	1.03E-06	2.00E-05	0.86	0.11	Phosphatidylcholines
C(22:1;7Z,10Z,13Z,16Z,19Z-pentanoyl/20:0) aa L-Phosphatidylcholine	PC aa C42:5	8.12	1.42E-06	2.48E-05	0.77	0.10	Phosphatidylcholines
C16:0 Sphingomyelin	SM C16:0	7.59	1.65E-06	2.60E-05	0.77	0.10	Sphingomyelins
C(12:0/18:0) aa Phosphatidylcholine	PC aa C30:0	7.53	2.00E-06	2.60E-05	0.70	0.09	Phosphatidylcholines
C16:0 a Lysophosphatidylcholine	lysoPC a C16:0	-7.50	2.08E-06	2.60E-05	-0.69	0.09	Lyso-Phosphatidylcholines
C(20:3;5Z,8Z,11Z-trienoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	PC aa C38:4	7.55	1.78E-06	2.60E-05	0.95	0.13	Phosphatidylcholines
(5Z,8Z,11Z,14Z) a C20:4 Lysophosphatidylcholine	lysoPC a C20:4	-7.32	2.84E-06	3.31E-05	-1.07	0.15	Lyso-Phosphatidylcholines
C(16:0;e18:1;9Z-en) ae Phosphatidylcholine	PC ae C34:1	7.41	3.12E-06	3.41E-05	0.77	0.10	Phosphatidylcholines
C17:0 a Lysophosphatidylcholine	lysoPC a C17:0	-7.32	3.78E-06	3.89E-05	-0.75	0.10	Lyso-Phosphatidylcholines
Glycine	Gly	-6.94	4.85E-06	4.72E-05	-0.67	0.10	Amino Acids
Pimelylcarnitine	C7-DC	-6.85	8.48E-06	7.81E-05	-0.77	0.11	Acylcarnitines
C(20:4;5E,8E,11E,14E-tetraenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C38:4	7.07	1.01E-05	8.86E-05	0.51	0.07	Phosphatidylcholines
C(20:1;13Z-enoyl/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C38:3	6.28	1.49E-05	1.24E-04	1.17	0.19	Phosphatidylcholines
Valine	Val	-6.41	1.74E-05	1.38E-04	-0.95	0.15	Amino Acids
C(20:4;(5Z,8Z,11Z,14Z-tetraenoyl/e18:1;9Z-en) ae L-Phosphatidylcholine	PC ae C38:5	6.13	1.91E-05	1.46E-04	0.90	0.15	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C42:3	PC ae C42:3	-6.95	2.08E-05	1.52E-04	-1.19	0.17	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C40:1	PC ae C40:1	-6.90	3.54E-05	2.48E-04	-1.04	0.15	Phosphatidylcholines
Sphingomyelin with acyl residue sum C16:1	SM C16:1	5.73	4.00E-05	2.69E-04	0.36	0.06	Sphingomyelins
C18:1;9E-enoyl a Lysophosphatidylcholine	lysoPC a C18:1	-6.19	4.38E-05	2.84E-04	-0.97	0.16	Lyso-Phosphatidylcholines
Acetyl-L-ornithine	Ac-Orn	-5.53	5.91E-05	3.69E-04	-0.62	0.11	Biogenic Amines

Significantly altered metabolites between Ex4 vs. CR in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(22:0/e20:0) ae L-Phosphatidylcholine	PC ae C42:0	5.41	7.59E-05	4.43E-04	0.26	0.05	Phosphatidylcholines
C24:1 Sphingomyelin	SM C24:1	5.97	7.34E-05	4.43E-04	0.94	0.16	Sphingomyelins
C18:0 Sphingomyelin	SM C18:0	6.15	9.39E-05	5.14E-04	0.98	0.16	Sphingomyelins
Putrescine	Putrescine	5.49	9.16E-05	5.14E-04	1.97	0.36	Biogenic Amines
C(18:0/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C36:2	5.28	1.32E-04	7.02E-04	0.51	0.10	Phosphatidylcholines
lysoPC a C20:3	lysoPC a C20:3	-5.11	1.37E-04	7.04E-04	-0.90	0.18	Lyso-Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C40:4	PC ae C40:4	5.33	1.50E-04	7.52E-04	0.40	0.07	Phosphatidylcholines
C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/e16:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:5	4.96	1.92E-04	9.32E-04	0.45	0.09	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C44:5	PC ae C44:5	-4.79	2.38E-04	1.13E-03	-0.53	0.11	Phosphatidylcholines
2,6 Dimethylheptanoylcarnitine (Nonaylcarnitine)	C9	-4.92	2.56E-04	1.18E-03	-0.41	0.08	Acylcarnitines
C24:0 Sphingomyelin	SM C24:0	5.92	2.78E-04	1.25E-03	0.58	0.10	Sphingomyelins
C(16:0/20:1;11Z-enoyl) aa Phosphatidylcholine	PC aa C36:1	5.19	3.44E-04	1.47E-03	0.79	0.15	Phosphatidylcholines
Leucine	Leu	-4.70	3.44E-04	1.47E-03	-0.61	0.13	Amino Acids
Hydroxysphingomyelin with acyl residue sum C16:1	SM (OH) C16:1	4.75	4.15E-04	1.73E-03	0.52	0.11	Sphingomyelins
Octanoylcarnitine	C8	-4.66	5.02E-04	2.04E-03	-0.30	0.07	Acylcarnitines
Isoleucine	Ile	-4.41	5.73E-04	2.28E-03	-0.63	0.14	Amino Acids
C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	PC aa C38:5	4.49	6.77E-04	2.63E-03	0.35	0.08	Phosphatidylcholines
C(16:0/18:4;9Z,11E,13E,15Z-tetraenoyl) aa L-Phosphatidylcholine	PC aa C34:4	4.19	7.91E-04	3.01E-03	0.47	0.11	Phosphatidylcholines
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-pentaenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C38:6	4.24	9.32E-04	3.47E-03	0.24	0.06	Phosphatidylcholines
Hydroxytetradecenoylcarnitine	C14:1-OH	4.12	9.50E-04	3.47E-03	0.50	0.12	Acylcarnitines
Glutarylcarnitine	C5-DC (C6-OH)	-4.06	1.21E-03	4.31E-03	-0.52	0.13	Acylcarnitines
Dodecanedioylcarnitine	C12-DC	-4.05	1.26E-03	4.41E-03	-0.26	0.06	Acylcarnitines
C(20:3;11Z,14Z.17Z-trienoyl/20:0) aa L-Phosphatidylcholine	PC aa C40:3	3.85	1.61E-03	5.32E-03	0.42	0.11	Phosphatidylcholines
C(18:1;9E-enoyl/14:0) aa Phosphatidylcholine	PC aa C32:1	3.86	1.56E-03	5.32E-03	0.59	0.15	Phosphatidylcholines
Myristylcarnitine (Tetradecanoylcarnitine)	C14	3.87	1.61E-03	5.32E-03	0.62	0.16	Acylcarnitines
C(18:1;9Z,12Z-dienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C34:2	3.84	1.65E-03	5.34E-03	0.25	0.06	Phosphatidylcholines
Hexenoylcarnitine	C6:1	-3.81	1.71E-03	5.45E-03	-0.40	0.11	Acylcarnitines
Phenylalanine	Phe	-3.46	3.53E-03	1.10E-02	-0.51	0.15	Amino Acids

Significantly altered metabolites between Ex4 vs. CR in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Hydroxysphingomyelin with acyl residue sum C22:2	SM (OH) C22:2	3.79	4.20E-03	1.29E-02	0.64	0.17	Sphingomyelins
Carnosine	Carnosine	-3.35	4.62E-03	1.39E-02	-0.54	0.16	Biogenic Amines
C(16:0/20:4;Z,8E,11E,14E,17E-tetraenoyl) aa L-Phosphatidylcholine	PC aa C36:4	3.30	5.55E-03	1.65E-02	0.28	0.09	Phosphatidylcholines
C(20:3;8Z,11Z,14Z-trienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:3	3.26	7.38E-03	2.15E-02	0.20	0.06	Phosphatidylcholines
Hydroxyhexadecenoylcarnitine	C16:1-OH	3.17	7.74E-03	2.22E-02	0.45	0.14	Acylcarnitines
C16:1;9Z-enoyl a Lysophosphatidylcholine	lysoPC a C16:1	-3.06	8.12E-03	2.25E-02	-0.61	0.20	Lyso-Phosphatidylcholines
Hydroxyhexadecadienylcarnitine	C16:2-OH	3.09	8.04E-03	2.25E-02	0.21	0.07	Acylcarnitines
Tryptophan	Trp	-3.03	8.48E-03	2.32E-02	-0.45	0.15	Amino Acids
Palmitoleylcarnitine (Hexadecenoylcarnitine)	C16:1	3.07	1.06E-02	2.86E-02	0.56	0.18	Acylcarnitines
Stearoylcarnitine (Octadecanoylcarnitine)	C18	2.91	1.12E-02	2.97E-02	0.48	0.16	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C42:1	PC ae C42:1	2.92	1.14E-02	2.97E-02	0.32	0.11	Phosphatidylcholines
Palmitoylcarnitine	C16	2.88	1.33E-02	3.42E-02	0.50	0.17	Acylcarnitines
Malonylcarnitine	C3-DC (C4-OH)	-2.73	1.56E-02	3.95E-02	-0.61	0.22	Acylcarnitines
C(18:3;9Z,12Z,15Z-trienoyl/14:0) aa Phosphatidylcholine	PC aa C32:3	2.67	1.90E-02	4.75E-02	0.41	0.15	Phosphatidylcholines

Table A32: Significantly altered metabolites between Ex4 and H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C18:1;11Z-enoyl Sphingomyelin	SM C18:1	-26.17	6.39E-13	1.12E-10	1.84	0.07	Sphingomyelins
Phosphatidylcholine with acyl-alkyl residue sum C42:3	PC ae C42:3	22.43	3.19E-12	2.79E-10	-2.22	0.10	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C40:1	PC ae C40:1	23.63	8.30E-12	4.84E-10	-1.86	0.08	Phosphatidylcholines
C16:0 Sphingomyelin	SM C16:0	-13.27	4.96E-09	2.17E-07	1.48	0.11	Sphingomyelins
C18:1;9Z,12Z-dienoyl a Lysophosphatidylcholine	lysoPC a C18:2	12.70	6.62E-09	2.32E-07	-1.89	0.15	Lyso-Phosphatidylcholines
Aldohexose	H1	12.09	9.47E-09	2.76E-07	-1.83	0.15	Hexoses
C(20:4;5E,8E,11E,14E-tetraenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:4	-16.39	1.19E-08	2.96E-07	1.24	0.08	Phosphatidylcholines
Methylglutarylcarnitine	C5-M-DC	-13.18	1.53E-08	3.34E-07	3.09	0.23	Acylcarnitines
Hydroxysphingomyelin with acyl residue sum C14:1 (C14:1-OH Sphingomyelin)	SM (OH) C14:1	-10.72	3.96E-08	7.71E-07	1.19	0.11	Sphingomyelins
Hydroxysphingomyelin with acyl residue sum C16:1	SM (OH) C16:1	-10.45	5.76E-08	9.20E-07	1.30	0.12	Sphingomyelins
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-pentaenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C38:6	-10.42	5.78E-08	9.20E-07	0.66	0.06	Phosphatidylcholines
C(20:4;(5Z,8Z,11Z,14Z-tetraenoyl/e18:1;9Z-en) ae L-Phosphatidylcholine	PC ae C38:5	-12.87	8.23E-08	1.06E-06	1.39	0.11	Phosphatidylcholines
Acetyl-L-ornithine	Ac-Orn	10.92	7.61E-08	1.06E-06	-1.05	0.10	Biogenic Amines
Alanine	Ala	10.25	8.45E-08	1.06E-06	-1.20	0.12	Amino Acids
C(12:0/18:0) aa Phosphatidylcholine	PC aa C30:0	-10.08	1.05E-07	1.22E-06	1.07	0.11	Phosphatidylcholines
Valine	Val	9.83	1.17E-07	1.28E-06	-1.10	0.11	Amino Acids
C18:0 Sphingomyelin	SM C18:0	-10.15	1.59E-07	1.64E-06	1.83	0.18	Sphingomyelins
C(16:0;e18:1;9Z-en) ae Phosphatidylcholine	PC ae C34:1	-9.57	3.10E-07	3.02E-06	0.70	0.07	Phosphatidylcholines
C24:1 Sphingomyelin	SM C24:1	-9.36	3.82E-07	3.30E-06	1.60	0.17	Sphingomyelins
lysoPC a C20:3	lysoPC a C20:3	15.43	3.67E-07	3.30E-06	-1.66	0.11	Lyso-Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C44:5	PC ae C44:5	10.07	3.96E-07	3.30E-06	-1.52	0.15	Phosphatidylcholines
Stearoylcarnitine (Octadecanoylcarnitine)	C18	-8.73	4.95E-07	3.94E-06	1.53	0.18	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C42:2	PC ae C42:2	9.77	6.43E-07	4.89E-06	-1.17	0.12	Phosphatidylcholines
C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/e16:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:5	-8.49	8.51E-07	6.21E-06	0.75	0.09	Phosphatidylcholines
Sphingomyelin with acyl residue sum C16:1	SM C16:1	-8.97	2.89E-06	2.02E-05	0.94	0.11	Sphingomyelins
Spermidine	Spermidine	-7.29	4.00E-06	2.69E-05	1.05	0.14	Biogenic Amines
Phosphatidylcholine with acyl-alkyl residue sum C40:2	PC ae C40:2	7.37	7.70E-06	4.99E-05	-0.85	0.12	Phosphatidylcholines
Hydroxytetradecenoylcarnitine	C14:1-OH	-7.18	9.26E-06	5.79E-05	0.98	0.14	Acylcarnitines

Significantly altered metabolites between Ex4 vs. H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(16:0/18:4;9Z,11E,13E,15Z-tetraenoyl) aa L-Phosphatidylcholine	PC aa C34:4	-6.77	1.08E-05	6.32E-05	0.80	0.12	Phosphatidylcholines
C16:1;9Z-enoyl a Lysophosphatidylcholine	lysoPC a C16:1	6.70	1.06E-05	6.32E-05	-1.21	0.18	Lyso-Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C38:2	PC ae C38:2	6.62	1.29E-05	7.25E-05	-0.61	0.09	Phosphatidylcholines
C(20:4;5E,8E,11E,14E-tetraenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C38:4	-7.35	2.63E-05	1.40E-04	0.69	0.09	Phosphatidylcholines
Isoleucine	Ile	6.16	2.60E-05	1.40E-04	-0.67	0.11	Amino Acids
C(22:4;7Z,10Z,13Z,16Z,19Z-tetraenoyl/18:1;9Z-enoyl) aa Phosphatidylcholine	PC aa C40:5	-7.02	2.96E-05	1.52E-04	1.15	0.16	Phosphatidylcholines
(5Z,8Z,11Z,14Z) a C20:4 Lysophosphatidylcholine	lysoPC a C20:4	6.05	3.05E-05	1.53E-04	-0.75	0.12	Lyso-Phosphatidylcholines
Spermine	Spermine	5.94	4.34E-05	2.05E-04	-0.62	0.10	Biogenic Amines
Putrescine	Putrescine	-8.23	4.28E-05	2.05E-04	2.38	0.29	Biogenic Amines
Hydroxyhexadecanoylcarnitine	C16-OH	-6.35	4.57E-05	2.11E-04	0.87	0.14	Acylcarnitines
C(18:1;9Z-enoyl/e14:0) ae L-Phosphatidylcholine	PC ae C32:1	-6.14	4.89E-05	2.19E-04	0.78	0.13	Phosphatidylcholines
C(22:5;7Z,10Z,13Z,16Z,19Z-pentaenoyl/18:1;11Z-enoyl) aa L-Phosphatidylcholine	PC aa C40:6	-6.40	6.54E-05	2.86E-04	0.85	0.13	Phosphatidylcholines
Palmitoylcarnitine	C16	-6.09	7.89E-05	3.37E-04	1.10	0.18	Acylcarnitines
C16:0 a Lysophosphatidylcholine	lysoPC a C16:0	5.72	8.23E-05	3.43E-04	-0.46	0.08	Lyso-Phosphatidylcholines
C(20:3;5Z,8Z,11Z-trienoyl/18:1;9Z-enoyl) aa L-Phosphatidylcholine	PC aa C38:4	-5.75	8.69E-05	3.54E-04	0.91	0.16	Phosphatidylcholines
Leucine	Leu	5.43	9.32E-05	3.71E-04	-0.59	0.11	Amino Acids
C18:1;9E-enoyl a Lysophosphatidylcholine	lysoPC a C18:1	7.12	9.75E-05	3.79E-04	-0.97	0.14	Lyso-Phosphatidylcholines
C17:0 a Lysophosphatidylcholine	lysoPC a C17:0	5.42	1.42E-04	5.41E-04	-0.58	0.11	Lyso-Phosphatidylcholines
Hexenoylcarnitine	C6:1	5.45	1.87E-04	6.82E-04	-0.84	0.15	Acylcarnitines
C(22:5;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C40:6	-5.20	1.84E-04	6.82E-04	0.40	0.08	Phosphatidylcholines
Hydroxyhexadecenoylcarnitine	C16:1-OH	-5.56	2.08E-04	7.44E-04	0.92	0.17	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C42:4	PC ae C42:4	5.23	2.33E-04	8.14E-04	-1.07	0.20	Phosphatidylcholines
Octanoylcarnitine	C8	5.28	3.45E-04	1.16E-03	-0.42	0.08	Acylcarnitines
Butenylcarnitine	C4:1	4.77	3.40E-04	1.16E-03	-0.44	0.09	Acylcarnitines
Hydroxysphingomyelin with acyl residue sum C22:1	SM (OH) C22:1	-4.83	3.64E-04	1.20E-03	0.54	0.11	Sphingomyelins
Hydroxyoctadecenoylcarnitine	C18:1-OH	-4.79	4.58E-04	1.48E-03	0.89	0.18	Acylcarnitines
Butyrylcarnitine	C4	4.46	5.35E-04	1.70E-03	-1.15	0.26	Acylcarnitines
Phenylalanine	Phe	4.38	7.29E-04	2.28E-03	-0.56	0.13	Amino Acids

Significantly altered metabolites between Ex4 vs. H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Hydroxysphingomyelin with acyl residue sum C22:2	SM (OH) C22:2	-4.90	7.81E-04	2.40E-03	0.83	0.17	Sphingomyelins
Octadecadienylcarnitine	C18:2	-4.88	8.17E-04	2.46E-03	0.84	0.17	Acylcarnitines
C(16:0/20:4;Z,8E,11E,14E,17E-tetraenoyl) aa L-Phosphatidylcholine	PC aa C36:4	-4.29	8.58E-04	2.54E-03	0.32	0.07	Phosphatidylcholines
C(20:1;9Z-enoyl/e18:0) ae L-Phosphatidylcholine	PC ae C38:1	4.27	9.03E-04	2.63E-03	-0.45	0.11	Phosphatidylcholines
C(12:0/26:0) aa Phosphatidylcholine	PC aa C38:0	-4.28	9.34E-04	2.68E-03	0.44	0.10	Phosphatidylcholines
Glutamic acid	Glu	-4.31	1.02E-03	2.88E-03	0.70	0.16	Amino Acids
C(22:1;7Z,10Z,13Z,16Z,19Z-pentanoyl/20:0) aa L-Phosphatidylcholine	PC aa C42:5	-4.34	1.09E-03	2.97E-03	0.47	0.11	Phosphatidylcholines
Hexanoylcarnitine (Fumarylcarnitine)	C6 (C4:1-DC)	4.32	1.08E-03	2.97E-03	-1.06	0.25	Acylcarnitines
C(18:3;9Z,12Z,15Z-trienoyl/14:0) aa Phosphatidylcholine	PC aa C32:3	-4.31	1.26E-03	3.39E-03	0.78	0.18	Phosphatidylcholines
C(22:5;4Z,7Z,10Z,13Z,16Z-pentaenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C40:5	-4.25	1.34E-03	3.56E-03	0.43	0.10	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C44:4	PC ae C44:4	4.38	1.46E-03	3.81E-03	-0.64	0.15	Phosphatidylcholines
C(16:0/18:2;2E,4E-dienoyl) aa L-Phosphatidylcholine	PC aa C34:2	4.49	1.79E-03	4.60E-03	-0.31	0.07	Phosphatidylcholines
C(20:3;11Z,14Z.17Z-trienoyl/20:0) aa L-Phosphatidylcholine	PC aa C40:3	3.74	2.22E-03	5.62E-03	-0.39	0.11	Phosphatidylcholines
Tryptophan	Trp	3.75	2.60E-03	6.50E-03	-0.44	0.12	Amino Acids
Dodecanoylcarnitine	C12	-3.66	2.71E-03	6.68E-03	0.72	0.20	Acylcarnitines
C(20:0/e18:0) ae L-Phosphatidylcholine	PC ae C38:0	4.28	2.75E-03	6.69E-03	-0.53	0.12	Phosphatidylcholines
Pimelylcarnitine	C7-DC	3.95	3.02E-03	7.24E-03	-0.77	0.19	Acylcarnitines
Oleoylcarnitine (Octadecenoylcarnitine)	C18:1	-3.79	3.43E-03	8.11E-03	1.11	0.29	Acylcarnitines
C(20:3;8Z,11Z,14Z-trienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:3	3.62	3.51E-03	8.20E-03	-0.24	0.07	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C40:4	PC ae C40:4	-3.70	4.15E-03	9.57E-03	0.36	0.10	Phosphatidylcholines
C(22:0/e20:0) ae L-Phosphatidylcholine	PC ae C42:0	3.58	4.66E-03	1.06E-02	-0.31	0.09	Phosphatidylcholines
Myristylcarnitine (Tetradecanoylcarnitine)	C14	-3.52	4.81E-03	1.08E-02	0.76	0.22	Acylcarnitines
Hydroxypropionylcarnitine	C3-OH	-3.46	5.05E-03	1.12E-02	0.45	0.13	Acylcarnitines
Hexadecadienylcarnitine	C16:2	-3.37	5.32E-03	1.16E-02	0.44	0.13	Acylcarnitines
Tetradecadienylcarnitine	C14:2	-3.33	5.55E-03	1.20E-02	0.41	0.12	Acylcarnitines
C(18:0/14:0) aa Phosphatidylcholine	PC aa C32:0	-3.38	6.33E-03	1.35E-02	0.42	0.12	Phosphatidylcholines
C(18:3;6Z,9Z,12Z-trienoyl/18:0) aa Phosphatidylcholine	PC aa C36:3	3.17	6.87E-03	1.45E-02	-0.42	0.13	Phosphatidylcholines
C(18:1;11Z-enoyl/18:4;2E,4E,6E,11Z-tetraenoyl) aa Phosphatidylcholine	PC aa C36:5	3.63	7.84E-03	1.63E-02	-0.35	0.10	Phosphatidylcholines

Significantly altered metabolites between Ex4 vs. H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
Acetylcarnitine	C2	3.40	9.50E-03	1.96E-02	-0.38	0.11	Acylcarnitines
Dodecanedioylcarnitine	C12-DC	2.94	1.30E-02	2.61E-02	-0.17	0.06	Acylcarnitines
C(18:1;9Z-enoyl/e18:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:2	2.86	1.29E-02	2.61E-02	-0.19	0.07	Phosphatidylcholines
Hydroxyproline	OH-Pro	-2.95	1.63E-02	3.24E-02	0.43	0.15	Biogenic Amines
Carnitine	C0	2.76	1.67E-02	3.29E-02	-0.44	0.16	Acylcarnitines
Glutarylcarnitine	C5-DC (C6-OH)	-2.75	1.72E-02	3.35E-02	0.35	0.13	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C42:5	PC ae C42:5	-2.63	1.96E-02	3.77E-02	0.12	0.05	Phosphatidylcholines
2-Methylbutyrylcarnitine	C5	2.65	2.10E-02	4.00E-02	-0.53	0.20	Acylcarnitines

Table A33: Significantly altered metabolites between CR and H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C18:1;11Z-enoyl Sphingomyelin	SM C18:1	-10.03	2.14E-08	3.75E-06	0.95	0.09	Sphingomyelins
Alanine	Ala	8.54	1.48E-07	8.64E-06	-1.02	0.12	Amino Acids
Hydroxysphingomyelin with acyl residue sum C14:1 (C14:1-OH Sphingomyelin)	SM (OH) C14:1	-9.02	1.43E-07	8.64E-06	0.95	0.11	Sphingomyelins
C(20:3;8Z,11Z,14Z-trienoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:3	8.25	2.99E-07	1.31E-05	-0.44	0.05	Phosphatidylcholines
Methylglutarylcarnitine	C5-M-DC	-7.43	1.10E-06	3.47E-05	2.41	0.32	Acylcarnitines
C(20:3;11Z,14Z.17Z-trienoyl/20:0) aa L-Phosphatidylcholine	PC aa C40:3	7.30	1.25E-06	3.47E-05	-0.81	0.11	Phosphatidylcholines
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/20:0) aa Phosphatidylcholine	PC aa C42:6	7.71	1.39E-06	3.47E-05	-0.63	0.08	Phosphatidylcholines
C(22:6;4Z,7Z,10Z,13Z,16Z,19Z-pentaenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C38:6	-7.14	2.75E-06	6.02E-05	0.42	0.06	Phosphatidylcholines
Hydroxysphingomyelin with acyl residue sum C16:1	SM (OH) C16:1	-6.95	4.42E-06	8.60E-05	0.78	0.11	Sphingomyelins
Spermine	Spermine	-6.45	6.06E-06	1.06E-04	0.82	0.13	Biogenic Amines
Stearoylcarnitine (Octadecanoylcarnitine)	C18	-6.06	1.45E-05	2.13E-04	1.05	0.17	Acylcarnitines
C24:0 Sphingomyelin	SM C24:0	7.43	1.46E-05	2.13E-04	-0.78	0.11	Sphingomyelins
Spermidine	Spermidine	-5.97	1.64E-05	2.21E-04	0.88	0.15	Biogenic Amines
Phosphatidylcholine with acyl-alkyl residue sum C44:5	PC ae C44:5	6.22	2.34E-05	2.73E-04	-0.99	0.16	Phosphatidylcholines
C16:0 Sphingomyelin	SM C16:0	-5.84	2.34E-05	2.73E-04	0.71	0.12	Sphingomyelins
Glutarylcarnitine	C5-DC (C6-OH)	-5.56	3.49E-05	3.60E-04	0.86	0.16	Acylcarnitines
C18:0 Sphingomyelin	SM C18:0	-5.87	3.35E-05	3.60E-04	0.85	0.14	Sphingomyelins
C(22:0/e20:0) ae L-Phosphatidylcholine	PC ae C42:0	6.49	4.38E-05	4.26E-04	-0.57	0.09	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C42:2	PC ae C42:2	5.33	6.07E-05	5.59E-04	-0.91	0.17	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C42:3	PC ae C42:3	5.75	6.70E-05	5.86E-04	-1.03	0.18	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C38:2	PC ae C38:2	5.17	1.01E-04	8.42E-04	-0.59	0.11	Phosphatidylcholines
C(18:0;12Z-enoyl/22:4;7Z,10Z,13Z,16Z-tetraenoyl) aa Phosphatidylcholine	PC aa C40:4	5.57	1.18E-04	9.35E-04	-0.72	0.13	Phosphatidylcholines
Sphingomyelin with acyl residue sum C16:1	SM C16:1	-5.32	1.85E-04	1.41E-03	0.58	0.11	Sphingomyelins
Butyrylcarnitine	C4	4.91	1.96E-04	1.43E-03	-1.17	0.24	Acylcarnitines
Phosphatidylcholine with acyl-alkyl residue sum C40:1	PC ae C40:1	5.08	2.30E-04	1.61E-03	-0.82	0.16	Phosphatidylcholines
C(18:3;6Z,9Z,12Z-trienoyl/18:0) aa Phosphatidylcholine	PC aa C36:3	4.75	2.63E-04	1.77E-03	-0.61	0.13	Phosphatidylcholines
Aldohexose	H1	4.86	3.03E-04	1.97E-03	-1.30	0.27	Hexoses
C24:1 Sphingomyelin	SM C24:1	-4.52	3.52E-04	2.20E-03	0.66	0.15	Sphingomyelins

Significantly altered metabolites between CR vs. H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(20:0/e18:0) ae L-Phosphatidylcholine	PC ae C38:0	4.42	3.74E-04	2.26E-03	-0.30	0.07	Phosphatidylcholines
lysoPC a C20:3	lysoPC a C20:3	5.06	4.47E-04	2.61E-03	-0.76	0.15	Lyso-Phosphatidylcholines
Creatinine	Creatinine	-4.29	5.55E-04	3.13E-03	0.48	0.11	Biogenic Amines
C(18:1;9E-enoyl/14:0) aa Phosphatidylcholine	PC aa C32:1	4.52	6.62E-04	3.51E-03	-0.60	0.13	Phosphatidylcholines
C(18:1;11Z-enoyl/18:4;2E,4E,6E,11Z-tetraenoyl) aa Phosphatidylcholine	PC aa C36:5	4.74	6.54E-04	3.51E-03	-0.48	0.10	Phosphatidylcholines
C(16:0/18:2;2E,4E-dienoyl) aa L-Phosphatidylcholine	PC aa C34:2	4.23	6.98E-04	3.59E-03	-0.22	0.05	Phosphatidylcholines
C(18:1;9Z-enoyl/e18:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:2	4.10	7.58E-04	3.79E-03	-0.27	0.07	Phosphatidylcholines
Acetyl-L-ornithine	Ac-Orn	4.12	7.90E-04	3.84E-03	-0.43	0.11	Biogenic Amines
Glycine	Gly	-4.04	8.94E-04	4.23E-03	0.46	0.11	Amino Acids
Hydroxysphingomyelin with acyl residue sum C22:1	SM (OH) C22:1	-3.91	1.27E-03	5.86E-03	0.35	0.09	Sphingomyelins
Phosphatidylcholine with acyl-alkyl residue sum C40:2	PC ae C40:2	3.82	1.41E-03	6.34E-03	-0.61	0.16	Phosphatidylcholines
C(20:4;5E,8E,11E,14E-tetraenoyl/e16:0) ae L-Phosphatidylcholine	PC ae C36:4	-4.03	1.56E-03	6.83E-03	0.38	0.10	Phosphatidylcholines
C(20:1;13Z-enoyl/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C38:3	3.76	1.73E-03	7.39E-03	-0.87	0.23	Phosphatidylcholines
C(20:4;(5Z,8Z,11Z,14Z-tetraenoyl/e18:1;9Z-en) ae L-Phosphatidylcholine	PC ae C38:5	-3.87	1.86E-03	7.73E-03	0.50	0.13	Phosphatidylcholines
Hexanoylcarnitine (Fumarylcarnitine)	C6 (C4:1-DC)	3.70	2.96E-03	1.21E-02	-0.91	0.25	Acylcarnitines
C(20:4;5Z,8Z,11Z,14Z-tetraenoyl/e16:1;1Z-en) ae L-Phosphatidylcholine	PC ae C36:5	-3.43	3.20E-03	1.27E-02	0.31	0.09	Phosphatidylcholines
C(12:0/18:0) aa Phosphatidylcholine	PC aa C30:0	-3.43	3.52E-03	1.37E-02	0.37	0.11	Phosphatidylcholines
Decadienylcarnitine	C10:2	-3.30	4.24E-03	1.60E-02	0.21	0.06	Acylcarnitines
Hydroxypropionylcarnitine	C3-OH	-3.43	4.30E-03	1.60E-02	0.47	0.14	Acylcarnitines
Octadecadienylcarnitine	C18:2	-3.28	5.25E-03	1.91E-02	0.66	0.20	Acylcarnitines
Malonylcarnitine	C3-DC (C4-OH)	-3.17	5.62E-03	2.01E-02	0.76	0.24	Acylcarnitines
Hydroxytetradecenoylcarnitine	C14:1-OH	-3.10	6.77E-03	2.30E-02	0.48	0.15	Acylcarnitines
C(18:0/18:2;9Z,12Z-dienoyl) aa L-Phosphatidylcholine	PC aa C36:2	3.19	6.84E-03	2.30E-02	-0.39	0.12	Phosphatidylcholines
C(16:0/e20:0) ae L-Phosphatidylcholine	PC ae C36:0	-3.13	6.81E-03	2.30E-02	0.40	0.13	Phosphatidylcholines
Asymmetric dimethylarginine	ADMA	-3.02	7.69E-03	2.49E-02	1.01	0.33	Biogenic Amines
Acetylcarnitine	C2	3.22	7.68E-03	2.49E-02	-0.37	0.12	Acylcarnitines
C(18:2;9Z,12Z-dienoyl/16:1;7Z-enoyl) aa Phosphatidylcholine	PC aa C34:3	2.85	1.12E-02	3.50E-02	-0.33	0.12	Phosphatidylcholines
C(12:0/26:0) aa Phosphatidylcholine	PC aa C38:0	-3.00	1.11E-02	3.50E-02	0.23	0.08	Phosphatidylcholines

Significantly altered metabolites between CR vs. H>C in the liver samples.

MOLECULE	SHORT NAME	T	P-VALUE	ADJUSTED P-VALUE	FOLD CHANGE	SE	METABOLON CLASS
C(22:5;4Z,7Z,10Z,13Z,16Z,19Z-hexaenoyl/e18:0) ae L-Phosphatidylcholine	PC ae C40:6	-2.82	1.20E-02	3.69E-02	0.29	0.10	Phosphatidylcholines
Glutamic acid	Glu	-2.81	1.24E-02	3.74E-02	0.54	0.19	Amino Acids
Phosphatidylcholine with acyl-alkyl residue sum C42:1	PC ae C42:1	2.80	1.33E-02	3.93E-02	-0.44	0.16	Phosphatidylcholines
C16:1;9Z-enoyl a Lysophosphatidylcholine	lysoPC a C16:1	2.74	1.41E-02	4.11E-02	-0.60	0.22	Lyso-Phosphatidylcholines
Carnosine	Carnosine	-2.73	1.44E-02	4.12E-02	0.43	0.16	Biogenic Amines
Phosphatidylcholine with acyl-alkyl residue sum C42:4	PC ae C42:4	2.70	1.55E-02	4.36E-02	-0.79	0.29	Phosphatidylcholines
C16:0 a Lysophosphatidylcholine	lysoPC a C16:0	-2.70	1.57E-02	4.36E-02	0.23	0.08	Lyso-Phosphatidylcholines
Hexenoylcarnitine	C6:1	2.71	1.77E-02	4.66E-02	-0.44	0.16	Acylcarnitines
Citrulline	Cit	-2.72	1.78E-02	4.66E-02	0.56	0.20	Amino Acids
C(16:1;9E-enoyl/16:1;9E-enoyl) aa Phosphatidylcholine	PC aa C32:2	2.63	1.78E-02	4.66E-02	-0.42	0.16	Phosphatidylcholines
Phosphatidylcholine with acyl-alkyl residue sum C38:3	PC ae C38:3	2.66	1.77E-02	4.66E-02	-0.29	0.11	Phosphatidylcholines
Palmitoylcarnitine	C16	-2.61	1.85E-02	4.75E-02	0.59	0.23	Acylcarnitines

Table A34: Reactome pathways enriched by genes present in edges of the CR network

ID	DESCRIPTION	COUNT	BG COUNT	FDR	GENE ID
MMU-1430728	Metabolism	29	1685	0.0195	*AK157302/Acaa2/Acat1/Acsf2/Acsm5/Agmo/B4galt3/Bcat2/Cyb5b/Decr1/Decr2/Eno1/Fdx1/Fh1/Gatm/Gnai2/Grhpr/Gusb/Hsd17b12/Man2b2/Mdh1/Nme2/Paics/Papss1/Pik3ca/Psmd9/Serinc2/Slc22a1/Tyms*

Table A35: Reactome pathways enriched by genes present in edges of the Ex4 network

ID	DESCRIPTION	COUNT	BG COUNT	FDR	GENE ID
MMU-556833	Metabolism of lipids	27	615	3.14e-05	*Acad10/Acat1/Acox3/Acsf2/Acsf3/Akr1c6/Aldh3a2/Cyp2d34/Enpp6/Ept1/Fabp7/Ggps1/Hmgcs2/Hsd17b12/Hsd17b13/Lpcat1/Lpcat3/Mgll/Mid1ip1/Mmaa/Mtm1/Mtmr9/Ocrl/Pi4k2b/Pla1a/Scap/Tspo*
MMU-1430728	Metabolism	46	1685	0.00035	*Acad10/Acat1/Acox3/Acsf2/Acsf3/Adi1/Akr1c6/Aldh3a2/B4galt1/Cmbl/Cyp2d34/Dguok/Eno1/Enpp6/Ept1/Fabp7/Ggct/Ggps1/Grhpr/Gstp1/Gusb/Hmgcs2/Hsd17b12/Hsd17b13/Iscu/Kmo/Lpcat1/Lpcat3/Lrat/Mgll/Mid1ip1/Mmaa/Mocs1/Mtm1/Mtmr9/Nup133/Ocrl/Pgls/Pi4k2b/Pla1a/Psmd3/Scap/Slc19a2/Slc25a21/Tspo/mt-Nd3*
MMU-194315	Signaling by Rho GTPases	15	349	0.0131	*Arhgap17/Arhgap23/Cyba/Cyfip1/Gdi2/Iqgap1/Nup133/Ocrl/Pfn1/Rhoa/Rhoc/Rhoj/Rhou/Tuba4a/Ywhab*
MMU-1483257	Phospholipid metabolism	10	173	0.0173	*Enpp6/Ept1/Lpcat1/Lpcat3/Mgll/Mtm1/Mtmr9/Ocrl/Pi4k2b/Pla1a*
MMU-194840	Rho GTPase cycle	8	121	0.0282	*Arhgap17/Arhgap23/Gdi2/Ocrl/Rhoa/Rhoc/Rhoj/Rhou*

Table A36: SNPs associated with LRGs of CR graph

	BMI	Type 2 diabetes
Pisd	X	X
Hsd17b12	X	X
Cystm1	X	

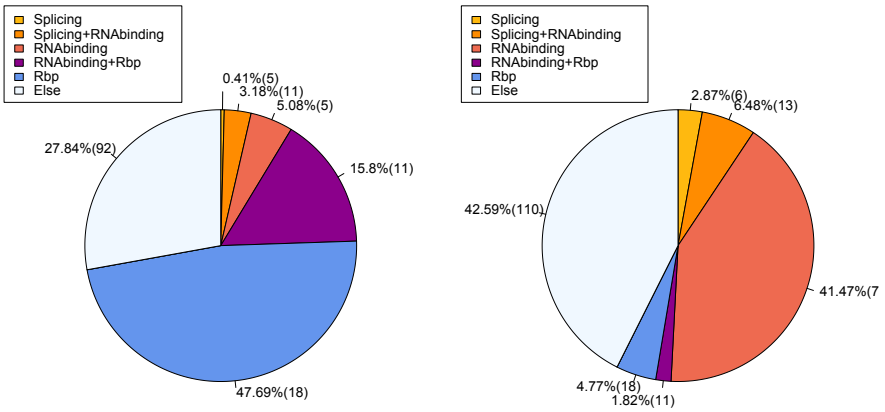
Table A37: SNPs associated with LRGs of Ex4 graph

	BMI	HDL cholesterol	Triglycerides	Fasting glucose
Nrbf2	X	X	X	
Tspo				X
Cast	X			
Cdc23	X			
Mknk2	X			
Scap	X			
Thap3	X			

Table A38: Metabolites influenced by the three LRGs selected for the Ex4 and CR networks.

SHORT NAME	MOLECULE	METABOLON CLASS	FOLD CHANGE	P-VALUE	ADJUSTED P-VALUE	LRG
C5.DC..C6.OH.	Glutarylcarnitine	Acylcarnitines	-0.515287934	0.001207698	0.004313208	*Cyba*
PC.aa.C40.2	C(20:1;11E-enoyl/20:1;11E-enoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.373809508	0.025051772	0.057685001	*Cyba*
PC.ae.C40.3	Phosphatidylcholine with acyl-alkyl residue sum C40:3	Phosphatidylcholines	0.249149564	0.024649566	0.057515654	*Cyba*
PC.ae.C40.5	C(22:5;4Z,7Z,10Z,13Z,16Z-pentaenoyl/e18:0) ae L-Phosphatidylcholine	Phosphatidylcholines	0.206677093	0.07358569	0.134755736	*Cyba*
PC.aa.C42.4	Phosphatidylcholine with diacyl residue sum C42:4	Phosphatidylcholines	0.44279081	0.022627417	0.053510782	*Cyba*
Ile	Isoleucine	Amino Acids	-0.626218372	0.000572769	0.00227806	*Magix*
Lys	Lysine	Amino Acids	-0.111954684	0.619678603	0.704180231	*Magix*
Leu	Leucine	Amino Acids	-0.60734028	0.000344155	0.001469221	*Magix*
Val	Valine	Amino Acids	-0.946678204	1.74E-05	0.000138153	*Magix*
lysoPC.a.C14.0	C14:0 a Lysophosphatidylcholine	Lyso-Phosphatidylcholines	-0.130123949	0.053392006	0.102676934	*Magix*
lysoPC.a.C16.0	C16:0 a Lysophosphatidylcholine	Lyso-Phosphatidylcholines	-0.687545399	2.08E-06	2.60E-05	*Magix, Nrbf2*
lysoPC.a.C16.1	C16:1;9Z-enoyl a Lysophosphatidylcholine	Lyso-Phosphatidylcholines	-0.610277577	0.008116629	0.02254619	*Magix, Nrbf2*
C3.1	Propenylcarnitine	Acylcarnitines	-0.610277577	0.008116629	0.02254619	*Magix*
lysoPC.a.C20.4	(5Z,8Z,11Z,14Z) a C20:4 Lysophosphatidylcholine	Lyso-Phosphatidylcholines	-1.074249143	2.84E-06	3.31E-05	*Magix, Nrbf2*
Phe	Phenylalanine	Amino Acids	-0.506966176	0.003532449	0.011038903	*Magix*
Pro	Proline	Amino Acids	-0.303818903	0.039196422	0.080698516	*Magix*
PC.aa.C34.4	C(16:0/18:4;9Z,11E,13E,15Z-tetraenoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	0.466929492	0.000791176	0.003009909	*Magix*
PC.aa.C36.6	C(18:2;9Z,12Z-dienoyl/18:4;9E,11E,13E,15E-tetraenoyl) aa L-Phosphatidylcholine	Phosphatidylcholines	-0.045980197	0.720629914	0.793146132	*Nrbf2*

A.4 SUPPLEMENTARY FIGURES - PART 2



(a) Commercial AB (b) PBS AB + 3P

Figure A1: Comparison of protein abundances detected in samples extracted with the **(a)** Commercial antibody and the **(b)** PBS antibody + 3P. The number in the brackets gives the number of quantified proteins.

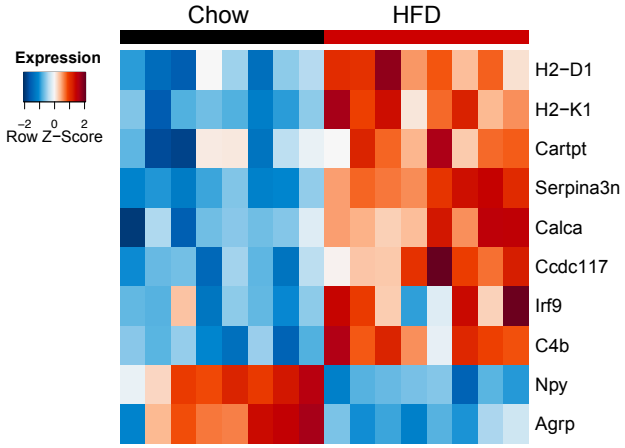
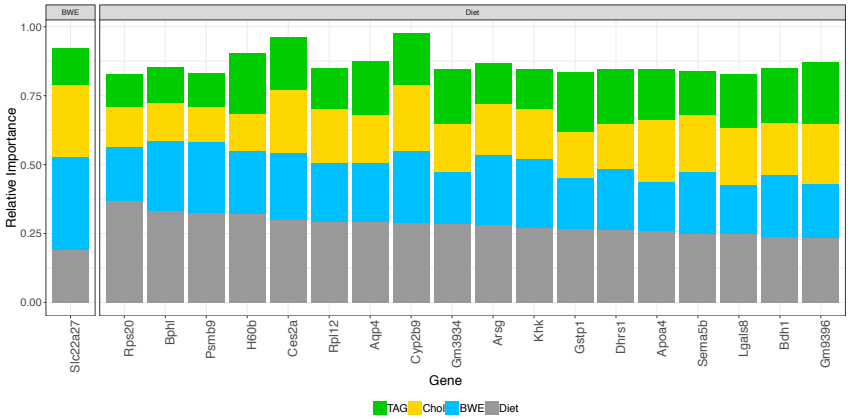
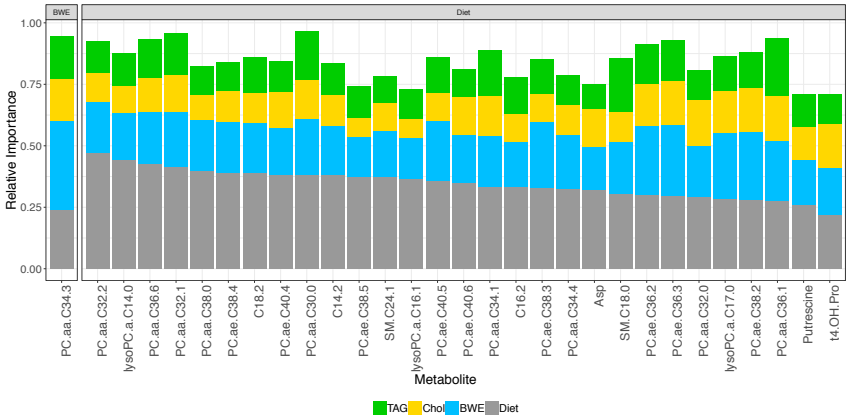


Figure A2: Significantly differentially expressed genes between HFD- and chow-fed mice.

A.5 SUPPLEMENTARY FIGURES - PART 3



(a) Relative importance of factors in genes



(b) Relative importance of factors in metabolites

Figure A3: Variance explained by phenotypic features for **(a)** genes and **(b)** metabolites in our dataset.

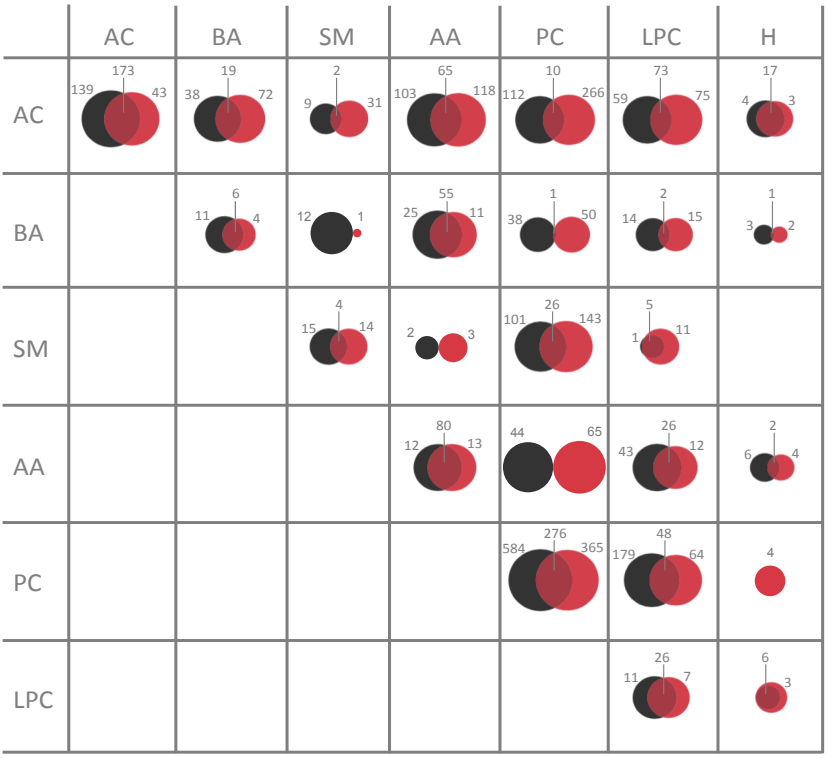
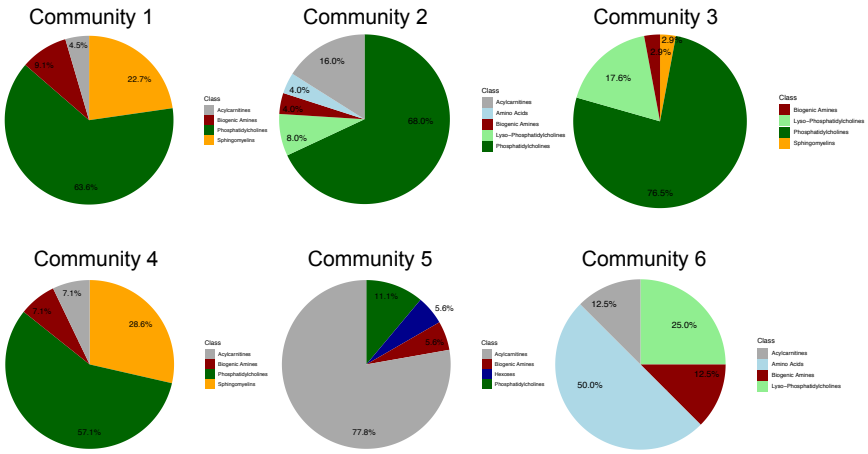
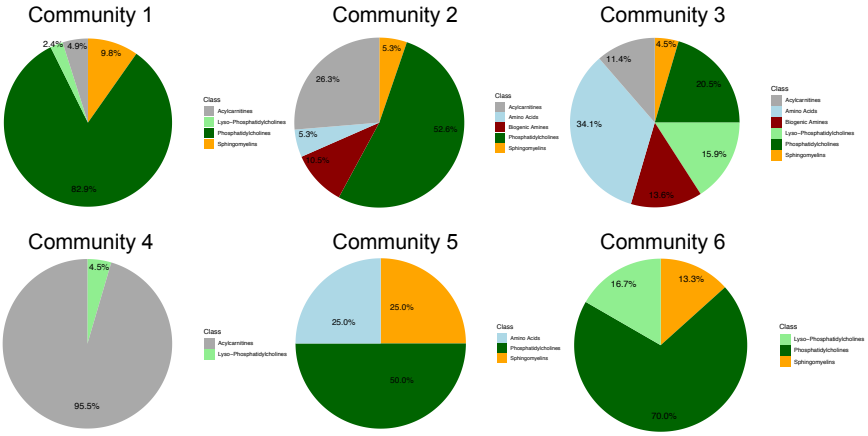


Figure A4: Class composition of correlated metabolite pairs obtained for the Chow (black) and HFD (red) setting.



(a) Chow Communities



(b) HFD Communities

Figure A5: Class composition of communities shown for the **(a)** Chow and **(b)** HFD networks.

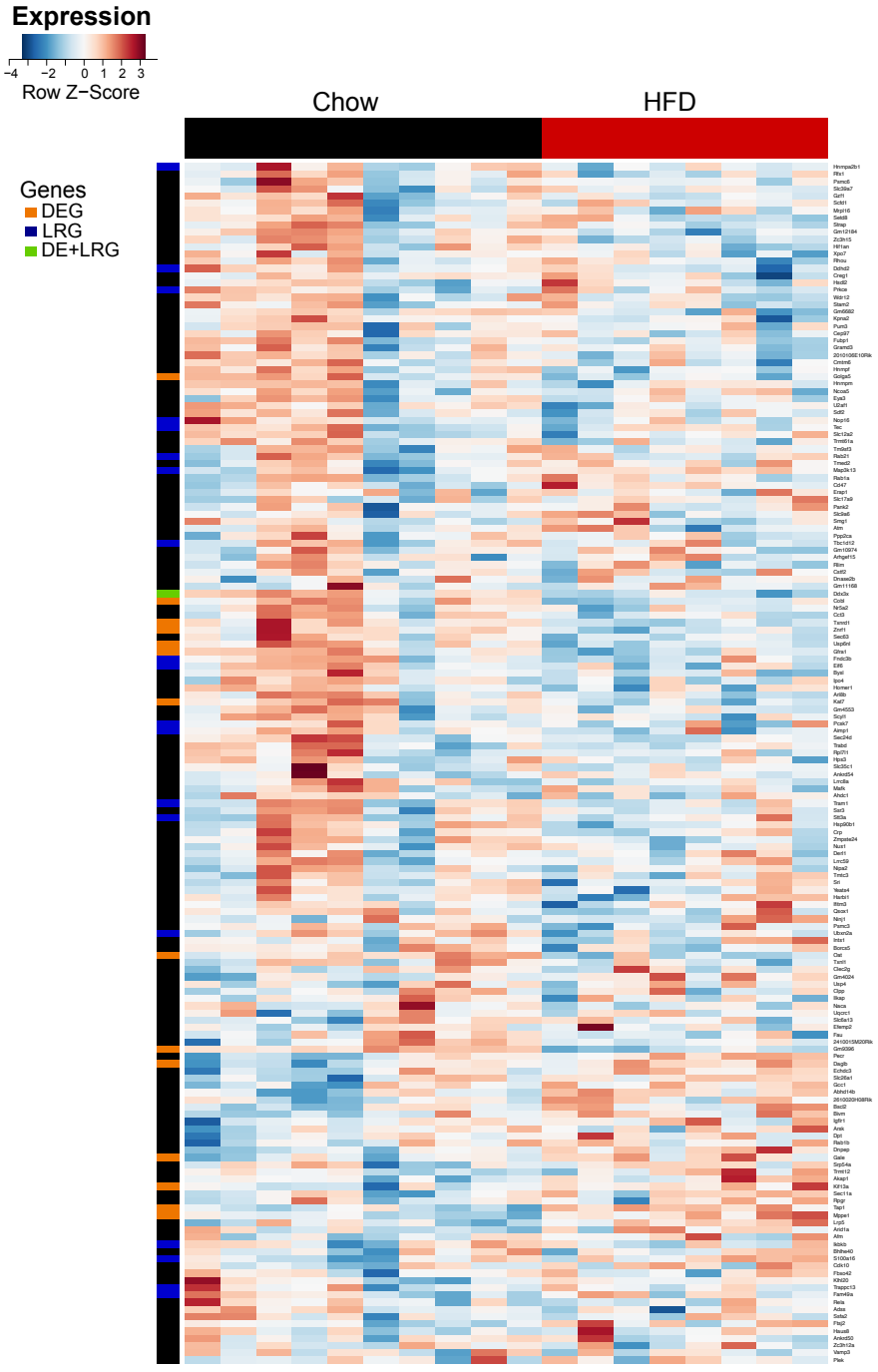


Figure A6: Genes present in the edges of the Chow network. LRGs are coloured in blue, differentially expressed genes in orange, and differentially expressed LRGs in green.

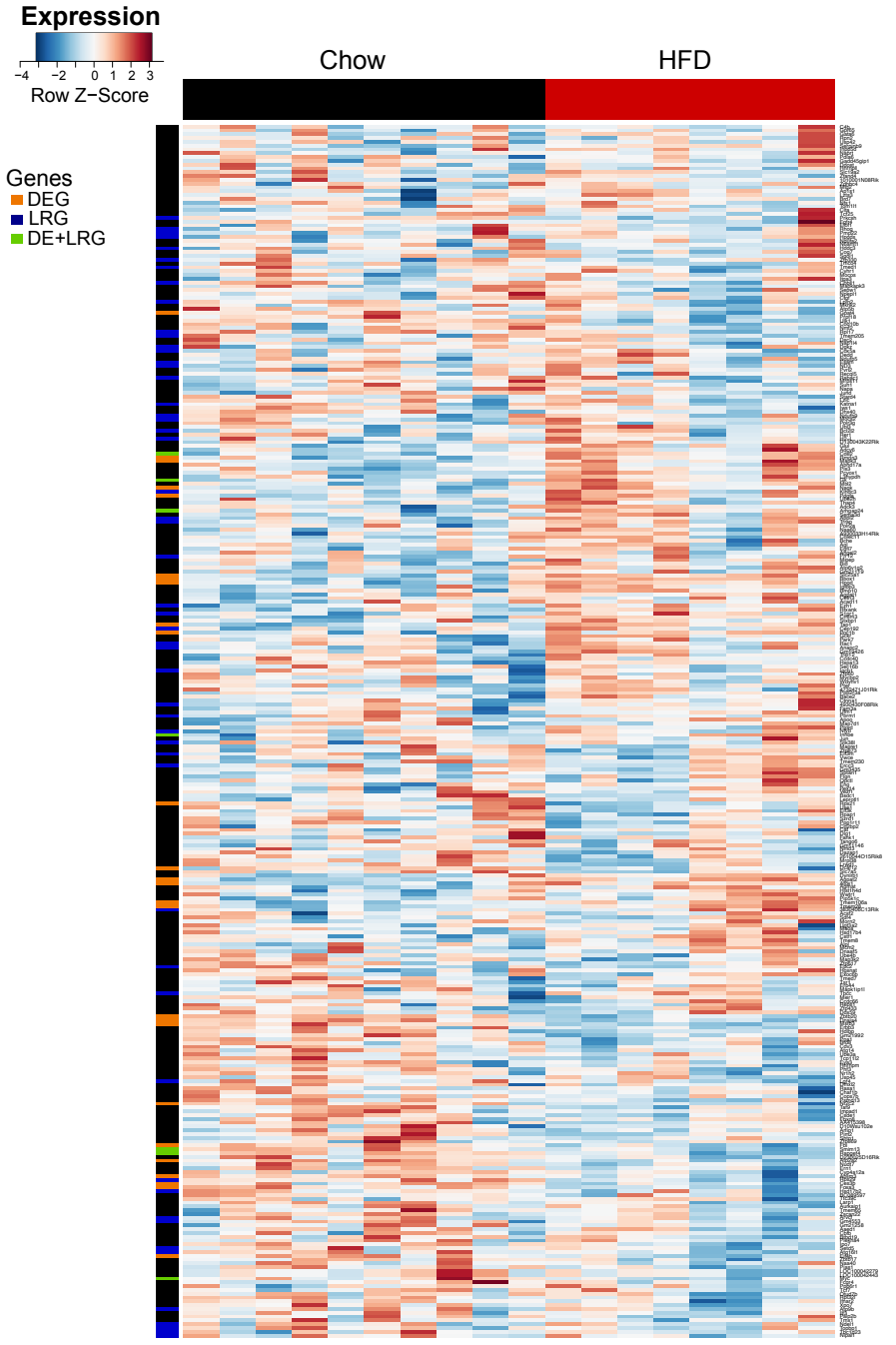


Figure A7: Genes present in the edges of the HFD network. LRGs are coloured in blue, differentially expressed genes in orange, and differentially expressed LRGs in green.

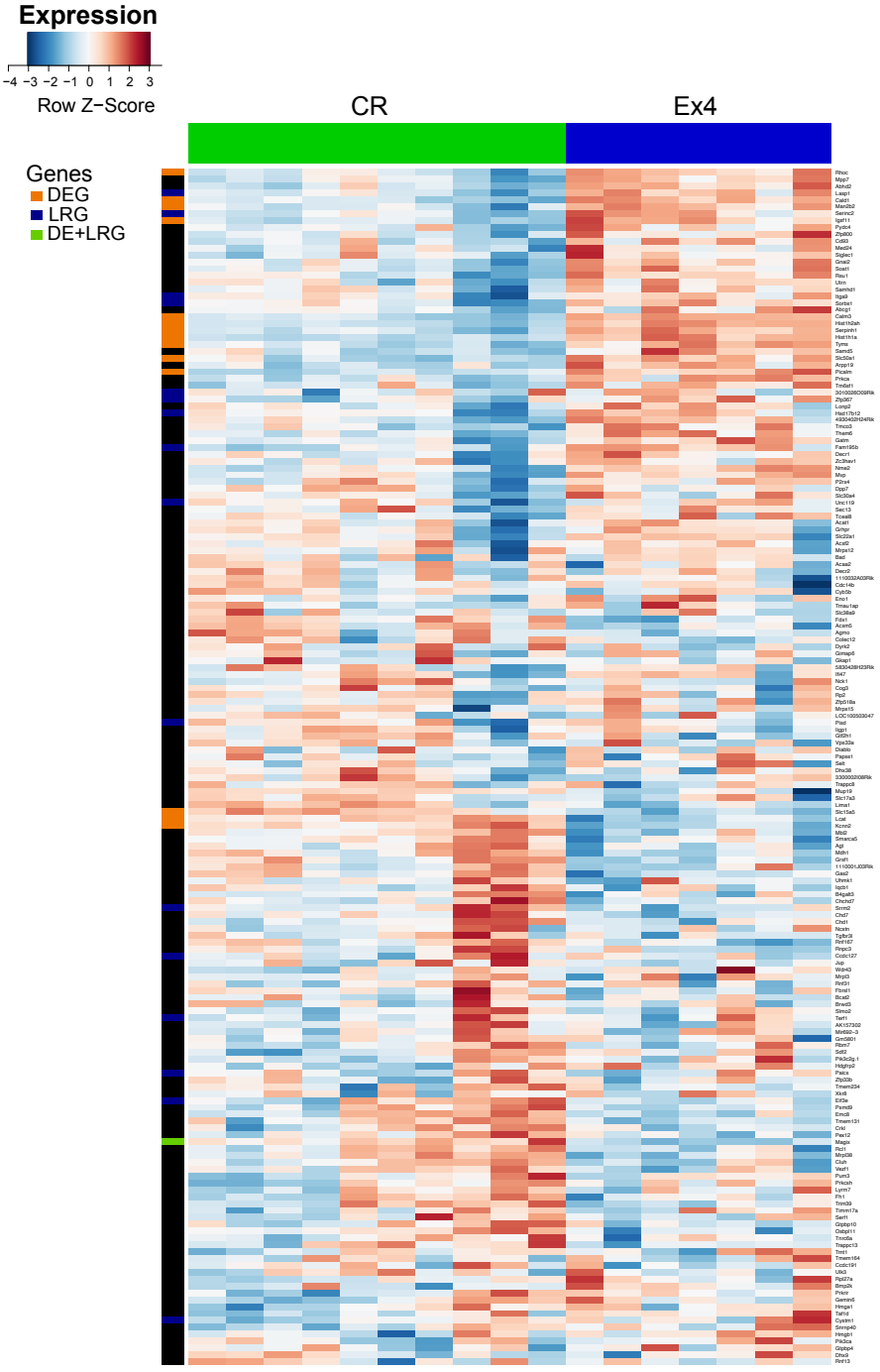


Figure A8: Genes present in the edges of the CR network. LRGs are coloured in blue, differentially expressed genes in orange, and differentially expressed LRGs in green.

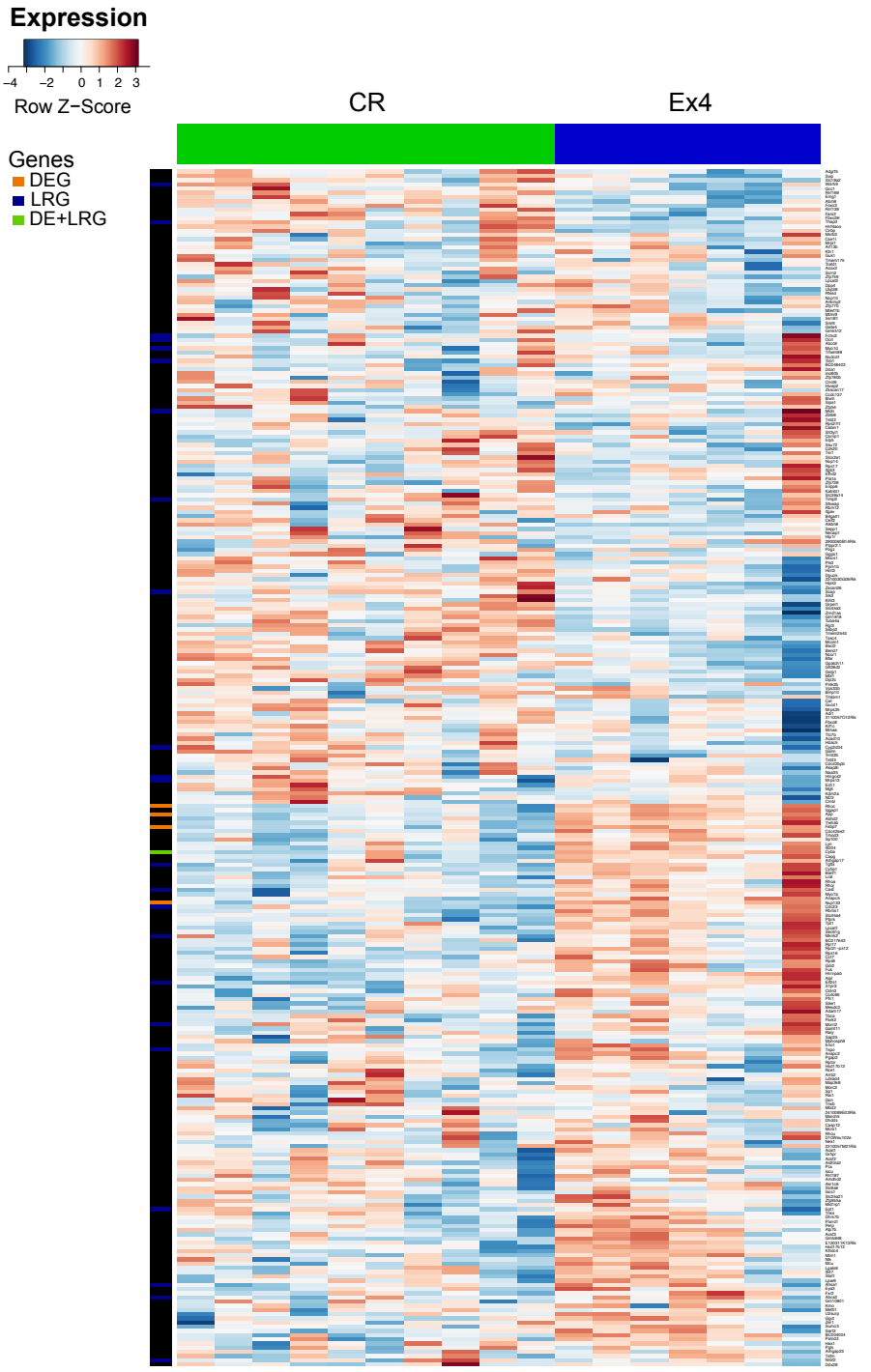


Figure A9: Genes present in the edges of the Ex4 network. LRGs are coloured in blue, differentially expressed genes in orange, and differentially expressed LRGs in green.

B

LIST OF POSTERS AND PUBLICATIONS

POSTERS:

Schriever SC, **Klaus V**, Lutter D, Pfluger PT: Differential effects of weight loss by calorie restriction vs exendin-4 on liver metabolism and eWAT inflammation, presented at the *Keystone Symposia on Molecular and Cellular Biology*, Track: , Banff, Canada, 2019

Klaus V, Schriever SC, Pfluger PT, Lutter D: Elucidating genetic control of metabolic networks under different dietary conditions, presented at the *International Conference on Systems Biology 2018*, Track: Multi-omics, Lyon, France, 2018

Klaus V, Lutter D: RNA-Seq read count correction applying intron signal estimation, presented at the *Intelligent Systems for Molecular Biology/European Conference on Computational Biology 2017*, Track: High Throughput Sequencing Algorithms & Application, Prague, Czech Republic, 2017

PUBLICATIONS:

Quarta C, Fisette A, Xu Y, Colldén G, Legutko B, Tseng YT, Reim A, Wierer M, De Rosa MC, **Klaus V**, Rausch R, Thaker VV, Graf E, Strom TM, Poher AL, Gruber T, Le Thuc O, Cebrian-Serrano A, Kabra D, Bellocchio L, Woods SC, Pflugfelder GO, Nogueiras R, Zeltser L, Grunwald Kadow IC, Moon A, García-Cáceres C, Mann M, Treier M, Doege CA, Tschöp MH: Functional identity of hypothalamic melanocortin neurons depends on Tbx3. *Nature Metabolism*, 2019

Klaus V*, Bastek H*, Damme K, Collins LB, Frötschl R, Benda N, Lutter D, Ellinger-Ziegelbauer H, Swenberg JA, Dietrich DR, Stemmer K: Time-matched analysis of DNA adduct formation and early gene expression as predictive tool for renal carcinogenesis in methylazoxymethanol acetate treated Eker rats. *Arch Toxicol*, 91(10):3427-3438, 2017.
*Equal contribution

Schriever SC, Zimprich A, Pfuhlmann K, Baumann P, Giesert F, **Klaus V**, Kabra DG, Hafen U, Romanov A, Tschöp MH, Wurst W, Conrad M, Hölter SM, Vogt Weisenhorn D, Pfluger PT: Alterations in

neuronal control of body weight and anxiety behavior by glutathione peroxidase 4 deficiency. *Neuroscience*, 357:241-254, 2017

Manuscript submitted:

Klaus VS*, Schriever SC*, Peter A, Monroy Kuhn JM, Irmler M, Torkaz J, Prehn C, Kastenmüller G, Beckers J, Adamski J, Königsrainer A, Müller TD, Heni M, Tschöp MH, Pfluger PT, Lutter D: Correlation guided Network Integration (CoNI) reveals novel genetic regulators of hepatic lipid metabolism.
*Equal contribution

BIBLIOGRAPHY

[1] Christian C. Abnet. "Carcinogenic Food Contaminants." In: *Cancer Investigation* 25.3 (2007), pp. 189–96.

[2] Isabel C. Abreu, Joyce F. Guerra, Renata R. Pereira, Maisa Silva, Wanderson G. Lima, Marcelo E. Silva, and Maria L. Pedrosa. "Hypercholesterolemic diet induces hepatic steatosis and alterations in mRNA expression of NADPH oxidase in rat livers." In: *Arquivos Brasileiros de Endocrinologia & Metabologia* 58 (2014), pp. 251–9.

[3] Cynthia A. Afshari, Hisham K. Hamadeh, and Pierre R. Bushel. "The evolution of bioinformatics in toxicology: advancing toxicogenomics." In: *Toxicological sciences : an official journal of the Society of Toxicology* 120 Suppl 1 (2011), pp. S225–37.

[4] Saeb Aliwaini, Jenna Bleloch, Serah Kimani, Sharon Prince, and M. A. Hayat. "Chapter 12 - Induction of Autophagy and Apoptosis in Melanoma Treated With Palladacycle Complexes." In: Academic Press, 2016, pp. 231–47.

[5] Elizabeth Allen, Annick Moing, Timothy Md Ebbels, Mickaël Maucourt, A. Deri Tomos, Dominique Rolin, and Mark A. Hooks. "Correlation Network Analysis reveals a sequential reorganization of metabolic and transcriptional states during germination and gene-metabolite relationships in developing seedlings of Arabidopsis." In: *BMC systems biology* 4 (2010), p. 62.

[6] Margaret B. Allison, Warren Pan, Alexander MacKenzie, Christa Patterson, Kimi Shah, Tammy Barnes, Wenwen Cheng, Alan Rupp, David P. Olson, and Martin G. Myers. "Defining the Transcriptional Targets of Leptin Reveals a Role for Atf3 in Leptin Action." In: *Diabetes* 67.6 (2018), pp. 1093–104.

[7] Muhammad Amir and Mark J. Czaja. "Autophagy in nonalcoholic steatohepatitis." In: *Expert review of gastroenterology & hepatology* 5.2 (2011), pp. 159–66.

[8] Simon Anders. *Counting reads in features with htseq-count*. `https://htseq.readthedocs.io/en/release_0.11.1/count.html`. Accessed: 2015-07-15.

[9] Simon Anders, Paul Theodor Pyl, and Wolfgang Huber. "HTSeq - a Python framework to work with high-throughput sequencing data." In: *Bioinformatics* 31.2 (2014), pp. 166–9.

[10] Simon Andrews. *FastQC A Quality Control tool for High Throughput Sequence Data*. `http://www.bioinformatics.babraham.ac.uk/projects/fastqc/`. 2014.

[11] Ina Aretz and David Meierhofer. "Advantages and Pitfalls of Mass Spectrometry Based Metabolome Profiling in Systems Biology." In: *International journal of molecular sciences* 17.5 (2016), p. 632.

[12] Shailendra Arindkar, Jashdeep Bhattacharjee, Jerald Mahesh Kumar, Barun Das, Pramod Upadhyay, Shajahan Asif, Ramesh C. Juyal, Subeer S. Majumdar, and Nagarajan Perumal. "Antigen peptide transporter 1 is involved in the development of fructose-induced hepatic steatosis in mice." In: *Journal of Gastroenterology and Hepatology* 28.8 (2013), pp. 1403–9.

[13] Carmen Arnal, Jose M. Lou-Bonafonte, María V. Martínez-Gracia, María J. Rodríguez-Yoldi, and Jesús Osada. "Transcriptomics and Nutrition in Mammalians." In: *Genomics, Proteomics and Metabolomics in Nutraceuticals and Functional Foods*. John Wiley & Sons, Ltd, 2015. Chap. 46, pp. 581–608. ISBN: 9781118930458.

[14] M. Ashburner et al. "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." In: *Nature genetics* 25.1 (2000), pp. 25–9.

[15] John-Paul Baird, Angela Choe, Jasmine L. Loveland, Janine Beck, Carrie E. Mahoney, Julia S. Lord, and Lindsay A. Grigg. "Orexin-A hyperphagia: hindbrain participation in consummatory feeding responses." In: *Endocrinology* 150.3 (2009), pp. 1202–16.

[16] Susan Barlow and Josef Schlatter. "Risk assessment of carcinogens in food." In: *Toxicology and Applied Pharmacology* 243.2 (2010), pp. 180–90.

[17] Tanya Barrett et al. "NCBI GEO: archive for functional genomics data sets—update." In: *Nucleic Acids Research* 41.D1 (2012), pp. D991–5.

[18] Alberto A. Barrios-Correa, Jose A. Estrada, and Irazu Contreras. "Leptin Signaling in the Control of Metabolism and Appetite: Lessons from Animal Models." In: *Journal of molecular neuroscience* 66 (2018), pp. 390–402.

[19] Jörg Bartel et al. "The Human Blood Metabolome-Transcriptome Interface." In: *PLoS genetics* 11.6 (2015), pp. e1005274.

[20] Diane Benford et al. "Application of the Margin of Exposure (MOE) approach to substances in food that are genotoxic and carcinogenic." In: *Food and Chemical Toxicology* 48 (2010), pp. S2–24.

[21] Yoaf Benjamini and Yosef Hochberg. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." In: *Journal of the Royal Statistical Society. Series B* 57.1 (1995), pp. 289–300.

[22] Hans-Rudolf Berthoud, Heike Münzberg, and Christopher D. Morrison. "Blaming the Brain for Obesity: Integration of Hedonic and Homeostatic Mechanisms." In: *Gastroenterology* 152.7 (2017), pp. 1728–38.

[23] Sivakama S. Bharathi et al. "Sirtuin 3 (SIRT3) Protein Regulates Long-chain Acyl-CoA Dehydrogenase by Deacetylating Conserved Lysines Near the Active Site." In: *Journal of Biological Chemistry* 288.47 (2013), pp. 33837–47.

[24] Christian Bjørbæk, Hugh J. Lavery, Sarah H. Bates, Ryan K. Olson, Sarah M. Davis, Jeffrey S. Flier, and Martin G. Myers. "SOCS3 Mediates Feedback Inhibition of the Leptin Receptor via Tyr985." In: *Journal of Biological Chemistry* 275.51 (2000), pp. 40649–57.

[25] Christian Bjørbæk, Joel K. Elmquist, J. Daniel Frantz, Steven E. Shoelson, and Jeffrey S. Flier. "Identification of SOCS-3 as a Potential Mediator of Central Leptin Resistance." In: *Molecular Cell* 1.4 (1998), pp. 619–25.

[26] Elena G. Bochukova et al. "A Transcriptomic Signature of the Hypothalamic Response to Fasting and BDNF Deficiency in Prader-Willi Syndrome." In: *Cell Reports* 22.13 (2018), pp. 3401–8.

[27] Ann M. Bode and Zigang Dong. "Toxic Phytochemicals and Their Potential Risks for Human Cancer." In: *Cancer Prevention Research* 8.1 (2015), pp. 1–8.

[28] Carlo E. Bonferroni. "Teoria statistica delle classi e calcolo delle probabilità." In: *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze.* (1936).

[29] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." In: *CA: A Cancer Journal for Clinicians* 68.6 (2018), pp. 394–424.

[30] Paul Brazhnik, Alberto de la Fuente, and Pedro Mendes. "Gene networks: how to put the function in genomics." In: *Trends in Biotechnology* 20.11 (2002), pp. 467–72.

[31] Alvis Brazma, Arek Kasprzyk, Bart De Moor, Sean Davis, Steffen Durinck, Wolfgang Huber, and Yves Moreau. "BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis." In: *Bioinformatics* 21.16 (2005), pp. 3439–40.

[32] Jean-Ánthelme Brillat-Savarin. *Physiologie du Goût, ou méditations de gastronomie transcendante*. Ouvrage théorique, historique et à l'ordre du jour. Dédié aux gastronomes parisiens. Quatrième Édition. Paris: Just Tessier, Libraire, 1834.

[33] Joerg M. Buescher and Edward M. Driggers. "Integration of omics: more than the sum of its parts." In: *Cancer & metabolism* 4 (2016), p. 4.

[34] Alice Cambiaghi, Manuela Ferrario, and Marco Masseroli. "Analysis of metabolomic data: tools, current strategies and future challenges for omics data integration." In: *Briefings in Bioinformatics* 18.3 (2016), pp. 498–510.

[35] John N. Campbell et al. "A molecular census of arcuate hypothalamus and median eminence cell types." In: *Nature Neuroscience* 20 (2017), pp. 484–96.

[36] Marc Carlson. *rat2302.db: Affymetrix Rat Genome 230 2.0 Array annotation data (chip rat2302)*. R package version 3.2.3. 2016.

[37] Rachel Cavill, Danyel Jennen, Jos Kleinjans, and Jacob Jan Briedé. "Transcriptomic and metabolomic data integration." In: *Briefings in Bioinformatics* 17.5 (2015), pp. 891–901.

[38] Ke Chen et al. "Induction of leptin resistance through direct interaction of C-reactive protein with leptin." In: *Nature Medicine* 12.4 (2006), pp. 425–32.

[39] Renchao Chen, Xiaoji Wu, Lan Jiang, and Yi Zhang. "Single-Cell RNA-Seq Reveals Hypothalamic Cell Diversity." In: *Cell Reports* 18.13 (2017), pp. 3227–41.

[40] Siyu Chen, Wenxiang Zhang, Chunqi Tang, Xiaoli Tang, Li Liu, and Chang Liu. "Vanin-1 Is a Key Activator for Hepatic Gluconeogenesis." In: *Diabetes* 63.6 (2014), p. 2073.

[41] Aaron Clauset, M. E .J. Newman, and Christopher Moore. *Finding community structure in very large networks*. `http://www.arxiv.org/abs/cond-mat/0408187`. Accessed 15th Dez 2018.

[42] Rebecca A. Clewell, Bin Sun, Yeyejide Adeleye, Paul Carmichael, Alina Efremenko, Patrick D. McMullen, Salil Pendse, O. J. Trask, Andy White, and Melvin E. Andersen. "Profiling Dose-Dependent Activation of p53-Mediated Signaling Pathways by Chemicals with Distinct Mechanisms of DNA Damage." In: *Toxicological Sciences* 142.1 (2014), pp. 56–73.

[43] Clontech. *SMARTer® Universal Low Input RNA Kit for Sequencing*. `http://info.clontech.com/rs/clontech/images/smarterUniversal_AN_0413_633600_IN.pdf`. Accessed: August 2015.

[44] Jeremy F. L. Cobbold et al. "Phenotyping murine models of non-alcoholic fatty liver disease through metabolic profiling of intact liver tissue." In: *Clinical Science* 116.5 (2009), p. 403.

[45] Samuel M. Cohen and Leon B. Ellwein. "Cell proliferation in carcinogenesis." In: *Science* 249.4972 (1990), pp. 1007–11.

[46] Laura K. Cole, Jean E. Vance, and Dennis E. Vance. "Phosphatidylcholine biosynthesis and lipoprotein metabolism." In: *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 1821.5 (2012), pp. 754–61.

[47] Mouse Genome Sequencing Consortium and NCBI. *Jul. 2007 assembly of the mouse genome*. `http://hgdownload.cse.ucsc.edu/goldenPath/mm9/chromosomes/`. Downloaded: Feb 2015. 2007.

[48] The Gene Ontology Consortium. "Expansion of the Gene Ontology knowledgebase and resources." In: *Nucleic acids research* 45.D1 (2017), pp. D331–8.

[49] Cyril Couturier et al. "Silencing of OB-RGRP in mouse hypothalamic arcuate nucleus increases leptin receptor signaling and prevents diet-induced obesity." In: *Proceedings of the National Academy of Sciences of the United States of America* 104.49 (2007), pp. 19476–81.

[50] Gabor Csardi and Tamas Nepusz. "The igraph software package for complex network research." In: *InterJournal Complex Systems* 1695 (2006).

[51] Mark J. Czaja. "Autophagy in health and disease. 2. Regulation of lipid metabolism and storage by autophagy: pathophysiological implications." In: *American journal of physiology. Cell physiology* 298.5 (2010), pp. C973–8.

[52] T. H. M. Da Costa, M. K. Ito, and Benjamin Caballero. "PHOSPHOLIPIDS | Physiology." In: *Encyclopedia of Food Sciences and Nutrition (Second Edition)*. Oxford: Academic Press, 2003, pp. 4523–31.

[53] A. M. De Livera and J. B. Bowne. *metabolomics: Analysis of Metabolomics Data. R package version 0.1.4.* `https://CRAN.R-project.org/package=metabolomics`. 2014.

[54] Alev Deli, Emanuel Kreidl, Stefan Santifaller, Barbara Trotter, Katja Seir, Walter Berger, Rolf Schulte-Hermann, Chantal Rodgarkia-Dara, and Michael Grusch. "Activins and activin antagonists in hepatocellular carcinoma." In: *World journal of gastroenterology* 14.11 (2008), pp. 1699–709.

[55] Don Delker et al. "Molecular biomarkers of oxidative stress associated with bromate carcinogenicity." In: *Toxicology* 221.2 (2006), pp. 158–65.

[56] Yalan Deng et al. "SGK1/FOXO3 Signaling in Hypothalamic POMC Neurons Mediates Glucocorticoid-Increased Adiposity." In: *Diabetes* 67.4 (2018), p. 569.

[57] Adel Derghal et al. "Leptin is required for hypothalamic regulation of miRNAs targeting POMC 3'UTR." In: *Frontiers in Cellular Neuroscience* 9 (2015), p. 172.

[58] Davina Derous, Sharon E. Mitchell, Cara L. Green, Luonan Chen, Jing-Dong J. Han, Yingchun Wang, Daniel E. L. Promislow, David Lusseau, John R. Speakman, and Alex Douglas. "The effects of graded levels of calorie restriction: VI. Impact of short-term graded calorie restriction on transcriptomic responses of the hypothalamic hunger and circadian signaling pathways." In: *Aging* 8.4 (2016), pp. 642–63.

[59] Birk Diedenhofen and Jochen Musch. "cocor: A Comprehensive Solution for the Statistical Comparison of Correlations." In: *PLOS ONE* 10.4 (2015), pp. e0121945.

[60] Xiaokun Ding, Neeraj K. Saxena, Songbai Lin, Nitika Arora Gupta, and Frank A. Anania. "Exendin-4, a glucagon-like protein-1 (GLP-1) receptor agonist, reverses hepatic steatosis in ob/ob mice." In: *Hepatology (Baltimore, Md.)* 43.1 (2006), pp. 173–81.

[61] Michael S. Donaldson. "Nutrition and cancer: A review of the evidence for an anti-cancer diet." In: *Nutrition Journal* 3.1 (2004), p. 19.

[62] Christoph Dorn, Julia C. Engelmann, Michael Saugspier, Andreas Koch, Arndt Hartmann, Martina Müller, Rainer Spang, Anja Bosserhoff, and Claus Hellerbrand. "Increased expression of c-Jun in nonalcoholic fatty liver disease." In: *Laboratory Investigation* 94 (2014), pp. 394EP–.

[63] Olive J. Dunn and Virgina A. Clark. "Correlation coefficients measured on the same individuals." In: *Journal of the American Statistical Association* 64 (1969), pp. 366–77.

[64] Steffen Durinck, Paul T. Spellman, Ewan Birney, and Wolfgang Huber. "Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt." In: *Nature Protocols* 4 (2009), pp. 1184–91.

[65] Kenneth A. Dyar et al. "Atlas of Circadian Metabolism Reveals System-wide Coordination and Communication between Clocks." In: *Cell* 174.6 (2018), pp. 1571–85.e11.

[66] Ron Edgar, Michael Domrachev, and Alex E. Lash. "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository." In: *Nucleic acids research* 30.1 (2002), pp. 207–10.

[67] L. Eriksson, E. Johansson, N. Kettaneh-Wold, and S. Wold. "Scaling." In: *Introduction to multi- and megavariate data analysis using projection methods (PCA & PLS)*. Umetrics, 1999, pp. 213–25.

[68] Antonio Fabregat et al. "The Reactome Pathway Knowledgebase." In: *Nucleic acids research* 46.D1 (2018), pp. D649–55.

[69] Mohd Farhan, Haitao Wang, Uma Gaur, Peter J. Little, Jiangping Xu, and Wenhua Zheng. "FOXO Signaling Pathways as Therapeutic Targets in Cancer." In: *Int J Biol Sci* 13.7 (2017), pp. 815–27.

[70] Robert D. Finn, Colin J. Henderson, Claire L. Scott, and C. Roland Wolf. "Unsaturated fatty acid regulation of cytochrome P450 expression via a CAR-dependent pathway." In: *The Biochemical journal* 417.1 (2009), pp. 43–54.

[71] Douglas A. Formolo, Joana M. Gaspar, Hiago M. Melo, Tuany Eichwald, Ramiro Javier Zepeda, Alexandra Latini, Michael S. Okun, and Roger Walz. "Deep Brain Stimulation for Obesity: A Review and Future Directions." In: *Frontiers in Neuroscience* 13 (2019), p. 323.

[72] Adam Frankish et al. "Ensembl 2018." In: *Nucleic Acids Research* 46.D1 (2017), pp. D754–61.

[73] Alberto de la Fuente, Nan Bing, Ina Hoeschele, and Pedro Mendes. "Discovery of meaningful associations in genomic data using partial correlation coefficients." In: *Bioinformatics* 20.18 (2004), pp. 3565–74.

[74] Brian L. Furman. "The development of Byetta (exenatide) from the venom of the Gila monster as an anti-diabetic agent." In: *Toxicon* 59.4 (2012), pp. 464–71.

[75] Dimos Gaidatzis, Lukas Burger, Maria Florescu, and Michael B. Stadler. "Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation." In: *Nature Biotechnology* 33 (2015), pp. 722–9.

[76] Cristina García-Cáceres, Esther Fuente-Martín, Jesús Argente, and Julie A. Chowen. "Emerging role of glial cells in the control of body weight." In: *Molecular metabolism* 1.1-2 (2012), pp. 37–46.

[77] Cristina García-Cáceres et al. "Role of astrocytes, microglia, and tanycytes in brain control of systemic metabolism." In: *Nature Neuroscience* 22.1 (2019), pp. 7–14.

[78] Robert C. Gentleman et al. "Bioconductor: open software development for computational biology and bioinformatics." In: *Genome biology* 5.10 (2004), pp. R80.

[79] Yves Gibon, Bjoern Usadel, Oliver E Blaesing, Beate Kamlage, Melanie Hoehne, Richard Trethewey, and Mark Stitt. "Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in Arabidopsis rosettes." In: *Genome biology* 7.8 (2006), pp. R76.

[80] Pieter Giesbertz, Josef Ecker, Alexander Haag, Britta Spanier, and Hannelore Daniel. "An LC-MS/MS method to quantify acylcarnitine species including isomeric and odd-numbered forms in plasma and tissues." In: *Journal of lipid research* 56.10 (2015), pp. 2029–39.

[81] David Gomez-Cabrero, Imad Abugessaisa, Dieter Maier, Andrew Teschendorff, Matthias Merkenschlager, Andreas Gisel, Esteban Ballestar, Erik Bongcam-Rudloff, Ana Conesa, and Jesper Tegnér. "Data integration in the era of omics: current and future challenges." In: *BMC Systems Biology* 8.2 (2014), p. I1.

[82] Bradley S. Gordon, Michael L. Rossetti, and Alexey M. Eroshkin. "Arrdc2 and Arrdc3 Elicit Divergent Changes in Gene Expression in Skeletal Muscle following Anabolic and Catabolic Stimuli." In: *Physiological Genomics* (2019).

[83] Carla Grandori, Shaun M. Cowley, Leonard P. James, and Robert N. Eisenman. "The Myc/Max/Mad Network and the Transcriptional Control of Cell Behavior." In: *Annual Review of Cell and Developmental Biology* 16.1 (2000), pp. 653–99.

[84] Julian L. Griffin, Stephanie A. Bonney, Chris Mann, Abdul M. Hebbachi, Geoff F. Gibbons, Jeremy K. Nicholson, Carol C. Shoulders, and James Scott. "An integrated reverse functional genomic and metabolic approach to understanding orotic acid-induced fatty liver." In: *Physiological Genomics* 17.2 (2004), pp. 140–9.

[85] Ulrike Grömping. "Relative Importance for Linear Regression in R: The Package relaimpo." In: *Journal of Statistical Software* 17.1 (2006), pp. 1–27.

[86] Ulrike Grömping. "Estimators of Relative Importance in Linear Regression Based on Variance Decomposition." In: *The American Statistician* 61.2 (2007), pp. 139–47.

[87] Douglas J. Guarnieri, Catherine E. Brayton, Sarah M. Richards, Jaime Maldonado-Aviles, Joseph R. Trinko, Jessica Nelson, Jane R. Taylor, Shannon L. Gourley, and Ralph J. DiLeone. "Gene Profiling Reveals a Role for Stress Hormones in the Molecular and Behavioral Response to Food Restriction." In: *Biological Psychiatry* 71.4 (2012), pp. 358–65.

[88] Ruth Gutierrez-Aguilar, Dong-Hoon Kim, Stephen C. Woods, and Randy J. Seeley. "Expression of New Loci Associated With Obesity in Diet-Induced Obese Rats: From Genetics to Physiology." In: *Obesity* 20.2 (2012), pp. 306–12.

[89] Samy L. Habib, Simona Simone, Jeff J. Barnes, and Hanna E. Abboud. "Tuberin haploinsufficiency is associated with the loss of OGG1 in rat kidney tumors." In: *Molecular Cancer* 7.1 (2008).

[90] Tina M. Hahn, John F. Breininger, Denis G. Baskin, and Michael W. Schwartz. "Coexpression of Agrp and NPY in fasting-activated hypothalamic neurons." In: *Nature Neuroscience* 1 (1998), pp. 271–2.

[91] Mariko Hara-Chikuma, Eisei Sohara, Tatemitsu Rai, Masahito Ikawa, Masaru Okabe, Sei Sasaki, Shinichi Uchida, and A. S. Verkman. "Progressive Adipocyte Hypertrophy in Aquaporin-7-deficient Mice: Adipocyte glycerol permeability as a novel regulator of fat accumulation." In: *Journal of Biological Chemistry* 280.16 (2005), pp. 15493–6.

[92] Niloofar Hariri and Louise Thibault. "High-fat diet-induced obesity in animal models." In: *Nutrition Research Reviews* 23.2 (2010), pp. 270–99.

[93] F. E. Jr Harrell, with contributions from Charles Dupont, and many others. *Hmisc: Harrell Miscellaneous. R package version 4.1-1.* `https://CRAN.R-project.org/package=Hmisc`. 2018.

[94] Traver Hart, H. Kiyomi Komori, Sarah LaMere, Katie Podshivalova, and Daniel R. Salomon. "Finding the active genes in deep RNA-seq gene expression studies." In: *BMC genomics* 14 (2013), p. 778.

[95] Osamu Hashimoto, Kazunari Sekiyama, Tsuyoshi Matsuo, and Yoshihisa Hasegawa. "Implication of activin E in glucose metabolism: Transcriptional regulation of the inhibin/activin βE subunit gene in the liver." In: *Life Sciences* 85.13 (2009), pp. 534–40.

[96] Osamu Hashimoto et al. "Activin E Controls Energy Homeostasis in Both Brown and White Adipose Tissues as a Hepatokine." In: *Cell Reports* 25.5 (2018), pp. 1193–203.

[97] Yehudit Hasin, Marcus Seldin, and Aldons Lusis. "Multi-omics approaches to disease." In: *Genome Biology* 18.1 (2017), p. 83.

[98] David W. Haslam and W. Philip T. James. "Obesity." In: *The Lancet* 366.9492 (2005), pp. 1197–1209.

[99] Ilkka Heinonen, P. Rinne, S. T. Ruohonen, S. Ruohonen, M. Ahotupa, and E. Savontaus. "The effects of equal caloric high fat and western diet on metabolic syndrome, oxidative stress and vascular endothelial function in mice." In: *Acta Physiologica* 211.3 (2014), pp. 515–27.

[100] Hari Hendarto, Toyoshi Inoguchi, Yasutaka Maeda, Noriko Ikeda, Jing Zheng, Ryoko Takei, Hisashi Yokomizo, Eiichi Hirata, Noriyuki Sonoda, and Ryoichi Takayanagi. "GLP-1 analog liraglutide protects against oxidative stress and albuminuria in streptozotocin-induced diabetic rats via protein kinase A-mediated inhibition of renal NAD(P)H oxidases." In: *Metabolism* 61.10 (2012), pp. 1422–34.

[101] Fredrick E. Henry, Ken Sugino, Adam Tozer, Tiago Branco, and Scott M. Sternson. "Cell type-specific transcriptomics of hypothalamic energy-sensing neuron responses to weight-loss." In: *eLife* 4 (2015), pp. e09800.

[102] Toshiyuki Hibuse et al. "Aquaporin 7 deficiency is associated with development of obesity through activation of adipose glycerol kinase." In: *Proceedings of the National Academy of Sciences of the United States of America* 102.31 (2005), pp. 10993–8.

[103] A. Hinsberger and B. K. Sandhu. "Digestion and absorption." In: *Current Paediatrics* 14.7 (2004), pp. 605–11.

[104] Masami Yokota Hirai, Mitsuru Yano, Dayan B. Goodenowe, Shigehiko Kanaya, Tomoko Kimura, Motoko Awazuhara, Masanori Arita, Toru Fujiwara, and Kazuki Saito. "Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana." In: *Proceedings of the National Academy of Sciences of the United States of America* 101.27 (2004), pp. 10205–10.

[105] Matthew D. Hirschey et al. "SIRT3 regulates mitochondrial fatty-acid oxidation by reversible enzyme deacetylation." In: *Nature* 464 (2010), pp. 121EP–.

[106] Matthew D. Hirschey et al. "SIRT3 Deficiency and Mitochondrial Protein Hyperacetylation Accelerate the Development of the Metabolic Syndrome." In: *Molecular Cell* 44.2 (2011), pp. 177–90.

[107] Ernest Hodgson. "Chapter Fourteen - Toxins and Venoms." In: *Toxicology and Human Environments*. Vol. 112. Progress in Molecular Biology and Translational Science. Academic Press, 2012, pp. 373–415.

[108] M Hwang et al. "Epac2a-null mice exhibit obesity-prone nature more susceptible to leptin resistance." In: *International Journal Of Obesity* 41 (2016), p. 279.

[109] Illumina. *TruSeq Technology*. http://www.illumina.com/science/education/truseq.html. Accessed: August 2015.

[110] Fumitaka Inoue, Walter L. Eckalbar, Yi Wang, Karl K. Murphy, Navneet Matharu, Christian Vaisse, and Nadav Ahituv. "Genomic and epigenomic mapping of leptin-responsive neuronal populations involved in body weight regulation." In: *Nature Metabolism* 1.4 (2019), pp. 475–84.

[111] Broad Institute. *Picard*. http://broadinstitute.github.io/picard.

[112] Broad Institute. *Type 2 Diabetes Knowledge Portal.* http://www.type2diabetesgenetics.org/. Accessed 15th Nov 2018.

[113] Julia Jäger et al. "Hepatic transforming growth factor-β1 stimulated clone-22 D1 controls systemic cholesterol metabolism." In: *Molecular metabolism* 3.2 (2014), pp. 155–66.

[114] Alexandra Jauhiainen, Olle Nerman, George Michailidis, and Rebecka Jörnsten. "Transcriptional and metabolic data integration and modeling for identification of active pathways." In: *Biostatistics (Oxford, England)* 13.4 (2012), pp. 748–61.

[115] Joy X. Jiang et al. "Reduced nicotinamide adenine dinucleotide phosphate oxidase 2 plays a key role in stellate cell activation and liver fibrogenesis in vivo." In: *Gastroenterology* 139.4 (2010), pp. 1375–84.

[116] Tao Jiang, Zhuowei Wang, Gregory Proctor, Shevie Moskowitz, Scott E. Liebman, Thomas Rogers, M. Scott Lucia, Jinping Li, and Moshe Levi. "Diet-induced Obesity in C57BL/6J Mice Causes Increased Renal Lipid Accumulation and Glomerulosclerosis via a Sterol Regulatory Element-binding Protein-1c-dependent Pathway." In: *Journal of Biological Chemistry* 280.37 (2005), pp. 32317–25.

[117] Szymon Jozefczuk, Sebastian Klie, Gareth Catchpole, Jedrzej Szymanski, Alvaro Cuadros-Inostroza, Dirk Steinhauser, Joachim Selbig, and Lothar Willmitzer. "Metabolomic and transcriptomic stress response of Escherichia coli." In: *Molecular systems biology* 6 (2010), p. 364.

[118] Dhiraj G. Kabra et al. "Hypothalamic leptin action is mediated by histone deacetylase 5." In: *Nature Communications* 7 (2016), p. 10782.

[119] Pâmela A. Kakimoto and Alicia J. Kowaltowski. "Effects of high fat diets on rodent liver bioenergetics and oxidative imbalance." In: *Redox Biology* 8 (2016), pp. 216–25.

[120] Atanas Kamburov, Rachel Cavill, Timothy M. D. Ebbels, Ralf Herwig, and Hector C. Keun. "Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA." In: *Bioinformatics* 27.20 (2011), pp. 2917–8.

[121] Minoru Kanehisa and S. Goto. "KEGG: kyoto encyclopedia of genes and genomes." In: *Nucleic acids research* 28.1 (2000), pp. 27–30.

[122] Minoru Kanehisa, Miho Furumichi, Mao Tanabe, Yoko Sato, and Kanae Morishima. "KEGG: new perspectives on genomes, pathways, diseases and drugs." In: *Nucleic acids research* 45.D1 (2017), pp. D353–61.

[123] Minoru Kanehisa, Yoko Sato, Miho Furumichi, Kanae Morishima, and Mao Tanabe. "New approach for understanding genome variations in KEGG." In: *Nucleic acids research* 47.D1 (2019), pp. D590–5.

[124] Mitsunori Kayano, Seiya Imoto, Rui Yamaguchi, and Satoru Miyano. "Multi-omics Approach for Estimating Metabolic Networks Using Low-Order Partial Correlations." In: *Journal of Computational Biology* 20.8 (2013), pp. 571–82.

[125] Kathleen Kelly, Brent H. Cochran, Charles D. Stiles, and Philip Leder. "Cell-specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor." In: *Cell* 35.3 (1983), pp. 603–10.

[126] Hyun-Jin Kim et al. "Metabolomic Analysis of Livers and Serum from High-Fat Diet Induced Obese Mice." In: *Journal of Proteome Research* 10.2 (2011), pp. 722–31.

[127] Min-Seon Kim et al. "Role of hypothalamic Foxo1 in the regulation of food intake and energy homeostasis." In: *Nature Neuroscience* 9.7 (2006), pp. 901–6.

[128] Irina A. Kirpich, Luis S. Marsano, and Craig J. McClain. "Gut–liver axis, nutrition, and non-alcoholic fatty liver disease." In: *Clinical Biochemistry* 48.13 (2015), pp. 923–30.

[129] Gemma M. Kirwan, Erik Johansson, Robert Kleemann, Elwin R. Verheij, Åsa M. Wheelock, Susumu Goto, Johan Trygg, and Craig E. Wheelock. "Building Multivariate Systems Biology Models." In: *Analytical Chemistry* 84.16 (2012), pp. 7064–71.

[130] Glen E. Kisby et al. "The Cycad Genotoxin MAM Modulates Brain Cellular Pathways Involved in Neurodegenerative Disease and Cancer in a DNA Damage-Linked Manner." In: *PLOS ONE* 6.6 (2011), pp. e20911–.

[131] Tadahiro Kitamura, Yun Feng, Yukari Ido Kitamura, Streamson C. Chua, Allison W. Xu, Gregory S. Barsh, Luciano Rossetti, and Domenico Accili. "Forkhead protein FoxO1 mediates Agrp-dependent effects of leptin on food intake." In: *Nature Medicine* 12.5 (2006), pp. 534–40.

[132] Zachary A. Knight, Keith Tan, Kivanc Birsoy, Sarah Schmidt, Jennifer L. Garrison, Robert W. Wysocki, Ana Emiliano, Mats I. Ekstrand, and Jeffrey M. Friedman. "Molecular Profiling of Activated Neurons by Phosphorylated Ribosome Capture." In: *Cell* 151.5 (2012).

[133] Eric Koesema and Thomas Kodadek. "Global analysis of gene expression mediated by OX1 orexin receptor signaling in a hypothalamic cell line." In: *PLOS ONE* 12.11 (2017), pp. e0188082.

[134] Greg M. Kowalski, Steven Hamley, Ahrathy Selathurai, Joachim Kloehn, David P. De Souza, Sean O'Callaghan, Brunda Nijagal, Dedreia L. Tull, Malcolm J. McConville, and Clinton R. Bruce. "Reversing diet-induced metabolic dysregulation by diet switching leads to altered hepatic de novo lipogenesis and glycerolipid synthesis." In: *Scientific reports* 6 (2016), p. 27541.

[135] Jan Krumsiek, Jörg Bartel, and Fabian J. Theis. "Computational approaches for systems metabolomics." In: *Current Opinion in Biotechnology* 39 (2016), pp. 198–206.

[136] Motoko Kuba et al. "Absence of Elovl6 attenuates steatohepatitis but promotes gallstone formation in a lithogenic diet-fed Ldlr- / -mouse model." In: *Scientific Reports* 5 (2015), pp. 17604EP.

[137] David K. La and James A. Swenberg. "DNA adducts: biological markers of exposure and potential applications to risk assessment." In: *Mutation Research/Reviews in Genetic Toxicology* 365.1 (1996), pp. 129–46.

[138] Brian Y. H. Lam et al. "Heterogeneity of hypothalamic pro-opiomelanocortin-expressing neurons revealed by single-cell RNA sequencing." In: *Molecular Metabolism* 6.5 (2017), pp. 383–92.

[139] D. Enette Larson-Meyer et al. "Effect of 6-month calorie restriction and exercise on serum and liver lipids and markers of liver function." In: *Obesity (Silver Spring, Md.)* 16.6 (2008), pp. 1355–62.

[140] Jinmi Lee et al. "Exendin-4 improves steatohepatitis by increasing Sirt1 expression in high-fat diet-induced obese C57BL/6J mice." In: *PloS one* 7.2 (2012), pp. e31394.

[141] Jeffrey T. Leek and John D. Storey. "Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis." In: *PLOS Genetics* 3.9 (2007), pp. 1724–35.

[142] Jeffrey T. Leek, W. Evan Johnson, Hilary S. Parker, Elana J. Fertig, Andrew E. Jaffe, John D. Storey, Yuqing Zhang, and Leonardo Collado Torres. *sva: Surrogate Variable Analysis*. R package version 3.30.1. 2019.

[143] Heng Li, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. "The Sequence Alignment/Map format and SAMtools." In: *Bioinformatics* 25.16 (2009), pp. 2078–9.

[144] Shuzhao Li, Andrei Todor, and Ruiyan Luo. "Blood transcriptomics and metabolomics for personalized medicine." In: *Computational and structural biotechnology journal* 14 (2015), pp. 1–7.

[145] Yong Li, Michael N. Corradetti, Ken Inoki, and Kun-Liang Guan. "TSC2: filling the GAP in the mTOR signaling pathway." In: *Trends in Biochemical Sciences* 29.1 (2004), pp. 32–8.

[146] Qingqing Liu, Bingbing Yuan, Kinyui Alice Lo, Heide C. Patterson, Yutong Sun, and Harvey F. Lodish. "Adiponectin regulates expression of hepatic genes critical for glucose and lipid metabolism." In: *Proceedings of the National Academy of Sciences of the United States of America* 109.36 (2012), pp. 14568–73.

[147] Zhuoying Liu, Ting Xiao, Xiaoyu Peng, Guangdi Li, and Fang Hu. "APPLs: More than just adiponectin receptor binding proteins." In: *Cellular Signalling* 32 (2017), pp. 76–84.

[148] Michael I. Love, Wolfgang Huber, and Simon Anders. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." In: *Genome Biology* 15 (12 2014), p. 550.

[149] Michael I. Love, Simon Anders, Vladislav Kim, and Wolfgang Huber. *RNA-seq workflow: gene-level exploratory analysis and differential expression*. `http://master.bioconductor.org/packages/release/workflows/vignettes/rnaseqGene/inst/doc/rnaseqGene.html`. 2018.

[150] Jiahong Lu et al. "NRBF2 regulates autophagy and prevents liver injury by modulating Atg14L-linked phosphatidylinositol-3 kinase III activity." In: *Nature Communications* 5 (2014), pp. 3920EP–.

[151] Xi Ma et al. "MTORC1-mediated NRBF2 phosphorylation functions as a switch for the class III PtdIns3K and autophagy." In: *Autophagy* 13.3 (2017), pp. 592–607.

[152] Norikazu Maeda, Tohru Funahashi, Toshiyuki Hibuse, Azumi Nagasawa, Ken Kishida, Hiroshi Kuriyama, Tadashi Nakamura, Shinji Kihara, Iichiro Shimomura, and Yuji Matsuzawa. "Adaptation to fasting by glycerol transport through aquaporin 7 in adipose tissue." In: *Proceedings of the National Academy of Sciences of the United States of America* 101.51 (2004), pp. 17801–6.

[153] Tobias Maier, Marc Güell, and Luis Serrano. "Correlation of mRNA and protein in complex biological samples." In: *FEBS Letters* 583.24 (2009), pp. 3966–73.

[154] Claudia Manzoni, Demis A. Kia, Jana Vandrovcova, John Hardy, Nicholas W. Wood, Patrick A. Lewis, and Raffaele Ferrari. "Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences." In: *Briefings in Bioinformatics* 19.2 (2016), pp. 286–302.

[155] Santiago Marco-Sola, Michael Sammeth, Roderic Guigo, and Paolo Ribeca. "The GEM mapper: fast, accurate and versatile alignment by filtration." In: *Nat Meth* 9.12 (2012), pp. 1185–8.

[156] D. R. Matthews, J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner. "Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man." In: *Diabetologia* 28.7 (1985), pp. 412–9.

[157] Colin S. McCoin, Trina A. Knotts, Kikumi D. Ono-Moore, Pieter J. Oort, and Sean H. Adams. "Long-chain acylcarnitines activate cell stress and myokine release in C2C12 myotubes: calcium-dependent and -independent effects." In: *American Journal of Physiology-Endocrinology and Metabolism* 308.11 (2015), pp. E990–1000.

[158] Ruth McPherson. "Genetic contributors to obesity." In: *The Canadian journal of cardiology* 23.Suppl A (2007), pp. 23A–27A.

[159] Kevin S. Mcdorman and Douglas C. Wolf. "Use of the Spontaneous Tsc2 Knockout (Eker) Rat Model of Hereditary Renal Cell Carcinoma for the Study of Renal Carcinogens." In: *Toxicologic Pathology* 30.6 (2002), pp. 675–80.

[160] Rebecca E. Mercer, Melissa J. S. Chee, and William F. Colmers. "The role of NPY in hypothalamic mediated food intake." In: *Frontiers in Neuroendocrinology* 32.4 (2011), pp. 398–415.

[161] Michele Milella and Alessandra Felici. "Biology of metastatic renal cell carcinoma." In: *Journal of Cancer* 2 (2011), pp. 369–73.

[162] Nabora Soledad Reyes de Mochel, Scott Seronello, Shelley Hsiuying Wang, Chieri Ito, Jasper Xi Zheng, T. Jake Liang, J. David Lambeth, and Jinah Choi. "Hepatocyte NAD(P)H oxidases as an endogenous source of reactive oxygen species during hepatitis C virus infection." In: *Hepatology (Baltimore, Md.)* 52.1 (2010), pp. 47–59.

[163] Masahiro Morita and Osamu Hashimoto. "Identification and expression of the medaka inhibin βE subunit." In: *Molecular Biology Reports* 46.2 (2019), pp. 1603–9.

[164] Olena Morozova, Martin Hirst, and Marco A. Marra. "Applications of New Sequencing Technologies for Transcriptome Analysis." In: *Annual Review of Genomics and Human Genetics* 10.1 (2009), pp. 135–51.

[165] Mark R. Morris and Farida Latif. "The epigenetic landscape of renal cancer." In: *Nature Reviews Nephrology* 13 (2016), pp. 47–60.

[166] Ali Mortazavi, Brian A. Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. "Mapping and quantifying mammalian transcriptomes by RNA-Seq." In: *Nature Methods* 5 (2008), pp. 621–8.

[167] Timo D. Müller et al. "Restoration of leptin responsiveness in diet-induced obese mice using an optimized leptin analog in combination with exendin-4 or FGF21." In: *J Pept Sci* 18 (2012), pp. 383–393.

[168] Angela A. Mulligan, Gunter G.C. Kuhnle, Kay-Tee Khaw, Marleen A.H. Lentjes, Robert N. Luben, and Heather A. Ward. "Breast, colorectal, and prostate cancer risk in the European Prospective Investigation into Cancer and Nutrition–Norfolk in relation to phytoestrogen intake derived from an improved database." In: *The American Journal of Clinical Nutrition* 91.2 (2009), pp. 440–8.

[169] Martin G. Myers Jr and David P. Olson. "Central nervous system control of metabolism." In: *Nature* 491 (2012), pp. 357–63.

[170] Miki Nagase and Toshiro Fujita. "Role of Rac1 – mineralocorticoid-receptor signalling in renal and cardiac disease." In: *Nature Reviews Nephrology* 9 (2013), pp. 86EP–.

[171] Christopher B. Newgard et al. "A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance." In: *Cell metabolism* 9.4 (2009), pp. 311–26.

[172] Jens Nielsen. "Systems Biology of Metabolism." In: *Annual Review of Biochemistry* 86.1 (2017), pp. 245–75.

[173] Victoria J. Nikiforova, Carsten O. Daub, Holger Hesse, Lothar Willmitzer, and Rainer Hoefgen. "Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response." In: *Journal of Experimental Botany* 56.417 (2005), pp. 1887–96.

[174] Takehiko Nohmi. "Thresholds of Genotoxic and Non-Genotoxic Carcinogens." In: *Toxicological research* 34.4 (2018), pp. 281–90.

[175] Katsunori Nonogaki, Kana Ohashi-Nozue, and Yoshitomo Oka. "Induction of hypothalamic serum- and glucocorticoid-induced protein kinase-1 gene expression and its relation to plasma des-acyl ghrelin in energy homeostasis in mice." In: *Biochemical and Biophysical Research Communications* 344.2 (2006), pp. 696–9.

[176] Erin P. O'Keefe. "siRNAs and shRNAs: Tools for Protein Knockdown by Gene Silencing." In: *MATER METHODS* (2013), p. 197.

[177] Marco Peluso et al. "DNA Adducts and Lung Cancer Risk: A Prospective Study." In: *Cancer Research* 65.17 (2005), pp. 8042–48.

[178] David N. Perkins, Darryl J. C. Pappin, David M. Creasy, and John S. Cottrell. "Probability-based protein identification by searching sequence databases using mass spectrometry data." In: *ELECTROPHORESIS* 20.18 (1999), pp. 3551–67.

[179] Belinda Phipson, Stanley Lee, Ian J. Majewski, Warren S. Alexander, and Gordon K. Smyth. "Robust hyperparameter estimation protects against hypervariable genes and improves power to detect differential expression." In: *The annals of applied statistics* 10.2 (2016).

[180] Massimo Pinzani and Krista Rombouts. "Liver fibrosis: from the bench to clinical targets." In: *Digestive and Liver Disease* 36.4 (2004), pp. 231–42.

[181] Alja Plošnik, Marjan Vračko, and Marija Sollner Dolenc. "Mutagenic and carcinogenic structural alerts and their mechanisms of action." In: *Archives of Industrial Hygiene and Toxicology* 67.3 (2016), pp. 169–82.

[182] Leona Plum et al. "The obesity susceptibility gene Cpe links FoxO1 signaling in hypothalamic pro-opiomelanocortin neurons with regulation of food intake." In: *Nature Medicine* 15 (2009), pp. 1195–201.

[183] Michal M. Poplawski, Xue-Jun Yang, Charles V. Mobbs, and Jason W. Mastaitis. "Hypothalamic Responses to Fasting Indicate Metabolic Reprogramming Away from Glycolysis Toward Lipid Oxidation." In: *Endocrinology* 151.11 (2010), pp. 5206–17.

[184] Susan Preston-Martin, Malcolm C. Pike, Ronald K. Ross, Peter A. Jones, and Brian E. Henderson. "Increased Cell Division as a Cause of Human Cancer." In: *Cancer Research* 50.23 (1990), pp. 7415–21.

[185] Aijuan Qu, Changtao Jiang, Yan Cai, Jung-Hwan Kim, Naoki Tanaka, Jerrold M. Ward, Yatrik M. Shah, and Frank J. Gonzalez. "Role of Myc in hepatocellular proliferation and hepatocarcinogenesis." In: *Journal of hepatology* 60.2 (2014), pp. 331–38.

[186] *RSeQC: An RNA-seq Quality Control Package*. `https://sourceforge.net/projects/rseqc/files/BED/Mouse_Mus_musculus/`. Accessed: 2015-06-30.

[187] Hongxia Ren, Ian J. Orozco, Ya Su, Shigetomo Suyama, Roger Gutiérrez-Juárez, Tamas L. Horvath, Sharon L. Wardlaw, Leona Plum, Ottavio Arancio, and Domenico Accili. "FoxO1 Target Gpr17 Activates AgRP Neurons to Regulate Food Intake." In: *Cell* 149.6 (2012), pp. 1314–26.

[188] Hongxia Ren, Joshua R. Cook, Ning Kon, and Domenico Accili. "Gpr17 in AgRP Neurons Regulates Feeding and Sensitivity to Insulin and Leptin." In: *Diabetes* 64.11 (2015), pp. 3670–9.

[189] Helen J. Renaud, Julia Y. Cui, Hong Lu, and Curtis D. Klaassen. "Effect of diet on expression of genes involved in lipid metabolism, oxidative stress, and inflammation in mouse liver – insights into mechanisms of hepatic steatosis." In: *PloS one* 9.2 (2014), pp. e88584.

[190] Kris Richardson et al. "The PLIN4 variant rs8887 modulates obesity related phenotypes in humans through creation of a novel miR-522 seed site." In: *PloS one* 6.4 (2011), pp. e17944.

[191] Matthew E. Ritchie, Belinda Phipson, Di Wu, Yifang Hu, Charity W. Law, Wei Shi, and Gordon K. Smyth. "limma powers differential expression analyses for RNA-sequencing and microarray studies." In: *Nucleic Acids Research* 43.7 (2015), pp. e47.

[192] Efren Riu, Fatima Bosch, and A. Valera. "Prevention of diabetic alterations in transgenic mice overexpressing Myc in the liver." In: *Proceedings of the National Academy of Sciences of the United States of America* 93.5 (1996), pp. 2198–202.

[193] Efren Riu, Tura Ferre, Alex Mas, Antonio Hidalgo, Sylvie Franckhauser, and Fatima Bosch. "Overexpression of c-myc in diabetic mice restores altered expression of the transcription factor genes that regulate liver metabolism." In: *The Biochemical journal* 368.Pt 3 (2002), pp. 931–7.

[194] Efren Riu, Tura Ferre, Alex Mas, Antonio Hidalgo, Sylvie Franckhauser, and Fatima Bosch. "Overexpression of c-myc in diabetic mice restores altered expression of the transcription factor genes that regulate liver metabolism." In: *Biochemical Journal* 368.3 (2002), pp. 931–7.

[195] Chantal Rodgarkia-Dara, Susanne Vejda, Natascha Erlach, Annemarie Losert, Wilfried Bursch, Walter Berger, Rolf Schulte-Hermann, and Michael Grusch. "The activin axis in liver biology and disease." In: *Mutation Research/Reviews in Mutation Research* 613.2 (2006), pp. 123–37.

[196] Alexander Roth et al. "STRING v10: protein–protein interaction networks, integrated over the tree of life." In: *Nucleic Acids Research* 43.D1 (2014), pp. D447–52.

[197] Maxwell A. Ruby, Julie Massart, Devon M. Hunerdosse, Milena Schönke, Jorge C. Correia, Sharon M. Louie, Jorge L. Ruas, Erik Näslund, Daniel K. Nomura, and Juleen R. Zierath. "Human Carboxylesterase 2 Reverses Obesity-Induced Diacylglyc-

erol Accumulation and Glucose Intolerance." In: *Cell reports* 18.3 (2017), pp. 636–46.

[198] Liangyou Rui. "Energy metabolism in the liver." In: *Comprehensive Physiology* 4.1 (2014), pp. 177–97.

[199] Jennifer M. Rutkowsky, Trina A. Knotts, Kikumi D. Ono-Moore, Colin S. McCoin, Shurong Huang, Dina Schneider, Shamsher Singh, Sean H. Adams, and Daniel H. Hwang. "Acylcarnitines activate proinflammatory signaling pathways." In: *American Journal of Physiology-Endocrinology and Metabolism* 306.12 (2014), pp. E1378–87.

[200] Jiyoon Ryu. "APPL2 Promotes Inflammation in Mouse Liver by Negative Regulation of Adiponectin Signaling." In: *Diabetes* 67.Supplement 1 (2018), pp. 1982–P.

[201] Kim S. *ppcor: Partial and Semi-Partial (Part) Correlation. R package version 1.1.* `https://CRAN.R-project.org/package=ppcor`. 2015.

[202] Kazuki Saito, Masami Y. Hirai, and Keiko Yonekura-Sakakibara. "Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'." In: *Trends in Plant Science* 13.1 (2008), pp. 36–43.

[203] Kazuki Saito and Fumio Matsuda. "Metabolomics for Functional Genomics, Systems Biology, and Biotechnology." In: *Annual Review of Plant Biology* 61.1 (2010), pp. 463–89.

[204] Juliane Schäfer and Korbinian Strimmer. "Learning Large-Scale Graphical Gaussian Models from Genomic Data." In: *AIP Conference Proceedings* 776.1 (2005), pp. 263–76.

[205] Domenico Sergi et al. "SerpinA3N is a novel hypothalamic gene upregulated by a high-fat diet and leptin in mice." In: *Genes & nutrition* 13 (2018), p. 28.

[206] Richard A. Shapiro, Donald G. Skinner, Philip Stanley, and Harold H. Edelbrock. "Renal tumors associated with tuberous sclerosis: the case for aggressive surgical management." In: *The Journal of urology* (1984).

[207] Majken Siersbæk, Lyuba Varticovski, Shutong Yang, Songjoon Baek, Ronni Nielsen, Susanne Mandrup, Gordon L. Hager, Jay H. Chung, and Lars Grøntved. "High fat diet-induced changes of mouse hepatic transcription and enhancer activity can be reversed by subsequent weight loss." In: *Scientific Reports* 7 (2017), pp. 40220EP.

[208] Rajat Singh, Susmita Kaushik, Yongjun Wang, Youqing Xiang, Inna Novak, Masaaki Komatsu, Keiji Tanaka, Ana Maria Cuervo, and Mark J. Czaja. "Autophagy regulates lipid metabolism." In: *Nature* 458.7242 (2009), pp. 1131–5.

[209] Stephanie Sisley and Darleen Sandoval. "Hypothalamic control of energy and glucose metabolism." In: *Reviews in endocrine & metabolic disorders* 12 (2011), pp. 219–33.

[210] Gordon K. Smyth. "Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments." In: *Statistical Applications in Genetics and Molecular Biology* 1.3 (2004), pp. 1–25.

[211] Gordon K. Smyth, Hamish S. Scott, and Joëlle Michaud. "Use of within-array replicate spots for assessing differential expression in microarray experiments." In: *Bioinformatics* 21.9 (2005), pp. 2067–75.

[212] Ock Soon Sohn, Emerich S. Fiala, Silvia P. Requeijo, John H. Weisburger, and Frank J. Gonzalez. "Differential Effects of CYP2E1 Status on the Metabolic Activation of the Colon Carcinogens Azoxymethane and Methylazoxymethanol." In: *Cancer Research* 61.23 (2001), p. 8435.

[213] Minsoo Song. "Recent developments in small molecule therapies for renal cell carcinoma." In: *European Journal of Medicinal Chemistry* 142 (2017). Current advances in cancer research: Therapeutics, Targets, and Chemical Biology, pp. 383–92.

[214] J. Speakman, C. Hambly, S. Mitchell, and E. Król. "Animal models of obesity." In: *Obesity Reviews* 8.s1 (2007), pp. 55–61.

[215] Peter S. Spencer, C. Edwin Garner, Valerie S. Palmer, Glen E. Kisby, Michael Aschner, and Lucio G. Costa. "Chapter 11 - Environmental Neurotoxins Linked to a Prototypical Neurodegenerative Disease." In: Boston: Academic Press, 2015, pp. 211–52.

[216] Ganesh Sriram, Lilly S. Parr, Lola Rahib, James C. Liao, and Katrina M. Dipple. "Moonlighting function of glycerol kinase causes systems-level changes in rat hepatoma cells." In: *Metabolic engineering* 12.4 (2010), pp. 332–40.

[217] James H. Steiger. "Tests for comparing elements of a correlation matrix." In: *Psychological Bulletin* 87 (1980), pp. 245–51.

[218] Kerstin Stemmer, Heidrun Ellinger-Ziegelbauer, Hans-Juergen Ahr, and Daniel R. Dietrich. "Carcinogen-Specific Gene Expression Profiles in Short-term Treated Eker and Wild-type Rats Indicative of Pathways Involved in Renal Tumorigenesis." In: *Cancer Research* 67.9 (2007), pp. 4052–68.

[219] Kerstin Stemmer, Heidrun Ellinger-Ziegelbauer, Hans-Juergen Ahr, and Daniel R. Dietrich. "Molecular characterization of preneoplastic lesions provides insight on the development of renal tumors." In: *The American journal of pathology* 175.4 (2009), pp. 1686–98.

[220] John D. Storey. "A direct approach to false discovery rates." In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.3 (2002), pp. 479–98.

[221] Gang Su, Charles F. Burant, Christopher W. Beecher, Brian D. Athey, and Fan Meng. "Integrated metabolome and transcriptome analysis of the NCI60 dataset." In: *BMC bioinformatics* 12 Suppl 1.Suppl 1 (2011), pp. S36.

[222] Kenji Sugawara, Tadao Shibasaki, Harumi Takahashi, and Susumu Seino. "Structure and functional roles of Epac2 (Rapgef4)." In: *Gene* 575.2, Part 3 (2016), pp. 577–83.

[223] Masakazu Sugiyama et al. "Inhibin βE (INHBE) is a possible insulin resistance-associated hepatokine identified by comprehensive gene expression analysis in human liver biopsy samples." In: *PLOS ONE* 13.3 (2018), pp. e0194798.

[224] Bandar Suliman, Dakang Xu, and Bryan Williams. "The Promyelocytic Leukemia Zinc Finger Protein: Two Decades of Molecular Oncology." In: *Frontiers in Oncology* 2 (2012), p. 74.

[225] Marc Sultan et al. "A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome." In: *Science* 321.5891 (2008), pp. 956–60.

[226] William R. Swindell. "Comparative analysis of microarray data identifies common responses to caloric restriction among mouse tissues." In: *Mechanisms of Ageing and Development* 129.3 (2008), pp. 138–53.

[227] R Core Team. "R: A Language and Environment for Statistical Computing." In: *R Foundation for Statistical Computing, Vienna, Austria* (2013). URL: `http://www.R-project.org`.

[228] Unnur P. Thorgeirsson, Dan W. Dalgard, Jeanette Reeves, and Richard H. Adamson. "Tumor Incidence in a Chemical Carcinogenesis Study of Nonhuman Primates." In: *Regulatory Toxicology and Pharmacology* 19.2 (1994), pp. 130–51.

[229] Dina G. Tiniakos, João Maurício, and Helen L. Reeves. "Fatty Liver Disease and Hepatocellular Carcinoma: The Pathologist's View." In: *Alcohol and Cancer*. Ed. by Vasilis Vasiliou, Samir Zakhari, Lopa Mishra, and Helmut K. Seitz. Cham: Springer International Publishing, 2018, pp. 55–69.

[230] Kelsey L. Tinkum, Lynn S. White, Luciano Marpegan, Erik Herzog, David Piwnica-Worms, and Helen Piwnica-Worms. "Forkhead Box O1 (FOXO1) Protein, but Not p53, Contributes to Robust Induction of p21 Expression in Fasted Mice." In: *Journal of Biological Chemistry* 288.39 (2013), pp. 27999–8008.

[231] Scott Tonidandel and James M. LeBreton. "Relative Importance Analysis: A Useful Supplement to Regression Analysis." In: *Journal of Business and Psychology* 26.1 (2011), pp. 1–9.

[232] Elijah Trefts, Maureen Gannon, and David H. Wasserman. "The liver." In: *Current biology : CB* 27.21 (2017), R1147–51.

[233] Takashi Ueno and Masaaki Komatsu. "Autophagy in the liver: functions in health and disease." In: *Nature Reviews Gastroenterology &Amp; Hepatology* 14 (2017), pp. 170EP–.

[234] Ewa Urbanczyk-Wochniak, Alexander Luedemann, Joachim Kopka, Joachim Selbig, Ute Roessner-Tunali, Lothar Willmitzer, and Alisdair R. Fernie. "Parallel analysis of transcript and metabolic profiles: a new approach in systems biology." In: *EMBO reports* 4.10 (2003), pp. 989–93.

[235] A. Valera, A. Pujol, X. Gregori, Efren Riu, J. Visa, and Fatima Bosch. "Evidence from transgenic mice that myc regulates hepatic glycolysis." In: *The FASEB Journal* 9.11 (1995), pp. 1067–78.

[236] Paul Van Hummelen and Jennifer Sasaki. "State-of-the-art genomics approaches in toxicology." In: *Mutation Research/Reviews in Mutation Research* 705.3 (2010), pp. 165–71.

[237] V Vauthier, T D. Swartz, P Chen, C Roujeau, M Pagnon, J Mallet, C Sarkis, R Jockers, and J Dam. "Endospanin 1 silencing in the hypothalamic arcuate nucleus contributes to sustained weight loss of high fat diet obese mice." In: *Gene Therapy* 21 (2014), pp. 638–44.

[238] Naomi van Vlies, Liqun Tian, Henk Overmars, Albert H. Bootsma, Willem Kulik, Ronald J. A. Wanders, Philip A. Wood, and Frédéric M. Vaz. "Characterization of carnitine and fatty acid metabolism in the long-chain acyl-CoA dehydrogenase-deficient mouse." In: *Biochemical Journal* 387.1 (2005), p. 185.

[239] WHO. *Obesity and overweight - Key facts*. `https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight`. Accessed: May 2019.

[240] C Walker, TL Goldsworthy, DC Wolf, and J Everitt. "Predisposition to renal cell carcinoma due to alteration of a cancer susceptibility gene." In: *Science* 255.5052 (1992), pp. 1693–95.

[241] Christopher S. Walker et al. "Mice Lacking the Neuropeptide α-Calcitonin Gene-Related Peptide Are Protected Against Diet-Induced Obesity." In: *Endocrinology* 151.9 (2010), pp. 4257–69.

[242] Guo-En Wang, Yi-Fang Li, Yu-Jia Zhai, Lian Gong, Jing-Yu Tian, Mo Hong, Nan Yao, Yan-Ping Wu, Hiroshi Kurihara, and Rong-Rong He. "Theacrine protects against nonalcoholic fatty liver disease by regulating acylcarnitine metabolism." In: *Metabolism - Clinical and Experimental* 85 (2018), pp. 227–39.

[243] Hong Wang, Yun Zhao, J. Alyce Bradbury, Joan P. Graves, Julie Foley, Joyce A. Blaisdell, Joyce A. Goldstein, and Darryl C. Zeldin. "Cloning, Expression, and Characterization of Three New Mouse Cytochrome P450 Enzymes and Partial Characterization of Their Fatty Acid Oxidation Activities." In: *Molecular Pharmacology* 65.5 (2004), pp. 1148–58.

[244] Liguo Wang, Shengqin Wang, and Wei Li. "RSeQC: quality control of RNA-seq experiments." In: *Bioinformatics* 28.16 (2012), pp. 2184–5.

[245] Xin Wang, Rui Wang, Miao Luo, Chen Li, Hua-Xia Wang, Chang-Chao Huan, Yu-Rong Qu, Ying Liao, and Xiang Mao. "(DEAD)-box RNA helicase 3 modulates NF-κB signal pathway by controlling the phosphorylation of PP2A-C subunit." In: *Oncotarget* 8.20 (2017), pp. 33197–213.

[246] Y. Wang et al. "Exendin-4 decreases liver inflammation and atherosclerosis development simultaneously by reducing macrophage infiltration." In: *British journal of pharmacology* 171.3 (2014), pp. 723–34.

[247] Zhong Wang, Mark Gerstein, and Michael Snyder. "RNA-Seq: a revolutionary tool for transcriptomics." In: *Nature Reviews Genetics* 10 (2009), pp. 57EP–.

[248] Kwanjeera Wanichthanarak, Johannes F. Fahrmann, and Dmitry Grapov. "Genomic, Proteomic, and Metabolomic Data Integration Strategies." In: *Biomarker Insights* 10s4 (2015), pp. BMI.S29511.

[249] E. Wiater and W. Vale. "Activins and Inhibins." In: *The TGF-β Family*. Derynck R. and Miyazono K. Eds., New York: Cold Spring Harbor Lab. Press, 2008, pp. 79–120.

[250] Frank Wilcoxon. "Individual Comparisons by Ranking Methods." In: *Biometrics Bulletin* 1.6 (1945), pp. 80–3.

[251] Gareth Williams, Chen Bing, Xue J. Cai, Joanne A. Harrold, Peter J. King, and Xiao H. Liu. "The hypothalamus and the control of energy homeostasis: Different circuits, different purposes." In: *Physiology & Behavior* 74.4 (2001), pp. 683–701.

[252] Pu Wu, Cecilia Jiang, Qian Shen, and Yinghe Hu. "Systematic gene expression profile of hypothalamus in calorie-restricted mice implicates the involvement of mTOR signaling in neuroprotective activity." In: *Mechanisms of Ageing and Development* 130.9 (2009), pp. 602–10.

[253] Jianguo Xia, Christopher D. Fjell, Matthew L. Mayer, Olga M. Pena, David S. Wishart, and Robert E. W. Hancock. "INMEX–a web-based tool for integrative meta-analysis of expression data." In: *Nucleic acids research* 41.Web Server issue (2013), pp. W63–70.

[254] Takakazu Yagi, Hirotaka Ueda, Haruka Amitani, Akihiro Asakawa, Shouichi Miyawaki, and Akio Inui. "The Role of Ghrelin, Salivary Secretions, and Dental Care in Eating Disorders." In: *Nutrients* 4.8 (2012), pp. 967–89.

[255] Ronghua Yang and Lili A. Barouch. "Leptin Signaling and Obesity." In: *Circulation Research* 101.6 (2007), pp. 545–59.

[256] Jin Yoo, In-Jin Cho, In-Kyung Jeong, Kyu Jeung Ahn, Ho Yeon Chung, and You-Cheol Hwang. "Exendin-4, a glucagon-like peptide-1 receptor agonist, reduces hepatic steatosis and endoplasmic reticulum stress by inducing nuclear factor erythroid-derived 2-related factor 2 nuclear translocation." In: *Toxicology and Applied Pharmacology* 360 (2018), pp. 18–29.

[257] Sooyeon Yoo, David Cha, Dong Won Kim, Thanh V. Hoang, and Seth Blackshaw. "Tanycyte-Independent Control of Hypothalamic Leptin Signaling." In: *Frontiers in Neuroscience* 13 (2019), p. 240.

[258] Guangchuang Yu, Li-Gen Wang, Yanyan Han, and Qing-Yu He. "clusterProfiler: an R package for comparing biological themes among gene clusters." In: *Omics : a journal of integrative biology* 16.5 (2012), pp. 284–87.

[259] Feng Yue et al. "A comparative encyclopedia of DNA elements in the mouse genome." In: *Nature* 515.7527 (2014), pp. 355–64.

[260] M. S. Zedeck and G. B. Brown. "Methylation of intestinal and hepatic DNA in rats treated with methylazoxymethanol acetate." In: *Cancer* 40.5 Suppl (1977), pp. 2580–83.

[261] Irene Sui Lan Zeng and Thomas Lumley. "Review of Statistical Learning Methods in Integrated Omics Studies (An Integrated Information Science)." In: *Bioinformatics and biology insights* 12 (2018), p. 1177932218759292.

[262] Suiliang Zhang, Liang Sui, Juhua Zhuang, Saifei He, Yanan Song, Ying Ye, and Wei Xia. "ARHGAP24 regulates cell ability and apoptosis of colorectal cancer cells via the regulation of P53." In: *Oncology letters* 16.3 (2018), pp. 3517–24.

[263] Ming Zhao, Lopa Mishra, and Chu-Xia Deng. "The role of TGF-β/SMAD4 signaling in cancer." In: *Int J Biol Sci* 14.2 (2018), pp. 111–23.

[264] Yu Zhong, Deanna H. Morris, Lin Jin, Mittul S. Patel, Senthil K. Karunakaran, You-Jun Fu, Emily A. Matuszak, Heidi L. Weiss, Brian T. Chait, and Qing Jun Wang. "Nrbf2 protein suppresses autophagy by modulating Atg14L protein-containing Beclin 1-Vps34 complex architecture and reducing intracellular phosphatidylinositol-3 phosphate levels." In: *The Journal of biological chemistry* 289.38 (2014), pp. 26021–37.

[265] Yiming Zuo, Guoqiang Yu, Mahlet G. Tadesse, and Habtom W. Ressom. "Biological network inference using low order partial correlation." In: *Methods (San Diego, Calif.)* 69.3 (2014), pp. 266–73.

[266] Tunahan Çakir, Kiran R. Patil, Zeynep I. Önsan, Kutlu Ö. Ülgen, Betül Kirdar, and Jens Nielsen. "Integration of metabolome data with metabolic networks reveals reporter reactions." In: *Molecular systems biology* 2 (2006), p. 50.

[267] Cell signaling technology. *Phospho-S6 Ribosomal Protein (Ser240/244) Antibody #2215*. `https://www.cellsignal.com/products/primary-antibodies/phospho-s6-ribosomal-protein-ser240-244-antibody/2215?hit=productId&Ntt=2215`.

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die bei der Fakultät für Medizin der TUM zur Promotionsprüfung vorgelegte Arbeit mit dem Titel **RNA Sequencing and data integration in the context of nutrition and weight loss intervention** am Helmholtz Zentrum München, Institut für Diabetes und Adipositas unter der Anleitung und Betreuung durch Prof. Dr. Paul Pfluger ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 6 und 7 Satz 2 angegebenen Hilfsmittel benutzt habe.

[X] Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

[X] Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

[ ] Die vollständige Dissertation wurde in _____ veröffentlicht. Die promotionsführende Einrichtung _____ hat der Vorveröffentlichung zugestimmt.

[X] Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

[ ] Ich habe bereits am _____ bei der Fakultät für _____ der Hochschule _____ unter Vorlage einer Dissertation mit dem Thema _____ die Zulassung zur Promotion beantragt mit dem Ergebnis: _____

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.

Mit der Aufnahme meiner personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich [X] einverstanden, [ ] nicht einverstanden.

München, den _____ _____

Valentina Klaus