A panomics step to unravel the role of gut microbiota in metabolic disease

Meijnikman, A.S.

Publication date
2023
Document Version
Final published version

A panomics step *to unravel* the role of

GUT MICROBIOTA in METABOLIC DISEASE

Abraham Stijn Meijnikman

A panomics step to unravel the role of gut microbiota in metabolic disease

Abraham Stijn Meijnikman

Colofon

A panomics step to unravel the role of gut microbiota in metabolic disease
Dissertation, University of Amsterdam, Amsterdam, The Netherlands

ISBN: 978-94-92332-53-0
Author: Abraham Stijn Meijnikman
Cover: Esther Scheide, Proefschriftomslag.nl
Design: Esther Scheide, Proefschriftomslag.nl
Printing: Proefschriftmaken

Printing of this thesis was financially supported by Medische Specialistisch Bedrijf Gouda, Stichting tot Steun Promovendi Vasculaire Geneeskunde, Spaarne Gasthuis Academy.

A panomics step to unravel the role of gut microbiota in metabolic disease

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op woensdag 19 april 2023, te 16.00 uur

door Abraham Stijn Meijnikman
geboren te Alkmaar

Promotiecommissie

Promotores:	prof. dr. M. Nieuwdorp	AMC-UvA
	prof. dr. A.K. Groen	AMC-UvA
Copromotores:	dr. H.J. Herrema	AMC-UvA
	dr. V.E.A. Gerdes	AMC-UvA
Overige leden:	prof. dr. F. Kuipers	Rijksuniversiteit Groningen
	prof. dr. P.C.N. Rensen	Universiteit Leiden
	prof. dr. U.H.W. Beuers	AMC-UvA
	prof. dr. W.J. de Jonge	AMC-UvA
	prof. dr. J.J.G.H.M. Bergman	AMC-UvA
	prof. dr. R.H.L. Houtkooper	AMC-UvA

Faculteit der Geneeskunde

Voor Hugo en Louise

TABLE OF CONTENTS

1

General introduction and outline of this thesis

GENERAL INTRODUCTION

The rates of overweight and obesity have reached pandemic proportions and has major consequences for public health and the economies of today's world. Estimates suggest that 300 million individuals are affected by obesity and almost half of the global population has overweight[1]. The global obesity pandemic is a driver for obesity-related complications such as insulin resistance, type 2 diabetes (T2D), cardiovascular disease and nonalcoholic fatty liver disease (NAFLD)[1]. In fact, the prevalence of these obesity-related complications is proportional to the increase in body mass index (BMI)[2]. In 2040, the predicted global prevalence of T2D in adults is a staggering 10%. For NAFLD, the prevalence at that time is predicted to be approximately 25-30% with the vast majority of cases (>80%) identified in individuals with obesity[3]. The ongoing rise in prevalence of these diseases and lack of effective treatment options, presents an unprecedented challenge for the public health[4,5]. Clearly, there is a high need to deeply investigate the complex pathophysiology of obesity, insulin resistance and NAFLD to identify new and effective treatment options.

The central cause of overweight and obesity can be brought down to an imbalance between the number of calories consumed and the number of calories used for energy expenditure by the human body[2]. This simple thermodynamic equation, however, is subject to and driven by many external regulatory factors. These have thus far been proven challenging to integrate in an all-encompassing approach to unravel the pathogenesis of obesity and -related disorders. Despite decades of dedicated research and billions of euros spent annually, there are only a few registered therapeutic treatment options available for obesity and its complications with unfortunately limited efficicacy.

Currently, numerous biomedical studies aim to assess how (epi)genetic, molecular, behavioral, developmental, and environmental forces affect the intake and the processing of these excessive calories[2]. Nevertheless, for most of its existence, the focus of biomedical research has been on the ability to first identify and second to target specific diseases-associated signaling pathways. Although this approach has generated new and exciting therapeutic strategies that target specific pathways, it does not account for the interindividual variability in disease development and response to therapy, which is particularly relevant for multifactorial diseases. Moreover, this approach cannot provide an overview of multiple pathways at the same time, neither is it possible to identify which pathways are proximal and which are dominant in disease development. Nevertheless, in search of potential new treatment options for obesity and related-complications, the scientific community still hangs on to this one-size-fits all approach by performing large randomized controlled trials without accounting for sex, age, ethnicity, and other confounders. To successfully tackle the obesity pandemic and stop the seemingly never-ending rise in prevalence of obesity related

diseases, it is critical to move beyond this general approach and focus on the interindividual differences.

It is now increasingly recognized that every individual, potentially responds different to an overflood of calories suggesting that there are also differences in metabolic processes in each individual[6]. This starts in the gut, where a collection of bacteria, archaea and eukarya -collectively termed the gut microbiome- have colonized the gastrointestinal tract[7]. Microorganisms are among the oldest living organisms on Earth originating over 3700 million years ago[8]. Microorganisms, including the large number residing in the gastro-intestinal tract, have a profound impact on physiology and the ability to influence general health and disease states[9,10]. With the introduction of affordable high-throughput sequencing, interest in and identification of the role of the gut microbiota in modulation of host metabolism has grown exponentially[11]. Profound differences in gut microbial composition between disease stages were identified and by employing a translational approach, causal evidence of gut microbiota in human metabolism was obtained for several diseases ranging from metabolic to psychiatric disorders[12]. Hence, maybe Hippocrates was right when he said: "all disease begins in the gut". It is now crystal clear that the gut microbiota influences numerous physiological processes including digestion, absorption, metabolism, and immune system development and function[11,13]. Yet, the impact and importance of the gut microbiota in the development of metabolic diseases is still under debate. As of to date, it has been shown that a few interventions targeting the gut microbiota suggest beneficial effects on clinical manifestations of metabolic disease whereas other interventions did not yield any meaningful outcome[11,14,15]. However, it is possible that the large interindividual variation in response to gut microbiota targeted interventions could be attributed to the heterogeneity of the disease itself, combined with the variation of the patients baseline microbial composition, as shown in several studies[16]. These interindividual differences open up the prospect of precision medicine, yet it also highlights the complexity of metabolic disease development. In fact, the complexity of contributing factors can obscure various structural associations between metabolic activities in different tissues, preventing deeper insight into molecular mechanisms of disease development[17].

We need studies that characterize patients in depth before and after interventions, where clinical data is combined with high-dimensional data such as fecal metagenomics, metabolomics, host tissue transcriptomics with the overarching aim to identify patients' signatures that allow us to predict response to a specific intervention[18]. Moreover, by employing a systems biology approach using these panomics data, it is possible to identify the crosstalk between different relevant biological layers identifying the hierarchy of these biological layers.

In this thesis, we took a global approach to investigate factors that may contribute to the development of obesity, insulin resistance and NAFLD in humans. We used a machine learning approach to integrate transcriptomics, fecal metagenomics and plasma metabolomics datasets from individuals that underwent bariatric surgery. In addition, we used the steppingstones that were provided by the comprehensive panomics analyses to move from association to causation and performed two clinical trials targeting the gut microbiota. Last, we investigated an old but new player in metabolic disease development, cellular senescence and found that insulin is highly associated with markers of senescence in both the liver as in the adipocytes.

OUTLINE OF THE THESIS

Part 1. Systems biology in metabolic disease

This part starts with a review assessing the quality of the evidence that is supportive of a causal role for the gut microbiome in obesity and diabetes development in humans (**Chapter 2**). Challenging aspects in determining causality in humans are postulated together with strategies that might hold potential to indeed assess a driving role for gut microbiota in metabolic disease development. Furthermore, we discuss means to modify gut microbiome composition in humans to help establish causality and discuss systems biology approaches that might hold the key to unravel the role of gut microbiota in obesity and metabolic diseases. In **Chapter 3,** we describe the distinct differences in gut microbial composition and function between individuals with and without obesity. Moreover, we show that the gut microbiome can explain the variance in several clinical phenotypes associated with obesity. In **Chapter 4** we describe the design and the aims of the BARIA study. Using this longitudinal bariatric surgery cohort, we aim to perform a systems biology approach identifying gut microbial, immunological and metabolic markers in a large and well phenotyped bariatric surgery cohort to identify novel pathways in the pathogenesis of obesity, T2D and NAFLD. The next three chapters used both clinical and omics data of participants from the BARIA study. In **Chapter 5**, we describe that the heterogeneity of a bariatric surgery population can be stratified better phenotypically using metabotyping, i.e., monitoring the fasting plasma metabolome, that captures the functional output of a complex multi-organ system, human host and their microbes rather than traditional clinical classifiers. Using state of the art machine learning techniques, we identified five distinct metabotypes, which were differentially enriched for metabolic pathways related to immune function, fatty acid metabolism, protein-signaling, and obesity pathogenesis. Stratification into metabotypes identified specific biomarker signatures for each phenotype from the multi-omics data. In

Chapter 6, we performed plasma metabolomic profiling on both fasting and postprandial samples and investigated global metabolic responses to a mixed meal test. We identified abnormal metabolism of (branched chain) amino acids, fatty acids and acylcholines in individuals with (pre)diabetes. Furthermore, we traced the differences in metabolic responses back to other omics sets including fecal metagenomics and transcriptomics data of liver, adipose tissue, and jejunum. Finally, by applying integrative machine learning models, we identified possible new biomarkers for glycemic control including N–acetylaspartate, phenylalanine derived metabolites and butyrate producing bacterial species that may be useful for intervention and prevention of T2D. In **Chapter 7,** we took a global approach to investigate factors that may contribute to NAFLD development in humans and used a machine learning approach to integrate transcriptomics, fecal metagenomics and plasma metabolomics datasets from obese women with and without NAFLD. Analyses of these integrated omics sets revealed a robust NAFLD-signature and highlight the additive value of a multi-omics approach to study NAFLD pathophysiology.

Part 2. From association to causation

In this part we used the steppingstones provided by the systems biology approach of part 1, to assess the causal role of the gut microbiome in the pathogenesis of NAFLD. In **Chapter 8**, we reveal by performing four distinct experiments in humans that the gut microbiome of individuals with NAFLD produces incredible amounts of ethanol that are clinically revelant for the development and progression of NAFLD. In **Chapter 9**, we performed a double-blind randomized controlled trial in which individuals with NAFLD received either an allogenic or autologous fecal microbiota transplantation (FMT). We observed that an allogenic FMT using lean vegan donors in individuals with NAFLD induced changes in intestinal microbiome composition, which was associated with beneficial changes in plasma metabolites and markers of NAFLD.

Part 3. Cellular senescence, an old but new player in metabolic disease

Cellular senescence is a state of irreversible cell cycle arrest with important physiological functions. However, cellular senescence is also a hallmark of ageing and has been associated with several pathological conditions. In the past decade, cellular senescence gained significant interest due to its putative role in the development of NAFLD and the progression towards NASH. Until recently, it was suggested that hepatocyte cellular senescence is a mere consequence of the metabolic dysregulation and inflammatory phenomena in fatty liver disease. However, recent work in rodents has suggested that senescence may be a causal factor in NAFLD development. In **Chapter 10**, we aim to provide insight in the quality of the evidence supportive of a causal role of cellular senescence in the development of

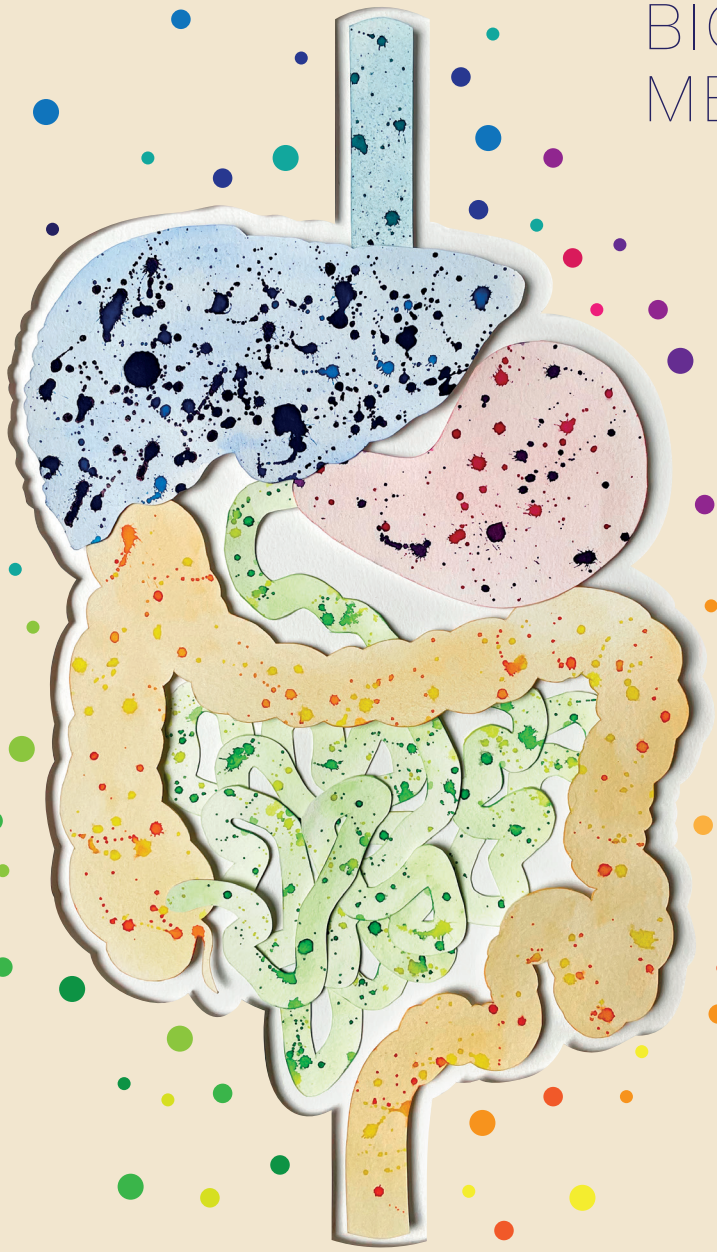
NAFLD in rodents and humans. In addition, we elaborate on key cellular and molecular features of senescence and discuss the efficacy and safety of novel senolytic drugs to treat or even prevent NAFLD. In **Chapter 11**, we show by using clinical data, portal vein plasma, immunohistochemistry, and transcriptomics data derived from individuals with obesity to establish a link between insulin and senescence in the liver. We found strong correlations between markers of cellular senescence and insulin, independent of NAFLD suggesting that hepatocytic senescence, driven by prevailing insulin concentrations, precedes NAFLD. These observations were validated using an external validation cohort. In **Chapter 12,** we used the same approach as in the previous chapter by using clinical, immunohistochemistry, and transcriptomic data derived from non-diabetic individuals with obesity, we have shown that peripheral insulin resistance is highly correlated with markers of cellular senescence in mesenteric adipose tissue prior to the presence of glycaemic dysregulation. These results confirm previous studies indicating that adipocytic cellular senescence may play an important role already in the earliest stages of insulin resistance before the onset of T2DM.

REFERENCES

1. Ng, M. *et al.* Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* (2014) doi:10.1016/S0140-6736(14)60460-8.
2. Schwartz, M. W. *et al.* Obesity Pathogenesis: An Endocrine Society Scientific Statement. *Endocr Rev* **38**, 267–296 (2017).
3. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
4. Marjot, T., Moolla, A., Cobbold, J. F., Hodson, L. & Tomlinson, J. W. Nonalcoholic Fatty Liver Disease in Adults: Current Concepts in Etiology, Outcomes, and Management. *Endocr Rev* **41**, 66–117 (2020).
5. Sanyal, A. J. Past, present and future perspectives in nonalcoholic fatty liver disease. *Nature Reviews Gastroenterology & Hepatology 2019 16:6* **16**, 377–386 (2019).
6. Apovian, C. M. *et al.* Approaches to addressing the rise in obesity levels. *Nature Reviews Endocrinology 2022* 1–6 (2022) doi:10.1038/s41574-022-00777-1.
7. Blaser, M. J. The Past and Future Biology of the Human Microbiome in an Age of Extinctions. *Cell* **172**, 1173–1177 (2018).
8. Dodd, M. S. *et al.* Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature 2017 543:7643* **543**, 60–64 (2017).
9. Young, V. B. The role of the microbiome in human health and disease: An introduction for clinicians. *BMJ (Online)* vol. 356 Preprint at https://doi.org/10.1136/bmj.j831 (2017).
10. Neish, A. S. Microbes in Gastrointestinal Health and Disease. *Gastroenterology* vol. 136 Preprint at https://doi.org/10.1053/j.gastro.2008.10.080 (2009).
11. Meijnikman, A. S., Gerdes, V. E., Nieuwdorp, M. & Herrema, H. Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. *Endocr Rev* **39**, 133–153 (2018).
12. Cani, P. D. Gut microbiota — at the intersection of everything? *Nature Reviews Gastroenterology & Hepatology 2017 14:6* **14**, 321–322 (2017).
13. Aron-Wisnewsky, J. *et al.* Gut microbiota and human NAFLD: disentangling microbial signatures from metabolic disorders. *Nat Rev Gastroenterol Hepatol* **17**, 279–297 (2020).
14. Dominguez-Bello, M. G., Godoy-Vitorino, F., Knight, R. & Blaser, M. J. Role of the microbiome in human development. *Gut* **68**, (2019).
15. Wortelboer, K. *et al.* From fecal microbiota transplantation toward next-generation beneficial microbes: The case of Anaerobutyricum soehngenii. *Front Med (Lausanne)* **9**, 3619 (2022).
16. Olofsson, L. E. & Bäckhed, F. The Metabolic Role and Therapeutic Potential of the Microbiome. *Endocr Rev* **43**, 907–926 (2022).
17. Meijnikman, A. S. *et al.* A systems biology approach to study non-alcoholic fatty liver (NAFL) in women with obesity. *iScience* **25**, 104828 (2022).
18. van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *J Intern Med* joim.13157 (2020) doi:10.1111/joim.13157.

PART 1

SYSTEMS BIOLOGY IN METABOLIC DISEASE



2

Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans

Abraham S. Meijnikman, Victor A.E. Gerdes, Max Nieuwdorp, Hilde Hererma

Endocrine Reviews, 2018 Apr 1;39(2):133-153

ABSTRACT

The pathophysiology of obesity and obesity-related diseases such as type 2 diabetes mellitus (T2DM) is complex and driven by many factors. One of the most recently identified factors in development of these metabolic pathologies is the gut microbiota. The introduction of affordable, high-throughput sequencing technologies has significantly expanded our understanding of the role of the gut microbiome in modulation of host metabolism and (cardio)metabolic disease development. Nevertheless, evidence for a role of the gut microbiome as causal, driving factor in disease development mainly originates from studies in mouse models: data showing causality in humans is scarce. In this review, we will discuss the quality of evidence supporting a causal role for the gut microbiome in the development of obesity and diabetes, in particular T2DM, in humans. Considering overlap in potential mechanisms, the role of the gut microbiome in type 1 diabetes mellitus will also be addressed. We will elaborate on factors that drive microbiome composition in humans and discuss how alterations in microbial composition or microbial metabolite production contribute to disease development. Challenging aspects in determining causality in humans will be postulated together with strategies that might hold potential to overcome these challenges. Furthermore, we will discuss means to modify gut microbiome composition in humans to help establish causality and discuss systems biology approaches that might hold the key to unravel the role of gut microbiome in obesity and T2DM.

INTRODUCTION

The global rise in prevalence of obesity presents an unprecedented challenge to public health and economies of today's world. Obesity has been associated with a plethora of metabolic disturbances including dyslipidaemia and insulin resistance; both are considered major risk factors for development of cardiovascular disease (CVD), non-alcoholic fatty liver disease (NAFLD) and several forms of cancer. Obesity therefore is considered one of the greatest public health threats of the 21st century[1]. Factors that strongly contribute to the obesity epidemic include decreased physical activity and increased (high-caloric) food intake. However, if the pathogenesis of obesity would have been this simple, Hippocrates' prescription for treatment of obesity: "eat only once a day and take no baths and sleep on a hard bed and walk naked as long as possible" would have been a successful prescription[2]. Unfortunately, treatment (and prevention) of obesity and obesity-related complications have been proven to be more complex. Despite extensive efforts in the field, successful strategies to tackle this pathology are still limited. The need to mechanistically unravel development of obesity and –related disease is therefore high and crucial for development of novel, effective treatment strategies.

The rise in prevalence of obesity coincides with the prevalence of type 2 diabetes mellitus (T2DM), which is a leading cause of CVD in almost all high-income countries[3]. It has been estimated that by the year 2040, a staggering 642 million people will suffer from this disease worldwide[3]. Numerous researchers have dedicated their careers to unravelling pathophysiological pathways that underlie the development of T2DM in obesity. In 2009, DeFronzo introduced a then-new paradigm in diabetes research: the ominous octet[4]. This paradigm describes that in addition to muscle, liver and beta-cells (triumvirate[5]), adipocytes, the gastrointestinal tract, alpha-cells, kidney and brain all play important roles in the development of T2DM. The dogma also describes the complexity of development of T2DM: numerous determinants drive disease development; however, the hierarchy of these driving factors remains largely unknown. Additionally, determinants other than those described in the 'ominous octet' might play a role in the development of T2DM. In the past decade, the gut microbiome has been identified as a novel, potentially driving, factor in the pathophysiology of T2DM.

In addition to T2DM, the incidence of type 1 diabetes mellitus (T1DM) is rapidly increasing worldwide as well[6,7]. Genetic predisposition or children being born from genetically susceptible mothers cannot simply explain this phenomenon[8]. The disproportionate increase in T1DM incidence has therefore largely been attributed to environmental influences such as early enterovirus infection[9]. In addition, the clinical onset of T1DM is usually preceded by years of enhanced systemic inflammation and augmented autoimmunity that associate with

shifts in gut microbial composition[10]. The gut microbiome has therefore been put forward as driving force in pathogenesis of T1DM.

Interest in and identification of the role of the gut microbiome in modulation of host metabolism have grown exponentially since the introduction of affordable, high-throughput sequencing technologies. These technologies allowed for compositional as well as functional analysis of intestinal microbiota in humans and mouse models. Murine models have provided crucial insight in determinants of gut microbiome composition and the role of the gut microbiome in health and disease. Although studies performed in murine models support the hypothesis that the gut microbiome might play a causal role in development of obesity and diabetes, data showing causality in humans is still scarce.

In this review, we aim to provide insight in (the quality of) evidence that is supportive of a causal role for the gut microbiome in obesity and diabetes development in humans. We will elaborate on factors that drive microbiome composition in humans and discuss possible mechanisms through which the gut microbiome and microbial metabolites affect host metabolism. Challenging aspects in determining causality in humans will be postulated together with strategies that might hold potential to indeed assess a driving role for gut microbiota in metabolic disease development. Furthermore, we will discuss means to modify gut microbiome composition in humans to help establish causality and discuss systems biology approaches that might hold the key to unravel the role of gut microbiota in obesity and diabetes.

Factors shaping the gut microbiota

The human gut microbiota is a complex ecosystem consisting of an estimated 10^{14} bacteria[11]. This number equals the number of human cells[12]. The combined genetic material of the gut microbiota, collectively called the gut microbiome, exceeds the human genome approximately 100 times[13,14,15]. The gut microbial community is dominated by five bacterial phyla; Actinobacteria, Bacteroidetes, Firmicutes, Proteobacteria and Verrucomicrobia. Variable pH (pH increases from proximal to distal intestine) and oxygen concentration (decreases from proximal to distal intestine) affect both relative and absolute abundance of bacteria across the gastrointestinal (GI) tract. The proximal GI tract is enriched in bacteria belonging to the phyla Firmicutes and Proteobacteria and the genus Lactobacilli, whereas the distal GI tract mainly comprises bacteria belonging to the phyla Bacteroidetes, Firmicutes and the Akkermansia municiphilia species[16].

Although a definition of what a healthy gut microbiome comprises still has to be defined, it is clear that in healthy individuals, the composition of the intestinal microbiota is highly diverse[17]. Interestingly, together with increased industrialization, an overall decline in gut microbiota diversity can been observed[18]. This decline likely is a consequence of modern

lifestyle and driven by introduction of new medication and increased availability of processed foods. Importantly, gut microbial composition is highly variable between individuals and is continuously modified by both endogenous and exogenous factors. This interindividual variability already starts at birth and is mainly determined by the microbiota composition of the mother[19]. Interestingly, the intestine of a new-born is not sterile; it has been suggested that intrauterine exposure to the mothers' microbiome is one of the first shaping factors of the gut microbiome[20]. The intestine is further colonized by bacteria as soon as the amniotic fluid disappears[21] and predominantly determent by the mode of delivery. Children born through natural (vaginal) delivery have a gut microbiota composition resembling the vaginal microbiota composition of the mother. Children born through Caesarean section on the other hand, have a gut microbiota composition that resembles the skin microbiota composition of the mothers[22]. These first determinants of gut microbiota composition persist for months, potentially even longer[23]. Whether differences in gut microbiota this early in life affect disease development later in life remains to be determined.

A recent study that combined whole genome sequencing (WGS) with 16S rRNA sequencing showed that there are significant interindividual differences in gut microbial diversity and richness depending on age and ethnicity of the host[24]. In addition to interindividual changes in microbiota composition, functional analysis of the gut microbiota of children and adults indicated age-related differences in the abundance of genes involved in amino acid metabolism, lipopolysaccharide (LPS) biosynthesis, RNA degradation and steroid hormone biosynthesis[25]. Although these results indicate that gut microbiota output differs between subjects in an age-dependent manner, these results have to be interpreted with caution. WGS is indicative only of potential functions but does not assess the actual gene expression levels. Rather, these results indicate that age and ethnicity are associated with differences in functionality of the gut microbiome.

One of the most important modulators of the gut microbiota is diet. Intervention studies in humans have revealed the extent to which the microbiota can be modulated by dietary changes[26,27]. The influence of diet on gut microbiome composition and functionality can be described in three different themes[28]. *First*, the response of the gut microbiota to (major) changes in dietary composition is very fast. Several studies have shown an acute shift in gut microbial composition and functionality as soon as two days after the start of a dietary intervention[26,27]. Switching from plant- and meat-based diets to a diet with a daily add-on of 30 grams of dietary fibres induced both compositional and functional changes in the gut microbiota[26]. In addition, compositional and functional changes were observed in subjects who followed either a high–fibre, low-fat or low–fibre, high-fat diet for ten days[26,27]. *Second*, despite rapid changes in composition and function following (major) changes in dietary composition, long-term dietary habits are required to induce major changes in

gut microbiota composition. This is most clearly exemplified by observations that certain microbial taxa found in traditional populations that stick to a plants-only diet (unique abundance of bacteria from the genus Prevotella and Xylanibacter) are absent in Western populations[29]). Furthermore, several studies have shown acute effects of diet on microbiome composition soon after the start of a dietary intervention but failed to show major changes at later time points. One study, for example, reported that dietary intervention for ten days did not mediate major compositional changes in gut microbiota composition whereas changes were detectable 24 hours after the start of the dietary intervention[27]. *Third*, there is high interindividual variability in response of microbiome composition to changes in dietary composition[30,31,32]. The fact that dietary interventions to treat obesity have variable effects could therefore potentially be due to differences in microbiota composition at the start of the diet[33]. Increased intake of fibres and decreased total caloric intake have been shown to increase microbial diversity in subjects with low microbial gene richness, at baseline. In contrast, subjects with high microbial gene richness at baseline remain unaffected by this dietary interventio[31].

Medication also significantly influences the gut microbial composition. Antibiotics treatment is well-known for influencing the gut microbiota[34,35]. Moreover, antibiotics use early in life has been associated with weight gain later in life[36]. A recent study showed that oral antibiotic treatment leads to specific expansion of Firmicutes[37], which might have unfavourable effects since an increased abundance of Firmicutes has been associated with obesity[38] and T2DM[39]. A single blinded randomized controlled trial in 20 male obese subjects who received either vancomycin or amoxicillin for seven days showed that vancomycin-treated subjects had significantly decreased peripheral insulin sensitivity compared to amoxicillin-treated subjects[40]. Vancomycin treatment, which specifically eradicates Gram+ bacteria, shifts the gut microbial community to a community dominated by Gram- bacteria which might negatively affect host metabolism, including insulin sensitivity. In another study, however, these metabolic effects of vancomycin treatment were not observed[41]. In this randomized double-blind, placebo-controlled trial, 57 obese human subjects were treated (oral) with vancomycin, amoxicillin, or placebo for seven days[41]. Amoxicillin treatment did not significantly affect microbiota composition whereas treatment with vancomycin had major impact on microbial diversity and composition with a decrease of Gram+ bacteria and a compensatory increase in Gram- bacteria. Although this was accompanied by changes in microbiota-mediated metabolic processes (*i.e.,* reduced conversion of primary to secondary bile acids and reduced production of short chain fatty acids), insulin sensitivity, energy metabolism and systemic low-grade inflammation were unaffected[41]. The discrepancy in metabolic outcome in these two studies is potentially due to differences in fecal bacterial richness at baseline. In a recent study from our group, we show that microbiota

composition at baseline is indeed an accurate determinant of the effect of fecal microbiota transplantation (FMT) on insulin sensitivity. We were able to predict with high accuracy if obese, insulin resistant recipients should be categorized as responders or non-responders following an allogenic FMT from a lean donor[42]. The fact that FMT-mediated improvements in metabolism do not sustain in recipients with high bacterial diversity at baseline (prior to FMT) can be attributed to the presence of a highly personal core microbiome[43]. In this subset of recipients, intestinal microbiota composition following FMT quickly returns to baseline composition, potentially driven by resilience of the host immune system[44]. It cannot be excluded, however, that lifestyle (particularly dietary habits) of these subjects plays a large part in this phenomenon. Nevertheless, higher bacterial diversity is likely accompanied by a more pronounced personal core microbiome composition which is difficult to change with FMT or antibiotics treatment. Therefore, microbiota-mediated effects on metabolism following FMT or antibiotic treatment are more challenging and less sustainable in subjects with high bacterial diversity

Metformin is currently the most prescribed oral antidiabetic medication and known to affect intestinal microbiota composition[45,46]. In a recent double-blind, placebo-controlled trial in patients with T2DM it was indeed shown that metformin-treated subjects had significantly altered gut microbiome composition compared to patients receiving placebo[47]. Interestingly, germ-free (GF) mice that received an FMT from metformin-treated subjects had improved glucose tolerance compared to mice that received an FMT from placebo-treated controls suggesting that metformin-induced changes in gut microbiome composition mediate part of the beneficial effects of this drug on glucose homeostasis[47]. It has been suggested that the beneficial effects of metformin are, at least in part, mediated by the production of short chain fatty acids (SCFA's) by the gut microbiota[46]. Functional shifts in LPS biosynthesisand SCFA metabolism in patients treated with metformin were observed[46]. Interestingly, known adverse events of metformin such as diarrhoea, nausea, vomiting and bloating were associated with a relative increase in abundance of Escherichia species[46]. In a recent study assessing the gut metagenome in faecal samples of 748 human subjects with and without T2DM, it was shown that metformin is a strong confounding factor in metagenomic analysis[46]. Thus, when assessing microbiota composition in T2DM subjects, it is of critical importance to correct for metformin use.

Proton pump inhibitors (PPI) are frequently used oral anti-acid medication that have also been implicated to modulate gut microbiota composition[48]. Although fecal microbial diversity did not change significantly, certain taxa known to have high potential to overgrow (*e.g.*, Clostridium difficile), were increased after four weeks of omeprazole (40 milligrams/day) treatment. These results suggest that PPI treatment might predispose to Clostridium difficile infection[48].

The role of human genetics in shaping the composition of the gut microbiota remains largely associative. A recent study in monozygotic twins suggested heritability of a number of microbial species[49], in part based on the association between the human gene locus that encodes lactase and the Bifidobacterium genus. Other associations between human genetic make-up and microbiome composition were found in genome-wide association studies in which genetic loci, microbial taxa and functional pathways were linked[50,51,52]. Recently, a novel, computational method applied on cross-sectional datasets from two large metagenomic studies, was used to investigate regulatory factors driving individual microbial composition[53]. Interestingly, it was suggested that gut microbial composition, at least at the species level, was independent of host genetics[53]. This conclusion challenges the assumption that along with host genetics and host immunity has a smaller role in shaping the gut microbiome than was previously considered. If true, this conclusion will have major influence on the development of successful generic procedures and products to manipulate microbiota composition[54].

Although the hierarchy of factors that drive gut microbial composition remains largely unknown, it is evident that a complex interplay between ethnicity, host genetics, mode of delivery, dietary habits and (history of) medication use all play an important role in shaping the microbial community. We will discuss in the next paragraphs available evidence that implicate a role for the gut microbiota in development of metabolic diseases such as obesity and diabetes in humans.

Gut microbiome composition and T2DM development in humans: why it is challenging to determine causality in humans

Interest in the role of the gut microbiota in development of metabolic disturbances such as obesity and T2DM in humans has risen significantly over the past decade. This is in part due to the introduction of novel and more affordable next-generation sequencing (NGS) techniques combined with increased availability of fecal samples and tissue biopsies obtained from human subjects. Nevertheless, studies reporting a link between gut microbiome composition and metabolic disease development in humans are still largely associative/correlative in nature and mostly based on the differences in relative abundance of bacterial strains in the accessible fecal compartment. Furthermore, reproducibility of results from studies in humans has been shown to be fairly low[55], which challenges a one-directional interpretation of the role of the gut microbiota in disease development. Discrepancies in study set up, geographical location of sample preparation and inconsistencies in data analysis all play part in low reproducibility. Full transparency of study details including open access to methodology and raw data in online repositories and collaborative initiatives between research groups (*e.g.*, exchange of samples and analysis on different sites) will enhance the reproducibility of data in the field.

Scarcity of biopsies or luminal/mucosal material from proximal parts of the intestine leaves the microbial composition and function of this important part of the GI tract relatively unexplored. In addition, gut microbiota composition has been mainly linked to clinical parameters obtained from observational (retrospective) studies. Often times, it cannot be concluded if gut microbiota composition was affected prior to disease development (causal) or whether the microbiota composition is a reflection (consequence) of the disease itself. This chicken-egg situation can in part be clarified in large, *prospective* studies such as the Dutch Life Lines[56] and HELIUS[57] cohorts. Although prospective studies will provide insight in the timeline of disease development linked to changes in gut microbiota composition, a causal contribution (*i.e.*, microbiota as driving factor for disease development) can only be concluded from intervention studies. However, controlled intervention studies with significant effect on microbiota composition in humans are rare and have thus far been limited to FMT, antibiotic treatment, diet and probiotic therapy. Although FMT in particular holds potential to serve as efficient intervention strategy to study causality in humans[58], other intervention studies in humans have thus far shown limited causal evidence for a role of the gut microbiota in metabolic disease development[41]. A top-down approach to determine a causal role of the gut microbiome in the development of (cardio)metabolic disease is presented in figure 1.

Causality: insight from studies in mice

Causal evidence that link the intestinal microbiota to host health and development of metabolic disease mostly originates from rodent studies[59]. The GI tract of humans and mice are anatomically, genetically and physiologically quite similar. Composition of sectional tissue of small and large intestine from mice resembles sections from humans. In addition, Goblet and Paneth cells fulfil the same unique role in intestinal integrity and host-microbiota equilibrium in both humans and mice[60]. Nevertheless, important differences exist and therefore, care must be taken to draw direct parallels between mice and human studies. An important difference between human and mouse GI tract is that the mouse cecum is relatively large in comparison with the size of the total GI tract. Moreover, the cecum is an important site for fermentation in mice. Increased fermentation capacity in mice significantly affects gut microbial diversity, composition and functionality[60]. In contrast, the human cecum, is relatively small and does not have a clear function[61]. Genetic background is one of the main drivers of the metabolic phenotype in mice whereas in humans, obesity and insulin resistance are driven by a complex interaction of genetics, diet and lifestyle[4]. Furthermore, and in sharp contrast to most human studies, mouse studies can be strictly controlled to minimize confounding factors that often times complicate data interpretation in humans (*e.g.*, food intake, dietary composition, history of medication use). The ability to

genetically modify mice provides valuable mechanistic insight in how the gut microbiota affects host metabolism and augments metabolic disease development.

Studies in germ free (GF) mice, which lack microbiota, provided first important evidence that the gut microbiota potentially plays a causal role in development of obesity and –related diseases. It was demonstrated that, despite a higher food intake, GF mice are leaner compared to conventionally raised mice[62]. In addition, GF mice are fairly resistant to HFD induced obesity[63]. GF mice allow for generation of gnotobiotic models: GF mice colonized with a specific microbe of interest or harbouring a strictly defined microbial community. FMT using a fecal transplant from conventionally raised mice, increased body fat by 60% and reduced food intake in GF recipients[62]. The gut microbiota thus increases the ability to derive energy from food (particularly from indigestible carbohydrates) thereby fueling energy metabolism of the host. A follow up study, in which fecal microbiota was transplanted from conventionally raised obese mice to GF mice, further accelerated establishment of a causal role for the gut microbiota in development of obesity[38]. Interestingly, GF recipient mice that received a transplant from an obese donor gained more weight on the same diet compared to recipients that received a transplant form a lean donor. These results suggested that the microbiome of obese mice harvest more energy from dietary components. Additionally, these data implicated that an obese phenotype can be transferred from donor to recipient indicating causality. In line, a study where fecal microbiota from twins discordant for obesity was transplanted to GF mice, showed that recipients of the faecal microbial transplant from the obese co-twin gained significantly more weight gain compared to counterparts that received a transplant from the lean co-twin[64]. Although studies in GF mice have provided crucial insight in the contribution of the gut microbiota to host metabolism, there are substantial differences in metabolism of germ free versus conventionally raised mice[60]. For example, GF mice have the tendency to consume more calories, excrete more lipids and weigh less than conventionally raised mice[65]. Importantly, lack of microbiota has significant consequences for maturation and capacity of the immune system and intestinal physiology[66]. Since immune system and intestinal function are crucial players in development of (microbiome-mediated) obesity and T2DM, results obtained from GF mice should be interpreted with some restraint. Despite convincing evidence from studies in mice, data implicating a causal role for the gut microbiota in obesity development in this model system cannot be projected on humans. To exemplify difficulties in interpreting mouse and human data and to underscore the challenges of translational research approaches, a recent study in mice reported that a membrane protein of the mucin-degrading bacterium Akkermansia mucinophilia improved obesity and T2DM[67]. This was conflicting with a study in humans where both Akkermansia mucinophilia and Desulfovirbrio were enriched in samples of T2DM patients[68] thus underscoring the question whether decreased relative abundance

of specific strains is a driving factor or merely a reflection of the disease. It is therefore relevant to ask where we currently stand in our understanding and evidence of the role of the gut microbiome in cardiometabolic disease development in human disease and look into strategies to tackle these challenges.

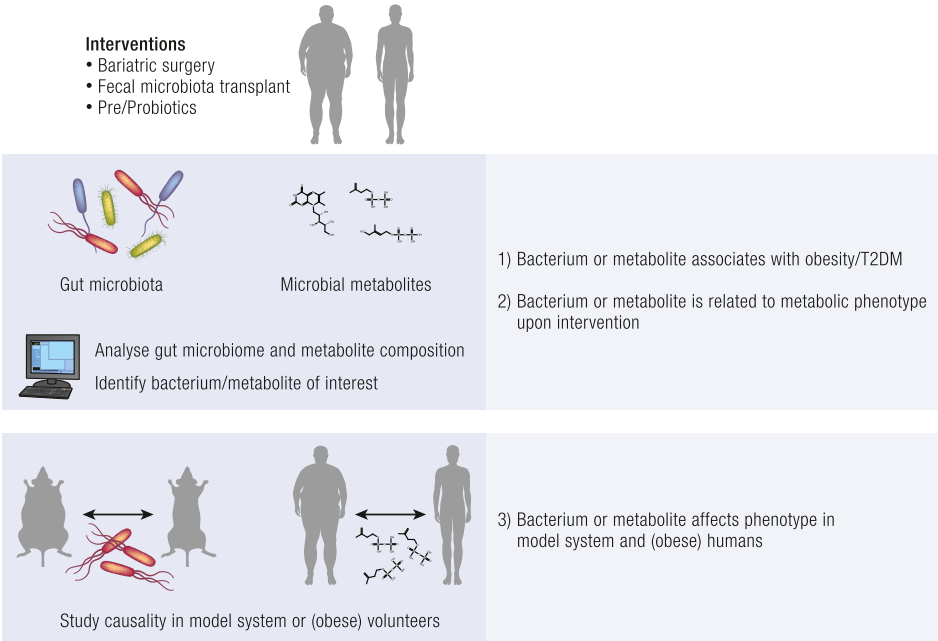


Figure 1: how to determine causality of the gut microbiome in cardiometabolic disease.

A top-down approach to determine a causal role of the microbiome/microbial metabolite in obesity and T2DM in humans. Microbiome and microbial metabolite composition is determined in lean and obese/T2DM subjects. Bacteria or metabolites of interest have to associate with these conditions (1). Following interventions that impact on microbiome composition of functional output and affect metabolic phenotype of the host, the bacterium/metabolite of interest has to correlate with the changes in phenotype (2). Bacterium or metabolite of interest accelerates or improves phenotype in a model system (*e.g.,* mouse models for obesity and T2DM) and in healthy or obese, insulin resistant volunteers, resp. (3).

Gut microbiome composition and function in cardiometabolic disease development: evidence from human studies

In line with studies in rodents, an increased ratio of Firmicutes/Bacteroidetes, which reduces with weight loss[69], has been associated with obesity in humans. Increased abundance of

Firmicutes was suggested to extract more energy from food[70]. In contrast, other research groups were not able to find differences in the ratio between Firmicutes/Bacteroidetes in obese versus lean subjects[71,72]. It is important to point out that technical difficulties and methodological discrepancies have been suggested to facilitate underrepresentation of bacterial groups, in particular of Bacteroides[73], thereby incorrectly indicating affected abundance between phyla. Furthermore, despite interesting findings on differences in the ratio Firmicutes/Bacteroidetes, it remains to be determined if this is a reflection of dietary intake or a driving factor of obesity. The relevance of these findings is therefore debatable.

Insulin resistance precedes development of T2DM and several metabolic markers thereof have been associated with Lactobacillus and Clostridium species[45]. Fasting glucose and HbA1c levels showed a positive correlation with Lactobacillus whereas Clostridium showed a negative correlation with these parameters[45]. Additionally, it has been shown that T2DM patients had reduced abundance of bacteria that produce the presumably beneficial short chain fatty acid butyrate[39].

Interestingly, in three (independent) metagenomic studies[17,32,74], obesity was associated with a reduced bacterial gene richness. Subjects with a less diverse gut microbiota composition were shown to have higher BMI, increased fat mass, reduced insulin sensitivity, dyslipidaemia, and increased markers of inflammation[17]. In addition, low bacterial richness was predictive for weight gain in a ten-year follow-up in which subjects with low bacterial richness had gained more weight compared to subjects with higher bacterial richness. As is the case for reported associations between improved ratio of Firmicutes/Bacteroidetes following weight loss, it remains to be determined if increased bacterial richness is a mere reflection of a healthy and varied diet, or that it directly contributes to the protection from obesity. Nevertheless, bacterial richness was simultaneously reported to have predictive potential for dietary interventions aiming to lose weight[31]. Metagenomics studies should be interpreted with caution, since a recent meta-analysis indicated that the reproducibility of metagenomics studies in humans is limited[75]. The authors concluded -after pooling datasets from several separate studies[69,74,76]- that there was no association between Body Mass Index (BMI) and taxonomic composition.

Bacterial metabolites and bacteria-derived components as modifiers of human metabolism

Changes in gut microbial output (metabolite production) or host exposure to bacterial-derived components (*e.g.*, endotoxin) have been suggested to play a larger role in metabolic disease development than microbial composition on the genome level *per se*[77]. A chronic, low-grade inflammatory state is often found in patients with obesity, insulin resistance and T2DM[78]. This increased inflammatory state has been proposed to be a driving factor in development of insulin resistance. In particular by reducing insulin sensitivity in muscle

and adipose tissue[79] and by impairing pancreatic islet function[80]. Although increased inflammatory tone in obesity is likely driven by multiple factors, studies in mouse models indicate that the gut microbiota is a causal factor in increasing inflammatory tone in obesity[81,82,83]. These findings lead to the hypothesis of metabolic endotoxemia in obesity and T2DM; a low-grade inflammatory state resulting from translocation of toxic, bacteria-derived components of mainly Gram-negative bacteria (*e.g.*, endotoxin)[81]. Significantly higher concentrations of LPS have indeed been measured in plasma from patients with T2DM compared to non-diabetic subjects[84,85,86]. Nevertheless, the risk of exogenous LPS contamination during analysis remains a topic of debate[87]. Blood from the GI tract drains into the portal vein; highest concentrations of LPS are therefore likely to be found in this compartment of the circulation. The portal vein is the main (75%) supplier of blood to the liver. High LPS influx into the liver potentially has significant consequences for inflammatory signalling pathways and insulin signalling in the liver. This hypothesis remains to be addressed since portal vein blood and liver biopsies are difficult to obtain. Increased translocation of endotoxin is potentially facilitated by a diet-induced increase in gut permeability and subsequent reduction in protective gut barrier function[88]. In line with increased gut permeability, humans predisposed to develop T2DM, had increased circulating levels of bacterial DNA[89].

The gut microbiota produces numerous organic compounds such as nitric oxide, ammonia, carbon oxide, indole and hydrogen sulphide, that possess pro- and anti-inflammatory properties and might be able to alter gut permeability[90]. Hydrogen sulphide has specifically gained interest in the past decades for its role in GI diseases[91] and CVD[92]. However, the role of these organic compounds in cardiometabolic disease is still under debate, partially due to the numerous conflicting studies which have been published. For example, H2S can potentially alter gut permeability[91] and increased levels of H2S are found in patients with ulcerative colitis[93], whereas H2S could have a protective role against nonsteroidal anti-inflammatory drug induced gastritis[94]. Interestingly, a recent paper showed that H2S possess cardioprotective effects during the cardiac remodelling process post myocardial infarction in rats by increasing macrophages infiltration into the infarcted myocardium and thus antagonizing hypoxia induced damage of cardiomyocytes[95,96]. In addition, H2S might have a beneficial role in the immune-inflammatory processes in atherosclerosis by inhibiting the macrophage-derived foam cell formation[97]. However, these studies are performed in murine models, in vitro or ex vivo, therefore the (causal) role of H2S in CVD and GI diseases has to be defined. The conflicting results regarding the inflammatory properties of H2S suggests that H2S may be a double-edged sword. Future research therefore needs to be focused on resolving these discrepancies and further investigate the role of this gaseous molecule on immune-inflammatory responses in CVD and GI disease.

Gut microbiota produce a large number of (yet to be defined) small molecules through primary (direct) or secondary (indirect) metabolic pathways[98]. It has therefore been suggested that the composition of gut-derived metabolites (as a measure for microbial output and functionality) is largely dependent on the diet of the host[18]. Although some of these metabolites might be retained within the gut ecosystem, others might be released in the circulation of the host and exert a diverse array of metabolic effects[99,100] . The bacterial metabolite trimethylamine (TMA) is an example of one of many gut-derived metabolites which has been associated with cardiovascular disease development in humans[101]. TMA is converted to trimethylamineoxide (TMAO) in the liver. High concentrations of TMAO were shown to accelerate atherosclerosis development in mice, and high concentrations of TMAO are correlated with a higher incidence of cardiovascular disease in humans[102]. Several studies have observed increased levels of TMAO in patients with T2DM than in healthy subjects[103,104,105]. Interestingly, increased levels of TMAO were observed in hepatic insulin receptor knockout mice (LIRKO mice) via upregulation of the TMAO-producing enzyme FMO3 in the liver[106]. Furthermore, knockdown of FMO3, the enzyme responsible for conversion of TMA to TMAO, prevented hyperglycemia, hyperlipidemia and atherosclerosis, suggesting that TMAO might be a potential player in diabetes associated atherosclerosis, at least in mice[106]. The first mechanistic link between TMAO and cardiovascular risk was provided in a study that showed that TMAO mediates blood platelet hyper-responsiveness and subsequent thrombosis[107].

Short chain fatty acids and (human) metabolic disease

Short chain fatty acids (SCFA) are produced by bacterial fermentation of non-digestible dietary fibres in the large intestine and mainly comprise acetate, propionate and butyrate. Studies in mice have shown that SFCA supplementation improves insulin-sensitivity and dyslipidaemia, prevents weight gain and increases energy expenditure in diet-induced obese mice[108,109]. SCFA-mediated activation of G-protein coupled receptor (GPR)-mediated signalling pathways are involved in several metabolic processes including enteroendocrine regulation[110]; glucagon-like-peptide (GLP) 1 secretion[111,112]; inflammatory response[113,114]; glucose uptake and fatty acid oxidation[115] and energy metabolism[39]. Murine and *ex vivo* experiments have shown that SCFA improve intestinal barrier function by a SCFA-mediated increase in transcription of mucin genes[116,117]. Improved gut barrier function prevents overt exposure to the innate immune system of the host, potentially reducing inflammatory tone. SCFA are inhibitors of histone deacetylases (HDACs). SCFA-mediated inhibition of HDAC in regulatory T cells (Treg) was shown to increase Forkhead box P3 (FOXP3) expression thereby affecting Treg generation[118,119]. In line, SCFA-mediated inhibition of HDAC has been shown to have anti-inflammatory properties by regulating intestinal macrophages[120] and dendritic

cells[121]. Depletion of SCFA, might therefore contribute to the increased inflammatory tone often found in patients with obesity and diabetes. The beneficial anti-inflammatory effects of SCFA in humans however remain to be further elucidated.

Through a complex intestine-brain-neural circuit, SCFA have been suggested to increase intestinal gluconeogenesis, thereby improving peripheral glucose production and insulin sensitivity[122]. In a recent study in rats, however, it was shown that the SCFA acetate increased food intake and promoted glucose-stimulated insulin secretion[109]. In humans, acetate supplementation was reported to facilitate short-term satiety[123] and reduce weight gain[124]. In line, SCFA reduced food intake and prevented weight by activating anorectic pathways in the brain[108]. Direct colonic delivery of propionate reduced weight gain in a randomized controlled study in 60 overweight subjects,[125]. In addition, fecal acetate levels have been inversely correlated to insulin resistance[126].

Despite these positive correlations between SCFA and metabolic health in humans, the fecal microbiota composition of obese subjects has been reported to be shifted towards increased numbers of SCFA-producing species compared to lean subjects[74,127]. In line, increased fecal concentrations of SCFA, especially butyrate, have been observed in obese subjects[38,99,128]. Interestingly, it was shown that in twins discordant for obesity, the gut microbiota of the obese twin was relatively enriched in SCFA-producing bacteria compared to the lean twin[64]. It has been proposed that an increased capacity to extract energy (in the form of bacterial SCFA production) from fibres might be a driving factor in obesity development. Thus, despite increased relative abundance of SCFA-producers and increased fecal SCFA content in obesity, it is difficult to interpret potential health benefits of SCFA in obese subjects. Following production, SCFA are rapidly absorbed by the host (at least in healthy subjects) where they regulate glucose and lipid metabolism. In addition, SCFA can be absorbed and converted by the gut microbiota itself. It can therefore be speculated that, despite increased SCFA production, the gut microbiota of obese subjects has reduced capacity to handle SCFA.

In contrast to obese subjects without diabetes, fecal microbiota from obese, T2DM patients, has been shown to be relatively depleted in SCFA-producing bacterial species[45,39]. In line, vancomycin treatment of metabolic syndrome patients reduced insulin sensitivity with a coinciding reduction in butyrate-producing bacteria[129]. A study in which the fecal microbiota of lean, insulin-sensitive donors was transplanted to insulin-resistant metabolic syndrome recipients, demonstrated that improved insulin sensitivity following fecal microbiota transplantation correlated positively with abundance of butyrate-producing bacteria[130]. A metagenomic study showed that metformin-naïve T2DM patients could be associated with a decrease in genera of butyrate-producers (*e.g.*, Roseburia spp, Subdoligranulum spp)[46]. In the same study, it was shown that the gut microbiota of metformin-treated T2DM patients

contains significantly more butyrate and proprionate-producers compared to T2DM patients not treated with metformine.

Bile acid signalling in host metabolism

Bile acids play a pivotal role in human health and metabolic disease development, mainly by their role as signalling molecules that can activate receptors in the gut, liver and adipose tissue[131]. Primary bile acids (cholic and chenodeoxycholic acids) are produced from cholesterol in the liver via a complex pathway including at least 17 enzymes and is under control of the nuclear Farnesoid X receptor (FXR) and its downstream targets FGF15/19 (in intestine) and SHP (in liver)[132]. Mice also produce alpha- and beta-muricholic acids (α/βMCAs) in addition to the primary bile acids found in humans[133]. Upon secretion into the intestine, bile acids are subject to modifications by the gut microbiota[133,134,133]. Primary bile acids are metabolized into secondary bile acids (deoxycholic and lithocholic acids) following 7α-dehydroxylation, which compromises numerous reactions carried out by bacteria that mainly belong to the Firmicutes[135,136]. It was shown that in mice, gut microbiota regulates expression of several key enzymes in bile acid formation including CYP7A1 and CYP27A1 by changing the composition of the bile acid pool, thereby alleviating FXR inhibition[137]. In addition to bile acid synthesis and modification, bile acid uptake in the gut has been suggested to be regulated by the microbiota. Expression of the apical sodium dependent bile acid transporter (ASBT), a transporter found in the small intestine responsible for the uptake of bile acids, is reduced in conventionally raised mice compared to their germ-free counterparts[137].

Data underscoring a role for bile acids in metabolic disease development originate in large from mouse studies. For example, cholic acid supplementation reduced HFD-induced weight gain and attenuated insulin resistance in mice, coinciding with increased circulating levels of bile acids[138]. FXR and downstream target FGF15/19 have been shown to regulate glucose and lipid metabolism[139]. Synthetic inhibition of FXR reduced bile acid pool size and attenuated weight gain and glucose intolerance in HFD-fed mice[140]. Furthermore, by acting on the GPR TGR5, bile acids have been shown to promote (antidiabetic) GLP1 secretion[141] and increase energy metabolism[138]. Higher concentrations of deoxycholic acid (DCA) have been associated with obesity in mice[142]. Tauro- β-muricholic acid, an endogenous FXR antagonist[143,137] is metabolized by the gut microbiota. Therefore, germ free mice are not able to metabolize this bile acid. This ability has been shown be a prerequisite to induce obesity, hepatic steatosis, impaired glucose tolerance and reduced insulin sensitivity[137]. The bile acids receptors FXR and TGR5 might play an important role in the development of metabolic diseases and have become major targets in translational and intervention studies[131]. Bile acids generated by the gut microbiota can modulate signalling through these bile acid

receptors and therefore might have the potential to alter lipid and glucose metabolism in humans.

In humans, bile acids have been implicated in regulation of food intake[144]. Furthermore, increased circulating levels of bile acids have been reported in obese, T2DM subjects[145] and were shown to correlate with BMI[144]. Rectal administration of TDCA improved glucose homeostasis and lowered food intake in obese, T2DM subjects[146]. Particular interest in a role for bile acids in regulation of host (energy) metabolism, however, arose from observations that (postprandial) bile acid metabolism is severely affected following bariatric surgery[147]. Circulating levels of primary and secondary bile acids are increased after bariatric surgery and correlate with improved glucose control[148,149,147]. Bile-acid mediated signalling events have been reported to be increased in post-Roux-en-Y Gastric Bypass (RYGB) subjects: this correlated with the release of satiety-promoting gut hormones such as GLP-1 and PYY[150,151,152]. Furthermore, supporting an important role for bile acids in RYGB-mediated improvements in glucose homeostasis: metabolism of FXR knock-out mice is not improved following vertical sleeve gastrectomy[153]. The beneficial effects of RYGB on energy metabolism were reproduced by diverting the biliary flow from duodenum to ileum in rats, suggesting that bile acids play an important role in adiposity, liver steatosis, lipid and glucose metabolism[154]. The animals in this study lost approximately 20% of their body weight, therefore these results must be interpreted with caution, since these results can be partially explained by weight loss.

Crosstalk between the gut microbiota and bile acids affect host metabolism. However, most of the studies that mechanistically assess pathways involved in this crosstalk were performed in animal models. The Human and rodent bile acid pools have major compositional differences. This has significant consequences for bile acid signalling properties and conclusions derived from rodent studies have to interpreted with caution. In humans, causal evidence supporting a role for changes in microbiota composition with subsequent bile acid-mediated changes in host metabolism remains largely unknown.

Gut microbiome alterations after bariatric surgery

Bariatric surgery is a last resort for the treatment of morbid obesity and -related complications such as T2DM and is superior to any other treatment regimen aiming to reduce weight[155,156]. The rapid improvement in metabolic parameters such as fasting glucose[157] and fasting insulin[158], (usually within days after surgery) can be explained in large by calorie restriction[159,160]. Bariatric surgery has significant effects on gut microbiome composition, induced by considerable alterations in the gastrointestinal tract (*i.e.,* reduced caloric intake, reduced gastric emptying, alterations in gastric acid production and bile acid[161]. Tremaroli *et al.*[162] showed that two distinct bariatric surgery procedures, *i.e.,* vertical

sleeve gastrectomy (VSG) (no intestinal diversion) and RYGB (with intestinal diversion), have similar effects on gut microbiome composition. Nevertheless, and despite small sample size, functional shifts were apparent and differed between the two surgical procedures and between the control group[162]. RYGB has significant effects on gut microbial composition, the abundance of Firmicutes, which is generally high in obesity, decreases and Proteobacteria increases following RYGB *REF. These effects differ strongly from effects of diet-mediated weight loss[71]. However, a recent meta-analysis showed that there is a high discrepancy in human studies investigating gut microbial alterations after bariatric surgery[163], therefore these results have to be validated in larger cohorts.

Altered microbiome composition and microbial metabolic output (*e.g.,* metabolite production) after bariatric surgery was hypothesized to add to the long-term beneficial effects of this surgical procedure on weight loss, diabetes remission and cardiovascular risk[164]. In support of this hypothesis, microbiota of murine RYGB donors augmented weight gain in GF recipients compared to GF mice that had received microbiota from sham-operated donors[164]. Similar effects were observed in GF mice that received fecal microbiota transplants from human RYGB or VSG donors compared to GF mice transplanted with feces from obese controls[162]. Mice colonized with microbiota from RYGB treated mice had higher lean mass and lower respiratory quotient (ratio between CO2 produced and O2 consumed) compared to VSG and control group, indicating decreased utilization of carbohydrates and increased utilization of lipids in the RYGB recipient mice[164].

Bariatric surgery is associated with significant changes in gut microbial composition and functionality[162,163]. However, large prospective studies are needed to validate these alterations and to further investigate whether the gut microbiome contributes to the beneficial metabolic effects of bariatric surgery. Bearing in mind the great dissimilarities in metabolic outcome[165] (responder, non-responder), it would be interesting to have follow up data available of the gut microbiome composition, diversity, and functionality years after the initial surgery.

The gut microbiome and type 1 diabetes mellitus

Although the main focus of this review is on the role of the gut microbiome in development of T2DM, the gut microbiome has also been implicated in the pathogenesis of T1DM. Both disorders are characterized by alterations in host immune response and have been linked to an immune system-gut microbiota interaction[166]. Interestingly, enhanced systemic inflammation and autoimmunity can be detected years before disease onset. This suggests that environmental factors, including changes in gut microbiota composition and output (*e.g.,* LPS, SCFA production), are determinants of disease progression and can have predictive value for those at risk to develop T1DM. It has been suggested that shifts in gut microbial

communities indeed precede disease development[167]. Nevertheless, in humans it is difficult to determine whether an altered microbiota, as observed in T1DM patients, is causal to or a consequence of compromised immune function. In addition, studies performed in humans are often subject to major confounding factors.

T1DM is generally considered to be driven by an (auto)immune-associated destruction of insulin producing pancreatic beta cells[168,169]. Approximately 70-90% of patients with T1DM show features of an immunological contribution (*e.g.,* self-reactive autoantibodies such as IA2 and GAD, genetic associations with genes controlling immune response)[170]. The remainder of T1DM cases can be classified as monogenic forms of T1DM, including certain types of maturity-onset diabetes of the young (MODY)[171] or have a yet to be determined pathogenesis. T1DM generally manifests early in life. Interestingly, most children are diagnosed in autumn and winter[172], and being born in spring is associated with a higher change of developing T1DM[173]. This suggests that the pathogenesis of T1DM is heterogeneous and environmental (seasonal) influences might initiate or even drive the pathogenic processes in T1DM. A plethora of environmental factors such as vitamin D deficiency[174,175], infant and adolescent nutrition[176] and early enterovirus infection[9] all have been postulated to contribute to the development of T1DM[6]. Improved sanitation and decreased incidence of childhood infections over the past decades, is associated with an increased incidence of autoimmune diseases such as T1DM and led to the hygiene hypothesis[177,178]. According to this hypothesis, infants may benefit from early exposure to specific microorganisms and parasites; this stimulation of the immune system early in life was indeed associated with lowered risk to develop allergies and autoimmune diseases later in life[177,178,179]. Removing microbes from an individual's living environment therefore has consequences for gut microbiome composition and development of the immune system. These associative studies have increased interest in the role of the gut microbiome in the development of T1DM in the past decade. As for T2DM, however, mechanistic evidence for a role of the gut microbiota in the pathophysiology of T1DM is mainly derived from studies in rodents.

Studies in BioBreeding Diabetes Prone (BB-DP) rats[180] and non-obese diabetic (NOD) mice[181] that were treated with antibiotics indicated that the subsequent alterations in gut microbial composition reduced the risk of T1DM development. In 2008, a landmark paper by Wen *et al*[182] showed that MyD88, which functions as a critical signal transducer in interleukin (IL)-1 and TLR signalling pathway, deficient NOD mice are protected from the development of T1DM[182]. Interestingly, the protection of developing T1DM is lost when deficient MyD88 mice are housed under GF conditions, suggesting that an interaction between the gut microbiota and the innate immune system has a role in the development of T1DM. In addition to shifts in gut microbial composition as contributing determinant for development of T1DM, microbial output in the form of SCFA's have been implicated to

elevate the number and enhance the function of intestinal Treg cells and T helper 17 cells (Th17 cells)[183,119,118]. Treg cells and Th17 are lymphocyte subsets with opposing actions[184]. An imbalance between Treg cells (anti-inflammatory) and Th17 cells (pro-inflammatory) has been shown to contribute to the pathophysiology of autoimmune diseases[184]. Since T1DM is a T-cell-mediated disease associated with a reduced number of dysfunctional Treg cells[185,186], an imbalance between Treg cells and Th17 cells could therefore augment an inflammatory response[184]. Interestingly, Th17 cells are important in maintenance of intestinal barrier function[187]. In a recent study, it was shown that antibiotic treatment reduced the number of Th17 cells in the lamina propria and increased T1DM incidence in NOD mice[188]. This result corresponds with earlier findings[189] and strengthens the hypotheses that an increased intestinal permeability might precede the clinical onset of T1DM[190]. The importance of a gut microbiome capable of producing sufficient SCFA was underscored by a study in which mice were fed diets supplemented with acetate and/or butyrate[191]. The acetate yielding diet decreased the number of activated diabetogenic T cells in lymphoid tissue. The butyrate-supplemented diet markedly increased the number and function of Treg cells and increased the expression of the tight junction protein occludin in the colon thereby preserving gut integrity. An intriguing interplay between genetics, altered gut microbiome/metabolites and immunity might play a role in the development of T1DM. Rodent studies have provided insight in this interplay, however, human data is scarce. In the next paragraph, we will discuss studies involving the gut microbiome in the development of T1DM in humans.

Comparison of the faecal bacterial composition of four pairs of T1DM infants and controls revealed a higher ratio of Bacteroidetes/Firmicutes approximately six months after birth in infants who developed T1DM compared to the controls[10]. This corresponds to other studies that reported increased ratio of Bacteroidetes/Firmicutes in T1DM children[192,193]. In addition, the diversity of the gut microbiome was less diverse in T1DM subjects compared to the controls[10]. Seroconversion is the time between development of a specific antibody till moment of detection of this antibody in the circulation[194]. In infants that later developed T1DM, detectability of anti-islet autoantibodies coincided with reduced abundance of bacterial genes associated with SCFA production and with gut intergrity[193]. This indicates that changes in early autoantibody production is related to changes in microbiome functional output.

In line, infants that expressed at least two diabetes-associated autoantibodies had low abundance of lactate and butyrate producing species compared to autoantibody-negative infants[195]. In the BABYDIET study[196], infants with first-degree relatives with T1DM and HLA genotypes associated with increased risk to develop T1DM at 6 or 12 months, had similar gut microbial composition and diversity compared to controls[197]. Interestingly, however, alterations in microbial interactions networks were observed in infants who developed

anti-islet cells autoantibodies[197]. In a longitudinal prospective cohort of 33 HLA-matched infants followed from birth until three years of age, decreased microbial diversity correlated with seroconversion, thus prior to the diagnosis of T1DM[198]. Furthermore, levels of human beta-defensin 2 (hBD2) were increased in infants who later developed T1DM[198]. Since hBD2 is an antimicrobial product produced by colonic epithelial cells during inflammation[199,200], this finding supports the hypothesis that development of T1DM is accompanied by intestinal inflammation. A case-control study in ten infants who were at risk to develop T1DM (*i.e.,* positive for at least two diabetes-associated autoantibodies) reported higher intestinal permeability as assessed by a lactulose/mannitol test in those infants compared to controls[201]. A possible mechanistic explanation for the contribution of the gut microbiome in the development of T1DM comes from a recent perspective study in Finland, Estonia and Russia[179]. Finland and, albeit to lesser extent, Estonia, have higher autoimmune disease prevalence, including T1DM, compared to neighbouring Russia. Gut microbiome development was followed from birth until the age of three in 222 infants and differed markedly between infants from Finland and Estonia compared to Russian infants. Of particular interest was a marked reduction in Bacteroides species in Russian infants compared to infants from Finland and Estonia. Functional pathway analysis suggested that early microbial communities of infants from Finland and Estonia produced more LPS compared to their Russian counterparts. However, LPS produced in this cohort was mainly derived from Bacteroides species: Bacteroides-derived LPS differs structurally and functionally from LPS derived from for example *E. coli* and has been shown to be *non-immunogenic* in mice[179]. Furthermore, in contrast to E. coli LPS, Bacteroides-derived LPS did not decrease incidence of autoimmune diabetes in NOD mice. Although a clear link between Bacteroides-derived LPS and T1DM could not be made in this study, these data raise the interesting hypothesis that the nature and composition of different LPS subtypes might determine the level of immune activation and serve protective roles in autoimmune disease development[179]. In line, based on the relation between celiac disease and T1DM, the intestine and its inhabitants might be a shared risk factor[202]. These findings however, have to be interpreted with caution since geographically, gut microbial composition varies considerably young children at risk to develop T1DM[203]. A framework for the potential role of the gut microbiome in the development in T1DM is given in figure 2.

Interventions in humans: diet, pro and prebiotics and fecal microbiota transplantation

As mentioned previously, gut microbiota composition as well as gut microbial function is highly related to dietary intake of the host[26]. A relative deprivation in plant-based dietary fibres in industrialized nations has been suggested to be a driving force behind the widespread change in functional capacity of the gut microbiota potentially contributing

to the increasing prevalence of obesity and -related complications[204,45,39]. In addition to macronutrient intake, it has been shown that food additives such as artificial sweeteners induce both compositional and functional changes in gut microbiota and augment features of the metabolic syndrome[205,206]. Since the gut microbiota is easily accessible and responds rapidly to changes in nutrient composition, dietary reinforcements have been put forward as an attractive therapeutic target for obesity. However, in addition to low overall adherence to diets, high inter-individual differences in response to diet makes this a challenging endeavour. For example, complex carbohydrate supplementation increased starch-degrading taxa in some but not all subjects who participated in a strictly controlled 10-week dietary intervention[30]. Another study showed that a low calorie, high fibre diet increased diversity only in subjects with high-gene count at baseline[31]. Subjects with improved glucose metabolism after a three-day intervention with whole grain bread had higher ratio of Prevotella/Bacteroidetes after the interventions than non-responders[32].

The inter-individual response to diet was particularly exemplified in a landmark study by Zeevi and coworkers who showed that the individual postprandial glycemic response to a high glycemic meal was highly variable[33]. This response correlated with individual microbiota composition. Interestingly, using a machine-learning approach, the individual response to and success of a particular dietary regimen could be predicted based on existing microbiota composition. This work provided crucial insight in the role of the gut microbiota in responsiveness to dietary strategies. Using novel, personalized pre- or probiotics to modify the gut microbiota composition of people predicted to have low response rate to diet-induced weight loss regimens into a more responsive composition might optimize the effectiveness of dietary strategies.

Probiotics are living microorganisms that either have potential to improve host metabolism directly (*e.g.*, by improving gut barrier function or increasing SCFA-production) or have the capacity to re-establish a more favourable intestinal balance by modulating pH, antibacterial compound production and competing with pathogens[207,208]. In mice, the probiotic strains Akkermansia muciniphila[209] and Lactobacillus planetarium[210] were both shown to lower endotoxemia and weight gain in HFD-fed mice. In humans, administration of Lactobacillus reuteri was associated with increased insulin secretion in obese, insulin resistant subjects[211]. Furthermore, a double–blinded, randomized, placebo-controlled intervention trial in overweight subjects showed beneficial effects of Lactobacillius gasseri on weight loss compared to fermented milk use only[212].

Prebiotics are nonmicrobial entities (usually dietary fibres) that elicit a favourable impact on microbial composition and function. Prebiotics might therefore be a feasible tool to modulate gut microbiota. Supplementation of prebiotics has been associated with improved plasma lipid levels and improved glycemic control in both humans[209,213] and

mice[214]. Oligofructose was shown to increase release of the satiety promoting hormone PPY and GLP1 in mice[215]. In T2DM humans, oligofructose proved to be a useful prebiotic: supplementation for three months increased weight loss and improved glucose control compared to patients receiving placebo[216]. Despite focus of (food) industry on development of novel pre- and probiotics to modulate microbiota composition and/or functional output to subsequently improve host metabolism, thus far, only minimal beneficial effects of gut microbiota-mediated on metabolism have been obtained. Improved engraftment of probiotic strains might help improve effectiveness of probiotic strain administration[217].

Fecal microbiota transplantation

FMT has a long medical history and has been used for treatment of several gastro-intestinal illnesses. As long as 1700 years ago, FMT was used to treat patients with food poisoning and diarrhoea in China[218]. After the realization that hygiene plays an important role in preventing infectious disease, FMT became obsolete. In 1958, after a long period of silence, FMT garnered interest again following a describing its use in treating fulminant enterocolitis[219]. The real breakthrough of FMT as treatment modality was after publication of an open-label, randomized, controlled trial, which demonstrated that the resolution of C. diff infection was 94% after FMT compared to 31% efficacy of conventional vancomycin treatment[220]. FMT is now the method of choice for treatment of recurrent Clostridium difficile infection. However, FMT is also of interest as therapeutic modality for a wide range of diseases including inflammatory bowel disease[221], obesity[58] and metabolic syndrome[58]. In addition, neurological[222] and psychological disorders[223], might benefit from FMT if correlations with altered gut microbiome composition are indeed valid. In two separate studies in humans, our group has shown that FMT has beneficial effects on insulin sensitivity[42,58]. Although effects are temporal and variable, FMT might have merit as intervention option for metabolic syndrome.

Peripheral insulin sensitivity of obese, insulin resistant subjects was significantly improved six weeks after receiving a transplant from a lean, insulin sensitive donor (allogenic transplant)[130]. Transplantation of own fecal microbiota (autologous transplant) did not affect insulin sensitivity. In a second, larger cohort of obese, metabolic syndrome subjects, we were able to reproduce these findings. Allogenic FMT improved insulin sensitivity compared to autologous FMT in metabolic syndrome recipients six weeks after transplantation. Interestingly, 12 weeks after transplantation, this beneficial effect could no longer be observed. Engraftment of donor microbiota in the gut of recipient negatively associated with metabolic outcome of FMT suggesting that specific donor-host interactions are important determinants of FMT efficacy. In line, based on baseline microbiota composition of the recipient, the metabolic response to FMT could be predicted.

Variation in experimental protocols, preparation of fecal samples and diurnal oscillations of the gut microbiota[224] are additional explanations for variable efficacy of FMT. This underscores the need for development of stringent standard operation procedures (SOPs)[54]. In addition to bacteria, viruses, fungi and bacteriophages that all reside in the fecal compartment are co-transplanted[225]. Although it has yet to be determined if and how these components contribute to FMT efficacy, it was recently shown that the virome (bacteriophages) has an important role in host health by modulating bacterial community and by direct interaction with host cells[226,227]. In addition, eukaryotic viruses and bacteriophages have been shown to modulate bacterial metabolism (*e.g.,* amino acid, lipid, and carbohydrate metabolism) and to affect signal transduction pathways and transcriptional regulation[226,228].

Modulation of the gut microbial composition and functionality by FMT only partly affects the intrinsic and complex pathophysiology of obesity and T2DM. Gut microbiome composition and function is influenced by many factors and therefore, it is unlikely that a single FMT can cure obesity or T2DM. Nevertheless, a combination of FMT with personalized prebiotics or treatment with 'missing' intestinal bacterial strains (drugging the microbiome) might enhance the effects of conventional treatment strategies[54]. Furthermore, early intervention in patients who are at risk to develop T2DM or patients who were recently diagnosed with these pathologies might benefit from gut microbial modulation in a personalized manner, such as a microbiota-based dietary strategies or personalized FMT.

The human gut microbial community is shaped by a complex interplay between host genetics, diet and (history of) medication use. Alterations in gut microbiota composition (*e.g.,* reduced diversity) or microbial output (*e.g.,* LPS subtypes, SCFA production or bile acid conversion) have been implicated in development of metabolic diseases such as obesity and T2DM in humans. Although mechanistic evidence for a causal role of the gut microbiota in the pathophysiology of these diseases in humans is scarce, currently available data suggests that an altered microbiota composition affects gut barrier function and induces (low-grade) inflammatory events, either locally in the intestine or systemically. Furthermore, bacterial metabolites including SCFA's and secondary bile acids, which serve important regulatory roles in energy homeostasis and regulation of peripheral glucose and lipid metabolism, have been hypothesized to be drivers of T2DM development. T1DM is generally considered to be driven by autoimmune antibodies that specifically destroy insulin-producing Beta cells in the pancreas. It has been hypothesized that autoimmune antibody generation is in part consequential to removal of microbes -with a crucial role for maturation of the immune system- from our living environment (hygiene hypothesis). In addition, some studies have observed increased autoimmune antibodies *prior* to T1DM diagnosis and have suggested that altered microbiota composition or microbial output and subsequent initiation of inflammatory events accelerates onset of T1DM. Please see text for details.

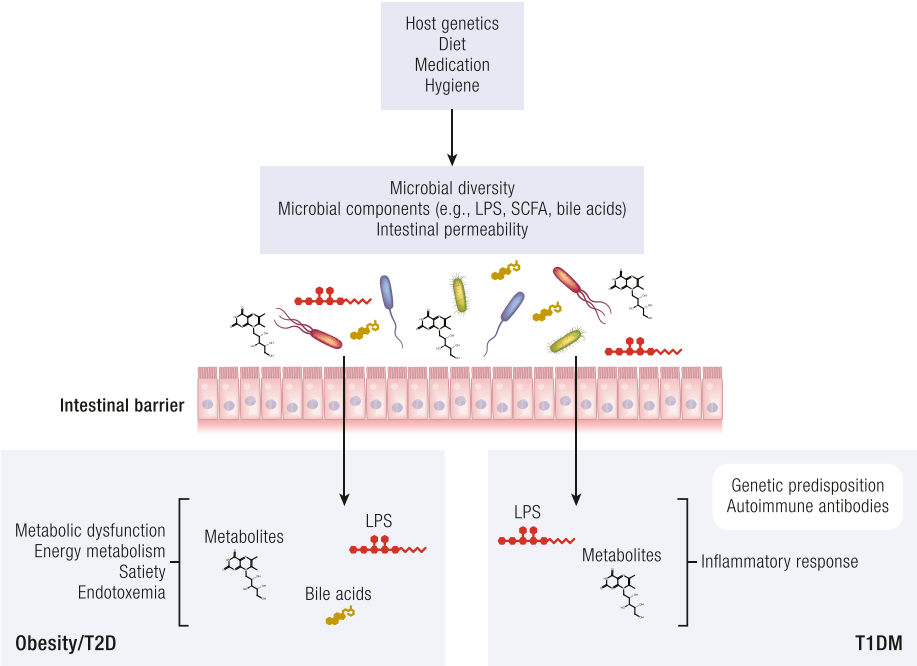


Figure 2: The role of the gut microbiome in the pathogenesis of cardiometabolic disease.

Systems approach potentially holds the key to establish driving role of the gut microbiota in obesity and T2DM

The introduction of DNA sequencing technologies significantly boosted the study of complex microbial communities and allowed for taxonomic identification of individual microbes. Nevertheless, early sequencing technologies were slow and expensive since large genomic fragments had to be cloned into plasmid vectors, transformed into suitable hosts for amplification prior to being sequenced[229].

PCR-based, massive parallel sequencing now allows for identification of previously undetectable bacteria within complex communities[229]. In addition, shotgun whole-genome sequencing (WGS) approaches have significantly enhanced detection of diversity, increased prediction of genes and taxa at species level can be identified[230]. Although WGS is currently more expensive and requires more extensive data analysis, this method is preferred above PCR based sequencing.

High through-put amplicon sequencing of isolated DNA samples or PCR amplification of regions within universally conserved 16S rRNA genes has generated an enormous amount of data on microbiome composition from different environments and conditions. Reference metagenomes of microbes were published in 2012 by the Human Microbiome Consortium (HMC) and showed that the dominant microbial taxa in the human gut

include Bacteroidetes, Firmicutes, and Proteobacteria, and that species including Bacteroides fragilis, Bacteroides melaninogenicus, Enterococcus faecalis, and E. coli are present in the majority of healthy human subjects[231]. It is important to note that taxonomic characterization of intestinal microbiota is based on relative (and not absolute) abundance and does not always translate into function. In order to effectively understand the impact of the microbiome on the host, it is critical to connect compositional to functional studies. This can be undertaken with a systems biology approach.

Systems biology approaches can be used to integrate omics data to untangle driving factors underlying gut microbiota composition (figure 3). Additionally, these approaches can provide insight in the hierarchy of mechanisms underlying the development of metabolic diseases. Taxonomic profiling can identify who's there, and complementary with metagenomic profiling: what are they capable of? Nevertheless, it is important to emphasize that the vast amount of data generated by high-throughput sequencing currently surpasses the ability to analyse those data with currently available bioinformatics tools.

A study in which the overlap between metagenomics (which microbes are there, and what genetic potential do they have) and metatranscriptomics (which genes are highest expressed) of the human gut microbiome community was systematically compared, revealed that only 41% of microbial transcripts corresponds to microbial genomic abundance[232]. This underscores the importance to move beyond metagenomics in order to understand 'what the community is really doing'. Moreover, the importance of post-transcriptional and translational regulation and the fact that protein abundance does not correspond with gene expression in either eukaryotes or bacteria has to be taken into account[233,234]. Metaproteomics and metabolomics might be the solution to better understand the functional capacities of the microbiota. The interpretation of metatranscriptomic and proteomic data is challenging due to incomplete information on the gut microbial genomes and proteomes and, hence, lack of (gut-specific) reference databases[229]. Moreover, metatranscriptomics provides only a snapshot of the dynamic interactions between host, gut microbiota and environment[235]. An example of the question 'who´s there' and more important: 'who's active', comes from a study where metaproteomics was combined with taxonomic profiles of gut microbiome obtained from obese and non-obese individuals[236]. This study demonstrated that, despite a lower abundance, Bacteroidetes had higher metabolic activity in obese individuals than in non-obese individuals. In addition, insulin sensitivity, as estimated by homeostasis model assessment (HOMA) index, was positively associated with peptides originating from a group of proteins derived from bacteria from the genus Ruminococcus[236].

Another approach for the study of microbiota functionality and host–microbiota co-metabolism is provided by metabolomics, which analyses the small-molecule composition of host fluids and tissues to detect gut micobial metabolites[236]. Metabolomics can be grouped

in targeted and untargeted methodologies. With untargeted metabolomics, up to 10.000 independent spectral features can be measured[237]. Thus far, however, only one third of these spectra can be identified, since translating the signals obtained by mass spectrometry to a specific/known chemical structure is still low throughput[237]. Targeted metabolomics is a quantitative technology since it measures known metabolites in clusters with similar chemical structure[238]. This approach has revealed the metabolic syndrome and other metabolic diseases are associated with microbial metabolites[238]. Targeted metabolomics has also shown that shifts in microbial gene functions are coupled to shifts in community functionality, as has been done by profiling SCFAs[239] and bile acids[240,241] in obesity and after bariatric surgery-mediated weight loss[162,153]. A systems biology approach with combined input of different omics data sets will accelerate our understanding of the contribution of the microbiome to human health and metabolic disease.

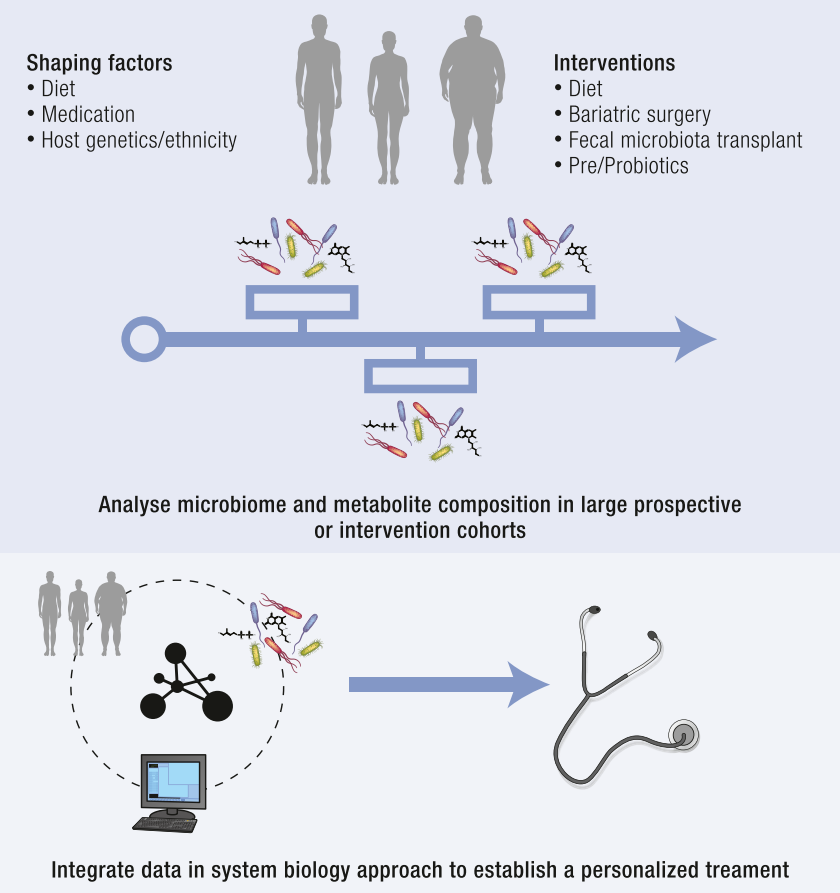


Figure 3: The role of systems biology approach in gut microbiome research.

Systems biology approaches that combine patient data with microbiome and microbial metabolite composition in pre- and postintervention settings and in large prospective cohorts of initially healthy subjects will reveal crucial insight in the role of the gut microbiome in metabolic disease development. Moreover, this strategy will allow for development of tools to predict metabolic disease development and to specify optimal treatment strategies to tackle these pathologies.

Concluding remarks

Since the introduction of NGS techniques, a plethora of studies has shown striking associations between the composition of the gut microbiota or gut microbial metabolites in the development of obesity and diabetes. Nevertheless, only few studies have provided mechanistic or causal evidence of the pivotal role of the gut microbiota in the development of metabolic diseases in humans. The complex interplay between ethnicity, host genetics, dietary habits and medication use all play an important role in shaping the microbial community and therefore makes it an intriguing yet challenging research field. Inconsistent or lack of corrective measures for confounding factors that might underlie changes in the gut microbial composition is a challenging aspect of data interpretation in humans. It likely is one of the main causes for the low reproducibility of research results between studies. Large prospective studies will be of critical importance to answer if gut microbial composition is a reflection of the disease itself or if the microbial composition was affected prior to disease development and hence was a driving factor. Although studies using antibiotic therapy or fecal microbiota transplantation are suggestive of causal linkage between the gut microbiota and metabolic disease development, effect size and evidence for causality are still marginal. Furthermore, these studies do not provide mechanistic insight in the interplay between the gut microbiota and host metabolism. Prospective and intervention studies in large human cohorts combined with dedicated mechanistic studies in model systems are required to understand if and how gut microbiota affects metabolic disease development. Using a multi-omics approach, a deeper understanding of host-microbe, microbe-microbe and diet-microbe interactions can be achieved. This will provide insight in the hierarchy of mechanisms underlying the development of metabolic diseases and lead to identification of a personalized intestinal microbiota signature. This will accelerate development of strategies to predict cardiometabolic disease development and, importantly, establish means to develop personalized, microbiota-based interventions to tackle metabolic pathologies in humans.

Conflicts of interest

M.Nieuwdorp is in the scientific advisory board of Caelus Health, the Netherlands.

Grant support

M. Nieuwdorp is supported by a ZONMW-VIDI grant 2013 (016.146.327) and CVON Young Talent grant 2012. H Herrema is supported by is supported by CVON 2012 grant (IN-CONTROL) (2012-03); to M.N.

REFERENCES

1. Palermo, A., Maggi, D., Maurizi, A. R., Pozzilli, P. & Buzzetti, R. Prevention of type 2 diabetes mellitus: is it feasible? *Diabetes/Metabolism Research and Reviews* **30**, 4–12 (2014).
2. Friedman, J. M. Modern science versus the stigma of obesity. *Nature Medicine* **10**, 563–569 (2004).
3. International Diabetes Federation (IDF). *IDF Diabetes Atlas 7th edition*. *idf.org* (2015). doi:10.1289/image.ehp.v119.i03.
4. Defronzo, R. A. Banting Lecture. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. *Diabetes* **58**, 773–95 (2009).
5. DeFronzo, R. A. The Triumvirate: β-Cell, Muscle, Liver: A Collusion Responsible for NIDDM. *Diabetes* **37**, (1988).
6. Atkinson, M. A., Eisenbarth, G. S. & Michels, A. W. Type 1 diabetes. *The Lancet* **383**, 69–82 (2014).
7. Dabelea, D. The accelerating epidemic of childhood diabetes. *The Lancet* **373**, 1999–2000 (2009).
8. Soltesz, G., Patterson, C., Dahlquist, G. & EURODIAB Study Group. Worldwide childhood type 1 diabetes incidence ? what can we learn from epidemiology? *Pediatric Diabetes* **8**, 6–14 (2007).
9. Yeung, W.-C. G., student, medical, Rawlinson, W. D., Craig, M. E. & professor, associate. Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. doi:10.1136/bmj.d35.
10. Giongo, A. *et al.* Toward defining the autoimmune microbiome for type 1 diabetes. *The ISME journal* **5**, 82–91 (2011).
11. Savage, D. C. Microbial Ecology of the Gastrointestinal Tract. *Annual Review of Microbiology* **31**, 107–133 (1977).
12. Sender, R., Fuchs, S. & Milo, R. Commentary Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–340 (2016).
13. Eckburg, P. B. *et al.* Diversity of the Human Intestinal Microbial Flora. *Science* **308**, 1635–1638 (2005).
14. Gill, S. R. *et al.* Metagenomic Analysis of the Human Distal Gut Microbiome. *Science* **312**, 1355–1359 (2006).
15. Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. Development of the Human Infant Intestinal Microbiota. *PLoS Biology* **5**, e177 (2007).
16. Scheithauer, T. P. M., Dallinga-Thie, G. M., de Vos, W. M., Nieuwdorp, M. & van Raalte, D. H. Causality of small and large intestinal microbiota in weight regulation and insulin resistance. *Molecular Metabolism* 1–12 (2016) doi:10.1016/j.molmet.2016.06.002.
17. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).
18. Sonnenburg, J. L. & Bäckhed, F. Diet–microbiota interactions as moderators of human metabolism. doi:10.1038/nature18846.
19. Koren, O. *et al.* Host Remodeling of the Gut Microbiome and Metabolic Changes during Pregnancy. *Cell* **150**, 470–480 (2012).
20. Prince, A. L. *et al.* The placental membrane microbiome is altered among subjects with spontaneous preterm birth with and without chorioamnionitis. *American journal of obstetrics and gynecology* **214**, 627.e1-627.e16 (2016).
21. Gosalbes, M. J. *et al.* Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. *Clinical and Experimental Allergy* **43**, 198–211 (2013).
22. Putignani, L., Del Chierico, F., Petrucca, A., Vernocchi, P. & Dallapiccola, B. The human gut microbiota: a dynamic interplay with the host from birth to senescence settled during childhood. (2014) doi:10.1038/pr.2014.49.
23. Penders, J. *et al.* Factors Influencing the Composition of the Intestinal Microbiota in Early Infancy. *Pediatrics* **118**, 511–521 (2006).
24. Huttenhower, C. & Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–14 (2012).
25. Hollister, E. B. *et al.* Structure and function of the healthy pre-adolescent pediatric gut microbiome. *Microbiome* **3**, 36 (2015).

26. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–63 (2014).

27. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).

28. Earle, K. A. *et al.* Quantitative Imaging of Gut Microbiota Spatial Organization. *Cell Host & Microbe* **18**, 478–488 (2015).

29. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14691–6 (2010).

30. Walker, A. W. *et al.* Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *The ISME journal* **5**, 220–30 (2011).

31. Cotillard, A. *et al.* Dietary intervention impact on gut microbial gene richness. *Nature* **500**, 585–588 (2013).

32. Kovatcheva-Datchary, P. *et al.* Dietary Fiber-Induced Improvement in Glucose Metabolism Is Associated with Increased Abundance of Prevotella. *Cell Metabolism* **22**, 971–982 (2015).

33. Zeevi, D. *et al.* Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079–1095 (2015).

34. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).

35. Maurice, C. F., Haiser, H. J. & Turnbaugh, P. J. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* **152**, 39–50 (2013).

36. Korpela, K. *et al.* Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nature communications* **7**, 10410 (2016).

37. Rutherford, S. T. & Bassler, B. L. Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control. *Cold Spring Harbor Perspectives in Medicine* **2**, a012427–a012427 (2012).

38. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–31 (2006).

39. Wang, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).

40. Vrieze, A. *et al.* Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity. *Journal of Hepatology* **60**, 824–831 (2014).

41. Reijnders, D. *et al.* Effects of Gut Microbiota Manipulation by Antibiotics on Host Metabolism in Obese Humans: A Randomized Double-Blind Placebo-Controlled Trial. *Cell Metabolism* **24**, 63–74 (2016).

42. Kootte, R. S. *et al.* Improvement of Insulin Sensitivity after Lean Donor Feces in Metabolic Syndrome Is Driven by Baseline Intestinal Microbiota Composition. *Cell metabolism* **26**, 611-619.e6 (2017).

43. Li, S. S. *et al.* Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science (New York, N.Y.)* **352**, 586–9 (2016).

44. Marques, R. E., Marques, P. E., Guabiraba, R. & Teixeira, M. M. Exploring the Homeostatic and Sensory Roles of the Immune System. *Frontiers in immunology* **7**, 125 (2016).

45. Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).

46. Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).

47. Wu, H. *et al.* Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nature Medicine* **23**, 850–858 (2017).

48. Freedberg, D. E. *et al.* Proton Pump Inhibitors Alter Specific Taxa in the Human Gastrointestinal Microbiome: A Crossover Trial. *Gastroenterology* **149**, 883–885 (2015).

49. Ridaura, V. K. *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214 (2013).

50. Bonder, M. J. *et al.* The effect of host genetics on the gut microbiome. *Nature Genetics* **48**, 1407–1412 (2016).

51. Turpin, W. *et al.* Association of host genome with intestinal microbial composition in a large healthy cohort. *Nature Genetics* **48**, 1413–1417 (2016).

52. Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. *Cell* **159**, 789–99 (2014).

53. Bashan, A. *et al.* Universality of human microbial dynamics. *Nature* **534**, 259–62 (2016).

54. Groen, A. K. & Nieuwdorp, M. An evaluation of the therapeutic potential of fecal microbiota transplantation to treat infectious and metabolic diseases. *EMBO molecular medicine* **9**, 1–3 (2017).

55. Clooney, A. G. *et al.* Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. *PLOS ONE* **11**, e0148028 (2016).

56. Tigchelaar, E. F. *et al.* Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ open* **5**, e006772 (2015).

57. Stronks, K. *et al.* Unravelling the impact of ethnicity on health in Europe: the HELIUS study. *BMC Public Health* **13**, 402 (2013).

58. Vrieze, A. *et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, (2012).

59. Herrema, H., IJzerman, R. G. & Nieuwdorp, M. Emerging role of intestinal microbiota and microbial metabolites in metabolic control. *Diabetologia* **60**, 613–617 (2017).

60. Nguyen, T. L. A., Vieira-Silva, S., Liston, A. & Raes, J. How informative is the mouse for human gut microbiota research? **8**, 1–16 (2015).

61. Treuting, P. M. & Dintzis, S. M. Lower Gastrointestinal Tract. in *Comparative Anatomy and Histology* 177–192 (Elsevier, 2012). doi:10.1016/B978-0-12-381361-9.00012-3.

62. Bäckhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15718–23 (2004).

63. Bäckhed, F., Manchester, J. K., Semenkovich, C. F. & Gordon, J. I. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 979–84 (2007).

64. Ridaura, V. K. *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214 (2013).

65. Wostmann, B. S., Larkin, C., Moriarty, A. & Bruckner-Kardoss, E. Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats. *Laboratory animal science* **33**, 46–50 (1983).

66. Tlaskalová-Hogenová, H. *et al.* The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: contribution of germ-free and gnotobiotic animal models of human diseases. *Cellular & molecular immunology* **8**, 110–20 (2011).

67. Dao, M. C. *et al.* Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. *Gut* **65**, 426–436 (2016).

68. Wang, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).

69. Ley, R., Turnbaugh, P., Klein, S. & Gordon, J. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–3 (2006).

70. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–31 (2006).

71. Duncan, S. H. *et al.* Human colonic microbiota associated with diet, obesity and weight loss. *International Journal of Obesity* **32**, 1720–1724 (2008).

72. Zhang, H. *et al.* Human gut microbiota in obesity and after gastric bypass. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 2365–70 (2009).

73. Hoyles, L. & McCartney, A. L. What do we mean when we refer to *Bacteroidetes* ???populations in the human gastrointestinal microbiota? *FEMS Microbiology Letters* **299**, 175–183 (2009).

74. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).

75. Finucane, M. M., Sharpton, T. J., Laurent, T. J., Pollard, K. S. & Zafar, N. A Taxonomic Signature of Obesity in the Microbiome? Getting to the Guts of the Matter. *PLoS ONE* **9**, e84689 (2014).

76. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).

77. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nature Reviews Immunology* **16**, 341–352 (2016).

78. Hotamisligil, G. S., Shargill, N. S. & Spiegelman, B. M. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. *Science* **259**, 87–91 (1993).

79. Hotamisligil, G. S., Shargill, N. S. & Spiegelman, B. M. Adipose expression of tumor necrosis factor-alpha: direct role in obesity-linked insulin resistance. *Science* **259**, 87–91 (1993).

80. Donath, M. Y. Targeting inflammation in the treatment of type 2 diabetes. *Diabetes Obes Metab* **15 Suppl 3**, 193–196 (2013).

81. Cani, P. D. *et al.* Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes* **56**, 1761–1772 (2007).

82. Cani, P. D. *et al.* Changes in Gut Microbiota Control Metabolic Endotoxemia-Induced Inflammation in High-Fat Diet-Induced Obesity and Diabetes in Mice. *Diabetes* **57**, 1470–1481 (2008).

83. Cani, P. D. *et al.* Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut* **58**, 1091–1103 (2009).

84. Pussinen, P. J., Havulinna, A. S., Lehto, M., Sundvall, J. & Salomaa, V. Endotoxemia Is Associated With an Increased Risk of Incident Diabetes. *Diabetes Care* **34**, 392–397 (2011).

85. Lassenius, M. I. *et al.* Bacterial Endotoxin Activity in Human Serum Is Associated With Dyslipidemia, Insulin Resistance, Obesity, and Chronic Inflammation. *Diabetes Care* **34**, 1809–1815 (2011).

86. Troseid, M. *et al.* Plasma Lipopolysaccharide Is Closely Associated With Glycemic Control and Abdominal Obesity: Evidence from bariatric surgery. *Diabetes Care* **36**, 3627–3632 (2013).

87. Glassing, A., Dowd, S. E., Galandiuk, S., Davis, B. & Chiodini, R. J. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathogens* **8**, 24 (2016).

88. Denou, E. *et al.* Defective NOD2 peptidoglycan sensing promotes diet-induced inflammation, dysbiosis, and insulin resistance. *EMBO molecular medicine* **7**, 259–74 (2015).

89. Amar, J. *et al.* Involvement of tissue bacteria in the onset of diabetes in humans: evidence for a concept. *Diabetologia* **54**, 3055–3061 (2011).

90. Tomasova, L., Konopelski, P. & Ufnal, M. Gut Bacteria and Hydrogen Sulfide: The New Old Players in Circulatory System Homeostasis. *Molecules* **21**, 1558 (2016).

91. Linden, D. R. Hydrogen sulfide signaling in the gastrointestinal tract. *Antioxidants & redox signaling* **20**, 818–30 (2014).

92. Hackfort, B. T. & Mishra, P. K. Emerging role of hydrogen sulfide-microRNA crosstalk in cardiovascular diseases. *American journal of physiology. Heart and circulatory physiology* **310**, H802-12 (2016).

93. Pitcher, M. C., Beatty, E. R. & Cummings, J. H. The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. *Gut* **46**, 64–72 (2000).

94. Wallace, J. L., Caliendo, G., Santagada, V., Cirino, G. & Fiorucci, S. Gastrointestinal Safety and Anti-Inflammatory Effects of a Hydrogen Sulfide–Releasing Diclofenac Derivative in the Rat. *Gastroenterology* **132**, 261–271 (2007).

95. Miao, L., Xin, X., Xin, H., Shen, X. & Zhu, Y.-Z. Hydrogen Sulfide Recruits Macrophage Migration by Integrin β1-Src-FAK/Pyk2-Rac Pathway in Myocardial Infarction. *Scientific Reports* **6**, 22363 (2016).

96. Pan, L.-L., Qin, M., Liu, X.-H. & Zhu, Y.-Z. The Role of Hydrogen Sulfide on Cardiovascular Homeostasis: An Overview with Update on Immunomodulation. *Frontiers in Pharmacology* **8**, 686 (2017).

97. Wang, X.-H. *et al.* Dysregulation of cystathionine γ-lyase (CSE)/hydrogen sulfide pathway contributes to ox-LDL-induced inflammation in macrophage. *Cellular Signalling* **25**, 2255–2262 (2013).

98. Donia, M. S. & Fischbach, M. A. Small molecules from the human microbiota. *Science* **349**, 1254766–1254766 (2015).

99. Meyer, T. W. & Hostetter, T. H. Uremic solutes from colon microbes. *Kidney international* **81**, 949–54 (2012).

100. Holmes, E., Li, J. V, Marchesi, J. R. & Nicholson, J. K. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab* **16**, 559–564 (2012).

101. Tang, W. H. W. & Hazen, S. L. Microbiome, trimethylamine N-oxide, and cardiometabolic disease. *Translational Research* (2016) doi:10.1016/j.trsl.2016.07.007.

102. Zhu, W. *et al.* Gut Microbial Metabolite TMAO Enhances Platelet Hyperreactivity and Thrombosis Risk. *Cell* **165**, 111–124 (2016).

103. Dambrova, M. *et al.* Diabetes is Associated with Higher Trimethylamine N-oxide Plasma Levels. *Experimental and Clinical Endocrinology & Diabetes* **124**, 251–256 (2016).

104. Tang, W. H. W. *et al.* Prognostic Value of Elevated Levels of Intestinal Microbe-Generated Metabolite Trimethylamine-N-Oxide in Patients With Heart Failure. *Journal of the American College of Cardiology* **64**, 1908–1914 (2014).

105. Lever, M. *et al.* Betaine and Trimethylamine-N-Oxide as Predictors of Cardiovascular Outcomes Show Different Patterns in Diabetes Mellitus: An Observational Study. *PLoS ONE* **9**, e114969 (2014).

106. Miao, J. *et al.* Flavin-containing monooxygenase 3 as a potential player in diabetes-associated atherosclerosis. *Nature Communications* **6**, 6498 (2015).

107. Zhu, W. *et al.* Gut Microbial Metabolite TMAO Enhances Platelet Hyperreactivity and Thrombosis Risk. *Cell* **165**, 111–124 (2016).

108. Lu, Y. *et al.* Short Chain Fatty Acids Prevent High-fat-diet-induced Obesity in Mice by Regulating G Protein-coupled Receptors and Gut Microbiota. *Scientific reports* **6**, 37589 (2016).

109. Perry, R. J. *et al.* Acetate mediates a microbiome-brain-beta-cell axis to promote metabolic syndrome. *Nature* **534**, 213–217 (2016).

110. Samuel, B. S. *et al.* Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. *Proc Natl Acad Sci U S A* **105**, 16767–16772 (2008).

111. Tolhurst, G. *et al.* Short-chain fatty acids stimulate glucagon-like peptide-1 secretion via the G-protein-coupled receptor FFAR2. *Diabetes* **61**, 364–371 (2012).

112. Fava, S. Glucagon-like peptide 1 and the cardiovascular system. *Current diabetes reviews* **10**, 302–10 (2014).

113. Thorburn, A. N. *et al.* Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites. *Nature communications* **6**, 7320 (2015).

114. Davie, J. R. Inhibition of histone deacetylase activity by butyrate. *The Journal of nutrition* **133**, 2485S-2493S (2003).

115. den Besten, G. *et al.* Gut-derived short-chain fatty acids are vividly assimilated into host carbohydrates and lipids. *Am J Physiol Gastrointest Liver Physiol* **305**, G900-10 (2013).

116. Willemsen, L. E. M., Koetsier, M. A., van Deventer, S. J. H. & van Tol, E. A. F. Short chain fatty acids stimulate epithelial mucin 2 expression through differential effects on prostaglandin E(1) and E(2) production by intestinal myofibroblasts. *Gut* **52**, 1442–7 (2003).

117. Gaudier, E. *et al.* Butyrate specifically modulates MUC gene expression in intestinal epithelial goblet cells deprived of glucose. *AJP: Gastrointestinal and Liver Physiology* **287**, G1168–G1174 (2004).

118. Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451–455 (2013).

119. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446–450 (2013).

120. Chang, P. V., Hao, L., Offermanns, S. & Medzhitov, R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proceedings of the National Academy of Sciences* **111**, 2247–2252 (2014).

121. Trompette, A. *et al.* Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature Medicine* **20**, 159–166 (2014).

122. De Vadder, F. *et al.* Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**, 84–96 (2014).

123.? stman, E., Granfeldt, Y., Persson, L. & Bj?rck, I. Vinegar supplementation lowers glucose and insulin responses and increases satiety after a bread meal in healthy subjects. *European Journal of Clinical Nutrition* **59**, 983–988 (2005).

124. Frost, G. *et al.* The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. *Nature Communications* **5**, 3611 (2014).

125. Chambers, E. S. *et al.* Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. *Gut* **64**, 1744–1754 (2015).

126. Yamaguchi, Y. *et al.* Association of Intestinal Microbiota with Metabolic Markers and Dietary Habits in Patients with Type 2 Diabetes. *Digestion* **94**, 66–72 (2016).

127. Schwiertz, A. *et al.* Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)* **18**, 190–195 (2010).

128. Schwiertz, A. *et al.* Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)* **18**, 190–195 (2010).

129. Vrieze, A. *et al.* Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity. *Journal of Hepatology* **60**, 824–831 (2014).

130. Vrieze, A. *et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, (2012).

131. Wahlstrom, A., Sayin, S. I., Marschall, H. U. & Backhed, F. Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism. *Cell Metabolism* vol. 24 41–50 Preprint at https://doi.org/10.1016/j.cmet.2016.05.005 (2016).

132. Sinal, C. J. *et al.* Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. *Cell* **102**, 731–744 (2000).

133. Russell, D. W. The enzymes, regulation, and genetics of bile acid synthesis. *Annual Review of Biochemistry* **72**, 137–174 (2003).

134. De Aguiar Vallim, T. Q., Tarling, E. J. & Edwards, P. A. Pleiotropic roles of bile acids in metabolism. *Cell Metabolism* vol. 17 657–669 Preprint at https://doi.org/10.1016/j.cmet.2013.03.013 (2013).

135. Doerner, K. C., Takamine, F., LaVoie, C. P., Mallonee, D. H. & Hylemon, P. B. Assessment of fecal bacteria with bile acid 7α-dehydroxylating activity for the presence of bai-like genes. *Applied and Environmental Microbiology* **63**, 1185–1188 (1997).

136. Ridlon, J. M., Kang, D.-J. & Hylemon, P. B. Bile salt biotransformations by human intestinal bacteria. *Journal of lipid research* **47**, 241–259 (2006).

137. Sayin, S. I. *et al.* Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metabolism* **17**, 225–235 (2013).

138. Watanabe, M. *et al.* Bile acids induce energy expenditure by promoting intracellular thyroid hormone activation. *Nature* **439**, 484–489 (2006).

139. Li, F. *et al.* ARTICLE Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. (2013) doi:10.1038/ncomms3384.

140. Watanabe, M. *et al.* Lowering Bile Acid Pool Size with a Synthetic Farnesoid X Receptor (FXR) Agonist Induces Obesity and Diabetes through Reduced Energy Expenditure. *Journal of Biological Chemistry* **286**, 26913–26920 (2011).

141. Thomas, C. *et al.* TGR5-Mediated Bile Acid Sensing Controls Glucose Homeostasis. *Cell Metabolism* **10**, 167–177 (2009).

142. Yoshimoto, S. *et al.* Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97–101 (2013).

143. Li, F. *et al.* Microbiome remodelling leads to inhibition of intestinal farnesoid X receptor signalling and decreased obesity. *Nat Commun* **4**, 2384 (2013).

144. Prinz, P. *et al.* Plasma bile acids show a positive correlation with body mass index and are negatively associated with cognitive restraint of eating in obese patients. *Frontiers in neuroscience* **9**, 199 (2015).

145. Vincent, R. P. *et al.* Higher circulating bile acid concentrations in obese patients with type 2 diabetes. *Annals of Clinical Biochemistry: An international journal of biochemistry and laboratory medicine* **50**, 360–364 (2013).

146. Adrian, T. E. *et al.* Rectal taurocholate increases L cell and insulin secretion, and decreases blood glucose and food intake in obese type 2 diabetic volunteers. *Diabetologia* **55**, 2343–2347 (2012).

147. Patti, M.-E. *et al.* Serum bile acids are higher in humans with prior gastric bypass: potential contribution to improved glucose and lipid metabolism. *Obesity (Silver Spring, Md.)* **17**, 1671–7 (2009).

148. Ahmad, N. N., Pfalzer, A. & Kaplan, L. M. Roux-en-Y gastric bypass normalizes the blunted postprandial bile acid excursion associated with obesity. *International journal of obesity (2005)* **37**, 1553–9 (2013).

149. Kohli, R. *et al.* Weight loss induced by Roux-en-Y gastric bypass but not laparoscopic adjustable gastric banding increases circulating bile acids. *Journal of Clinical Endocrinology and Metabolism* **98**, (2013).

150. Tremaroli, V. *et al.* Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation. *Cell Metabolism* **22**, 228–238 (2015).

151. Ryan, K. K. *et al.* FXR is a molecular target for the effects of vertical sleeve gastrectomy. *Nature* **509**, 183–188 (2014).

152. Miras, A. D. & le Roux, C. W. Mechanisms underlying weight loss after bariatric surgery. *Nature reviews. Gastroenterology & hepatology* **10**, 575–84 (2013).

153. Ryan, K. K. *et al.* FXR is a molecular target for the effects of vertical sleeve gastrectomy. *Nature* **509**, 183–188 (2014).

154. Flynn, C. R. *et al.* Bile diversion to the distal small intestine has comparable metabolic benefits to bariatric surgery. *Nature communications* **6**, 7715 (2015).

155. Schauer, P. R. *et al.* Bariatric Surgery versus Intensive Medical Therapy for Diabetes — 5-Year Outcomes. *New England Journal of Medicine* **376**, 641–651 (2017).

156. Buchwald, H. & Oien, D. M. Metabolic/bariatric surgery worldwide 2011. *Obesity Surgery* **23**, 427–436 (2013).

157. Pories, W. J. *et al.* Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus. *Annals of surgery* **222**, 339–50; discussion 350-2 (1995).

158. Reed, M. A. *et al.* Roux-en-Y Gastric Bypass Corrects Hyperinsulinemia Implications for the Remission of Type 2 Diabetes. *The Journal of Clinical Endocrinology & Metabolism* **96**, 2525–2531 (2011).

159. Jackness, C. *et al.* Very Low–Calorie Diet Mimics the Early Beneficial Effect of Roux-en-Y Gastric Bypass on Insulin Sensitivity and b-Cell Function in Type 2 Diabetic Patients. doi:10.2337/db12-1762.

160. Isbell, J. M. *et al.* The importance of caloric restriction in the early improvements in insulin sensitivity after Roux-en-Y gastric bypass surgery. *Diabetes care* **33**, 1438–42 (2010).

161. Aron-Wisnewsky, J. & Clement, K. The Effects of Gastrointestinal Surgery on Gut Microbiota: Potential Contribution to Improved Insulin Sensitivity. *Current Atherosclerosis Reports* **16**, 454 (2014).

162. Tremaroli, V. *et al.* Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation. *Cell Metabolism* **22**, 228–238 (2015).

163. Magouliotis, D. E., Tasiopoulou, V. S., Sioka, E., Chatedaki, C. & Zacharoulis, D. Impact of Bariatric Surgery on Metabolic and Gut Microbiota Profile: a Systematic Review and Meta-analysis. *Obesity Surgery* **27**, 1345–1357 (2017).

164. Liou, A. P. *et al.* Conserved shifts in the gut microbiota due to gastric bypass reduce host weight and adiposity. *Sci Transl Med* **5**, 178ra41 (2013).

165. Sjöström, L. *et al.* Effects of Bariatric Surgery on Mortality in Swedish Obese Subjects. *New England Journal of Medicine* **357**, 741–752 (2007).

166. Knip, M. & Siljander, H. The role of the intestinal microbiota in type 1 diabetes mellitus. *Nature Reviews Endocrinology* **12**, 154–167 (2016).

167. Paun, A., Yau, C. & Danska, J. S. The Influence of the Microbiome on Type 1 Diabetes. *The Journal of Immunology* **198**, 590–595 (2017).

168. Todd, J. A. Etiology of Type 1 Diabetes. *Immunity* **32**, 457–467 (2010).

169. Bluestone, J. A., Herold, K. & Eisenbarth, G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* **464**, 1293–300 (2010).

170. Eisenbarth, G. S. Update in Type 1 Diabetes. *The Journal of Clinical Endocrinology & Metabolism* **92**, 2403–2407 (2007).

171. Hattersley, A., Bruining, J., Shield, J., Njolstad, P. & Donaghue, K. C. The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatric Diabetes* **10**, 33–42 (2009).

172. Moltchanova, E. V., Schreier, N., Lammi, N. & Karvonen, M. Seasonal variation of diagnosis of Type 1 diabetes mellitus in children worldwide. *Diabetic Medicine* **26**, 673–678 (2009).

173. Kahn, H. S. *et al.* Association of Type 1 Diabetes With Month of Birth Among U.S. Youth: The SEARCH for Diabetes in Youth Study. *Diabetes Care* **32**, 2010–2015 (2009).

174. Blanton, D. *et al.* Reduced Serum Vitamin D-Binding Protein Levels Are Associated With Type 1 Diabetes. *Diabetes* **60**, 2566–2570 (2011).

175. Cooper, J. D. *et al.* Inherited variation in vitamin D genes is associated with predisposition to autoimmune disease type 1 diabetes. *Diabetes* **60**, 1624–31 (2011).

176. Knip, M., Virtanen, S. M. & Akerblom, H. K. Infant feeding and the risk of type 1 diabetes. *American Journal of Clinical Nutrition* **91**, 1506S-1513S (2010).

177. Bach, J.-F. & Chatenoud, L. The Hygiene Hypothesis: An Explanation for the Increased Frequency of Insulin-Dependent Diabetes. *Cold Spring Harbor Perspectives in Medicine* **2**, a007799–a007799 (2012).

178. Bach, J.-F. The Effect of Infections on Susceptibility to Autoimmune and Allergic Diseases. *New England Journal of Medicine* **347**, 911–920 (2002).

179. Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* **165**, 842–53 (2016).

180. Brugman, S. *et al.* Antibiotic treatment partially protects against type 1 diabetes in the Bio-Breeding diabetes-prone rat. Is the gut flora involved in the development of type 1 diabetes? *Diabetologia* **49**, 2105–2108 (2006).

181. Schwartz, R. F., Neu, J., Schatz, D., Atkinson, M. A. & Wasserfall, C. Comment on: Brugman S et al. (2006) Antibiotic treatment partially protects against type 1 diabetes in the Bio-Breeding diabetes-prone rat. Is the gut flora involved in the development of type 1 diabetes? Diabetologia 49:2105–2108. *Diabetologia* **50**, 220–221 (2006).

182. Wen, L. *et al.* Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* **455**, 1109–1113 (2008).

183. Smith, P. M. *et al.* The Microbial Metabolites, Short-Chain Fatty Acids, Regulate Colonic Treg Cell Homeostasis. *Science* **341**, 569–573 (2013).

184. Eisenstein, E. M. & Williams, C. B. The Treg/Th17 Cell Balance: A New Paradigm for Autoimmunity. *Pediatric Research* **65**, 26R-31R (2009).

185. Sakaguchi, S. Naturally arising Foxp3-expressing CD25+CD4+ regulatory T cells in immunological tolerance to self and non-self. *Nature Immunology* **6**, 345–352 (2005).

186. Turley, S. J., Lee, J.-W., Dutton-Swain, N., Mathis, D. & Benoist, C. Endocrine self and gut non-self intersect in the pancreatic lymph nodes. *Proceedings of the National Academy of Sciences* **102**, 17729–17733 (2005).

187. Blaschitz, C. & Raffatellu, M. Th17 Cytokines and the Gut Mucosal Barrier. *Journal of Clinical Immunology* **30**, 196–203 (2010).

188. Livanos, A. E. *et al.* Antibiotic-mediated gut microbiome perturbation accelerates development of type 1 diabetes in mice. *Nature Microbiology* **1**, 16140 (2016).

189. Candon, S. *et al.* Antibiotics in Early Life Alter the Gut Microbiome and Increase Disease Incidence in a Spontaneous Mouse Model of Autoimmune Insulin-Dependent Diabetes. *PLOS ONE* **10**, e0125448 (2015).

190. Bosi, E. *et al.* Increased intestinal permeability precedes clinical onset of type 1 diabetes. *Diabetologia* **49**, 2824–2827 (2006).

191. Mariño, E. *et al.* Gut microbial metabolites limit the frequency of autoimmune T cells and protect against type 1 diabetes. *Nature Immunology* **18**, 552–562 (2017).

192. Murri, M. *et al.* Gut microbiota in children with type 1 diabetes differs from that in healthy children: a case-control study. *BMC Medicine* **11**, 46 (2013).

193. Brown, C. T. *et al.* Gut Microbiome Metagenomics Analysis Suggests a Functional Model for the Development of Autoimmunity for Type 1 Diabetes. *PLoS ONE* **6**, e25792 (2011).

194. Gorus, F. K. *et al.* Twenty-Year Progression Rate to Clinical Onset According to Autoantibody Profile, Age, and *HLA-DQ* Genotype in a Registry-Based Group of Children and Adults With a First-Degree Relative With Type 1 Diabetes. *Diabetes Care* **40**, 1065–1072 (2017).

195. de Goffau, M. C. *et al.* Fecal Microbiota Composition Differs Between Children With -Cell Autoimmunity and Those Without. *Diabetes* **62**, 1238–1244 (2013).

196. Hummel, S., Pflüger, M., Hummel, M., Bonifacio, E. & Ziegler, A.-G. Primary Dietary Intervention Study to Reduce the Risk of Islet Autoimmunity in Children at Increased Risk for Type 1 Diabetes. *Diabetes Care* **34**, (2011).

197. Endesfelder, D. *et al.* Compromised Gut Microbiota Networks in Children With Anti-Islet Cell Autoimmunity. *Diabetes* **63**, 2006–2014 (2014).

198. Kostic, A. D. *et al.* The Dynamics of the Human Infant Gut Microbiome in Development and in Progression toward Type 1 Diabetes. *Cell Host & Microbe* **17**, 260–273 (2015).

199. O'Neil, D. A. *et al.* Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. *Journal of immunology (Baltimore, Md.: 1950)* **163**, 6718–24 (1999).

200. Wehkamp, J., Schmid, M., Fellermann, K. & Stange, E. F. Defensin deficiency, intestinal microbes, and the clinical phenotypes of Crohn's disease. *Journal of Leukocyte Biology* **77**, 460–465 (2004).

201. Maffeis, C. *et al.* Association between intestinal permeability and faecal microbiota composition in Italian children with beta cell autoimmunity at risk for type 1 diabetes. *Diabetes/Metabolism Research and Reviews* **32**, 700–709 (2016).

202. Craig, M. E. *et al.* Prevalence of Celiac Disease in 52,721 Youth With Type 1 Diabetes: International Comparison Across Three Continents. *Diabetes Care* **40**, 1034–1040 (2017).

203. Kemppainen, K. M. *et al.* Early Childhood Gut Microbiomes Show Strong Geographic Differences Among Subjects at High Risk for Type 1 Diabetes: Figure 1. *Diabetes Care* **38**, 329–332 (2015).

204. Forslund, K. *et al.* Country-specific antibiotic use practices impact the human gut resistome. *Genome Research* **23**, 1163–1169 (2013).

205. Chassaing, B. *et al.* Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* **519**, 92–96 (2015).

206. Suez, J. *et al.* Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* **514**, 181–186 (2014).

207. Tang, W. H. W., Kitai, T. & Hazen, S. L. Gut Microbiota in Cardiovascular Health and Disease. *Circulation Research* **120**, 1183–1196 (2017).

208. Ojetti, V. *et al.* Rifaximin pharmacology and clinical implications. *Expert opinion on drug metabolism & toxicology* **5**, 675–82 (2009).

209. Dao, M. C. *et al.* Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. *Gut* **65**, 426–436 (2016).

210. Park, D.-Y. *et al.* Supplementation of Lactobacillus curvatus HY7601 and Lactobacillus plantarum KY1032 in Diet-Induced Obese Mice Is Associated with Gut Microbial Changes and Reduction in Obesity. *PLoS ONE* **8**, e59470 (2013).

211. Simon, M. C. *et al.* Intake of lactobacillus reuteri improves incretin and insulin secretion in Glucose-Tolerant humans: A proof of concept. *Diabetes Care* **38**, 1827–1834 (2015).

212. Kadooka, Y. *et al.* Regulation of abdominal adiposity by probiotics (Lactobacillus gasseri SBT2055) in adults with obese tendencies in a randomized controlled trial. *European Journal of Clinical Nutrition* **64**, 636–643 (2010).

213. Robertson, M. D., Bickerton, A. S., Dennis, A. L., Vidal, H. & Frayn, K. N. Insulin-sensitizing effects of dietary resistant starch and effects on skeletal muscle and adipose tissue metabolism. *American Journal of Clinical Nutrition* **82**, 559–567 (2005).

214. Cani, P. D. *et al.* Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* **50**, 2374–2383 (2007).

215. Delmee, E. *et al.* Relation between colonic proglucagon expression and metabolic response to oligofructose in high fat diet-fed mice. *Life Sciences* **79**, 1007–1013 (2006).

216. Parnell, J. A. & Reimer, R. A. Weight loss during oligofructose supplementation is associated with decreased ghrelin and increased peptide YY in overweight and obese adults. *American Journal of Clinical Nutrition* **89**, 1751–1759 (2009).

217. Maldonado-G?mez, M. *et al.* Stable Engraftment of Bifidobacterium longum AH1206 in the Human Gut Depends on Individualized Features of the Resident Microbiome. *Cell Host & Microbe* **20**, 515–526 (2016).

218. Zhang, F., Luo, W., Shi, Y., Fan, Z. & Ji, G. Should we standardize the 1,700-year-old fecal microbiota transplantation? *The American journal of gastroenterology* **107**, 1755; author reply p.1755-6 (2012).

219. Eiseman, B., Silen, W., Bascom, G. & Kauvar, A. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery* **44**, 854–859 (1958).

220. van Nood, E. *et al.* Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med* **368**, 407–415 (2013).

221. Lopez, J. & Grinspan, A. Fecal Microbiota Transplantation for Inflammatory Bowel Disease. *Gastroenterology & hepatology* **12**, 374–9 (2016).

222. Evrensel, A. & Ceylan, M. E. Fecal Microbiota Transplantation and Its Usage in Neuropsychiatric Disorders. *Clinical psychopharmacology and neuroscience : the official scientific journal of the Korean College of Neuropsychopharmacology* **14**, 231–7 (2016).

223. Finegold, S. M. *et al.* Pyrosequencing study of fecal microflora of autistic and control children. *Anaerobe* **16**, 444–453 (2010).

224. Thaiss, C. A. *et al.* Microbiota Diurnal Rhythmicity Programs Host Transcriptome Oscillations. *Cell* **167**, 1495-1510.e12 (2016).

225. Liu, S. *et al.* The Host Shapes the Gut Microbiota via Fecal MicroRNA. *Cell Host and Microbe* **19**, 32–43 (2016).

226. Reyes, A. *et al.* Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. *Proceedings of the National Academy of Sciences* **112**, 201514285 (2015).

227. Focà, A. *et al.* Gut inflammation and immunity: What is the role of the human gut virome? *Mediators of Inflammation* vol. 2015 Preprint at https://doi.org/10.1155/2015/326032 (2015).

228. Reyes, A. *et al.* Viruses in the fecal microbiota of monozygotic twins and their mothers. *Nature* **466**, 334–338 (2010).

229. Arnold, J. W., Roach, J. & Azcarate-Peril, M. A. Emerging Technologies for Gut Microbiome Research. *Trends in Microbiology* **24**, 887–901 (2016).

230. Ranjan, R., Rani, A., Metwally, A., McGee, H. S. & Perkins, D. L. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and biophysical research communications* **469**, 967–77 (2016).

231. Consortium, H. M. P. *et al.* A framework for human microbiome research. *Nature* **486**, 215–221 (2012).

232. Franzosa, E. A. *et al.* Relating the metatranscriptome and metagenome of the human gut. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2329-38 (2014).

233. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* **13**, 227–232 (2012).

234. Maier, T. *et al.* Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Molecular systems biology* **7**, 511 (2011).

235. Selinger, D. W., Saxena, R. M., Cheung, K. J., Church, G. M. & Rosenow, C. Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. *Genome Research* **13**, 216–223 (2003).

236. Kolmeder, C. A. *et al.* Colonic metaproteomic signatures of active bacteria and the host in obesity. *Proteomics* **15**, 3544–3552 (2015).

237. Newgard, C. B. *et al.* Metabolomics and Metabolic Diseases: Where Do We Stand? *Cell Metabolism* **25**, 43–56 (2017).

238. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the Metabolome: Size, Flux, and Regulation. *Molecular Cell* **58**, 699–706 (2015).

239. Zheng, X. *et al.* A targeted metabolomic protocol for short-chain fatty acids and branched-chain amino acids. *Metabolomics* **9**, 818–827 (2013).

240. Tigchelaar, E. F. *et al.* Gut microbiota composition associated with stool consistency. *Gut* **65**, 540–2 (2016).

241. He, X., Ji, G., Jia, W., Li, H. & Han, T.-L. Gut Microbiota and Nonalcoholic Fatty Liver Disease: Insights on Mechanism and Application of Metabolomics. 1–808 doi:10.3390/ijms17030300.

3

Distinct differences in gut microbial composition and functional potential from lean to morbidly obese subjects

Abraham S Meijnikman, Omrum Aydin, Andrei Prodan, Valentina Tremaroli, Hilde Herrema, Evgeni Levin, Yair Acherman, Sjoerd Bruin, Victor E Gerdes, Fredrik Backhed, Albert K. Groen, Max Nieuwdorp

Journal of Internal Medicine, 2020 Dec;288(6):699-710.

ABSTRACT

Introduction

The gut microbiome may contribute to the development of obesity. So far, the extent of microbiome variation in people with obesity has not been determined in large cohorts and for a wide range of Body Mass Index (BMI). Here we aimed to investigate whether the fecal microbial metagenome can explain the variance in several clinical phenotypes associated with morbid obesity.

Methods

Caucasian subjects were recruited at our hospital. Blood pressure and anthropometric measurements were taken. Dietary intake was determined using questionnaires. Shotgun metagenomic sequencing was performed on fecal samples from 177 subjects.

Results

Subjects without obesity (n=82, BMI 24.7±2.9 kg/m2) and subjects with obesity (n=95, BMI 38.6±5.1 kg/m2) could be clearly distinguished based on fecal microbiota composition and microbial metabolic pathways. A total number of 52 bacterial species differed significantly in people with and without obesity. Independent of dietary intake, we found that microbial pathways involved in biosynthesis of amino acids were enriched in subjects with obesity, whereas pathways involved in the degradation of amino acids were depleted. Machine learning models showed that about half of the variance in BMI could be explained by the gut microbiome composition and microbial metabolic pathways, compared to 6% of variation explained in triglycerides and 9% in HDL.

Conclusion

Based on the fecal microbiota composition, we were able to separate subjects with and without obesity. In addition, we found strong associations between gut microbial amino acid metabolism and specific microbial species in relation to clinical features of obesity.

INTROUCTION

Obesity has become one of the most common and costly chronic disorders worldwide. Estimates suggest that 44% of the global population is overweight and more than 300 million individuals are affected by morbid obesity, defined as body mass index (BMI) over 40 kg/m^2 ref [1]. Type 2 diabetes (T2D) and Non-Alcoholic Fatty Liver Disease (NAFLD) are driven by the global obesity epidemic and their prevalence is proportional to the increase in BMI[1]. Lack of effective options for long-term weight reduction magnifies the burden of obesity, making it one of the greatest public health threats of the 21st century.

The pathogenesis of obesity may seem an issue of energy imbalance: calories consumed more than ongoing energy expenditure. However, emerging evidence suggests that the mechanism is more complex than passive accumulation of excess calories. In recent years, there have been substantial advances in the understanding of the development of obesity including genetics variants, epigenetic modifications, hormonal homeostasis, and environmental factors[3]. The gut microbiota has recently been identified as a potential contributing factor in the development of obesity and related morbidities[4,5]. Studies in humans have associated gut microbial alterations with moderate obesity and found lower bacterial diversity and gene richness in subjects with moderate obesity[6,7,8]. In an apparently healthy, large population cohort, it was shown that the gut microbiome can explain approximately 25% of the variance in BMI. Moreover, models using fecal metagenomics data showed superior performance in predicting several human phenotypes, such as BMI and plasma lipid levels, compared to models using host genetic or environmental data[9]. However, whether these results can be translated to a population with a wide variety in BMI (i.e., from lean to morbid obesity) remains to be investigated.

We therefore established a cohort of individuals with a broad range of BMI (from lean to morbidly obese) to determine the variation in gut microbiota composition and functional potential as a function of BMI. Shotgun metagenomic sequencing of fecal samples from 177 subjects and state of the art machine learning models were used to identify novel metabolic pathways and link these pathways to specific microbial species. In addition, we developed computational models using gut microbiome data to explain the variance in clinical human phenotypes such as BMI, lipid levels, and glucose. We show that almost 50% of the variance in BMI can be explained by the fecal metagenome.

METHODS

Subject recruitment and sample collection

Caucasian subjects with a wide variety in BMI (18.6 – 60.9 kg/m^2) were recruited at our hospital. Anthropometric measurements including height, weight, waist, and hip circumference were taken. In addition, body fat percentage using bioelectrical impedance and blood pressure were measured. Blood samples were taken after an overnight fast for hemoglobin, HbA1c, glucose, lipid profile, and creatinine levels. Patients with known diabetes were excluded to avoid confounding effects of long-term hyperglycemia or medication use (*e.g.,* metformin) on the gut microbiome. In addition, individuals on lipid-lowering medication were also excluded. In total, 177 subjects were included (see Table 1 for baseline characteristics). The study was executed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC. All participants provided written informed consent.

Stool sample collection

In brief, subjects were given a stool collection tube and a safety bag for transport. They were asked to bring a fresh stool sample to the research unit within 6 hours after production and collection. If this was not possible, subjects were instructed to keep the stool sample in their -20°C freezer overnight and to bring it on frozen icepacks to the research unit the next morning. Stool samples were stored at -80°C. Information regarding the use of probiotics, antibiotic use in the past three months and had diarrhea was registered. All participants filled out an online nutritional diary (https://mijn.voedingscentrum.nl/nl/eetmeter) to monitor daily caloric intake during a week and average carbohydrates, fat, protein and fibers amount was calculated.

Profiling of fecal microbiome composition by shotgun sequencing

Fecal genomic DNA was isolated from 100 mg of feces by 6 cycles of bead beating using a modification of the IHMS DNA extraction protocol Q (REF: Costea PI, Nat Biotechnol 2017: https://www.ncbi.nlm.nih.gov/pubmed/28967887). Briefly, fecal samples were extracted in Lysing Matrix E tubes (MP Biomedicals) containing ASL buffer (Qiagen). Cell lysis was induced by: 1) homogenization by vortexing for 2 minutes, followed by 2) heating at 90C for 10 minutes, and 3) three cycles of bead beating at 5.5 m/s for 60 s in a FastPrep®-24 Instrument (MP Biomedicals). After a bead-beating cycle, samples were placed on ice for 5 minutes. The supernatant was collected after the last bead beating cycle following centrifugation at 4C. Steps 2) and 3) described above were repeated one more time after addition of fresh ASL buffer. Supernatants from the two centrifugations steps were pooled and a 600 µL aliquot

from each sample was purified using the QIAamp DNA Mini kit (QIAGEN) in the QIAcube (QIAGEN) instrument using the procedure for human DNA analysis. Samples were eluted in 200 µL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0). Library preparation and sequencing was performed by Novogene. Extracted DNA quality was checked by agarose gel electrophoresis and Nanodrop (OD260/OD280). DNA was quantified using a Qubit 2.0. After fragmentation and index-containing adaptor ligation, samples were paired-end sequenced on an Illumina HiSeq (2 x 150 bp). Raw reads were checked using FastQC (v.0.11.8) and quality filtered using Trimmomatic (v.0.38). Trimmomatic was used to remove adaptors, apply sliding window quality trimming (window width 4 bp, threshold Q-score 15), and remove reads that, after trimming and adaptor removal, were shorter than 70 bp. Paired-end reads that passed the quality filtering were then mapped against the human genome (GRch37_hg19) using bowtie2 (v.2.3.4.3). Samtools (v.0.1.19), sambamba (v.0.6.8), and bedtools (v.2.26.0) were used to remove the reads that successfully mapped to the human genome. The remaining high quality, non-human reads were subsampled to 20 million paired-end reads per sample (using seqtk v.1.3) and fed to a humann2 pipeline[10] (v.0.11.2). For each sample, species-level microbial composition was inferred using MetaPhlaAn2[11] (v.2.7.7). After mapping the reads (using bowtie2) against the pangenomes selected based on inferred composition, unmapped reads were translated and mapped against the full Uniref90 protein database (using diamond v.0.8.38). MetaCyc pathway community-level abundance was normalized to copies per million (CPM).

Statistical analyses

Machine Learning (ML) models were constructed using gradient boosted trees[12]. To control overfitting and random sampling bias, a nested cross-validation structure was implemented in all models. For each model iteration, subjects were randomly split into a training set (80% of subjects) and a test set (20% of subjects). Hyperparameters were optimized within the training set using 5-fold cross-validation. The resulting optimized model was then used to predict the outcome on the test set, such that no models were ever trained and tested on the same data. This procedure was repeated 50 times per analysis, with different random test/train set splits. The performance of the model on the test set (i.e., R2 / explained variance for regression tasks; area under the ROC curve (AUC) for classification tasks) was recorded for each iteration. The final performance of the model was calculated as the average of the 50 iterations. Each analysis also produced a relative feature importance list, ranking the input variables by their importance for predicting the outcome. Two random variables were created and added as predictors in each analysis to serve as benchmarks (i.e., any predictor that is only as good as a random noise is of no interest). The ML workflows were implemented in Python (v.3.7.2), using the scikit-learn (v.0.20.3), numpy (v.1.16.2), pandas (v.0.24.2),

matplotlib (v.3.0.3), and xgboost (v.0.82) libraries and executed on a Linux workstation running Ubuntu 18.02.LTS. Statistical analysis was performed in R (v.3.5.2). Beta-diversity analysis was conducted using the vegan package to calculate Bray-Curtis distances and the 'pcoa' function of the ape package to perform Principal Coordinate Analysis (PCoA). The ggplot2 package was used to produce plots and visualizations. Correction of p-values to control the false discovery rate and account for multiple testing was done using the Benjamini-Hochberg method[13] for the Mann-Whitney U tests and the Bonferroni method for the microbe-pathways associations (Spearman's rho tests).

RESULTS

Gut microbiome phenotype for obesity

We analyzed intestinal microbiota composition of 95 obese subjects (BMI 38.6±5.1 kg/m^2, age 47.1±10.8 years) and 82 subjects without obesity (BMI 24.7±2.9 kg/m^2, age 53.2±2.1 years). Principal Coordinate Analysis (PCoA) of between-sample Bray-Curtis distances based on the fecal microbiome composition at species level (Figure 1A) and on the fecal microbial metabolic pathways (Figure 1B) revealed that subjects grouped together based on their obesity status (cutoff BMI either higher or lower than 30 kg/m^2). There was a small but significant negative association between microbial species alpha diversity (assessed by the Shannon index), and BMI (r=-0.181, p=0.0154), as well as a significant difference in the Shannon index between subjects with and without obesity (t-test p=0.0092), with obese individuals showing lower microbial diversity (Supplementary Figure 1). A total number of 52 bacterial species (out of the 196 species with mean relative abundance higher than 0.01%) differed significantly between subjects with and without obesity (Mann-Whitney U tests, after adjusting for multiple comparisons) (supplementary Table 1).

Next, we determined whether fecal microbiota composition could predict obesity status. As shown by the receiver operating characteristic (ROC) curve, we were able to predict obesity status using microbial composition with an AUC of 0.82 (Figure 2A). The most predictive bacterial species were *Actinomyces odontolyticus*, *Streptococcus australis*, *Streptococcus thermophilus*, *Collinsella aerofaciens,* and *Ruminococcus torques* (Figure 2B and 2C). The species enriched in subjects without obesity were *Alistipes shahii*, *Alistipes senegalensis*, *Butyrivibrio crossotus*, *Coprococcus eutactus, Oxalobacter formigenes,* and several Bacteroidales species (Figure 2B and 2C).

Variable	All	Non-Obese	Obese	p-value
BMI (kg/m^2)	32.2 (8.16)	24.7 (2.95)	38.6 (5.17)	5,60E-50
Age (years)	50.9 (9.8)	55.2 (6.3)	47.2 (10.8)	1,32E-08
Waist (cm)]	103.9 (17.9)	91.1 (11.7)	118.3 (11.8)	6,20E-30
Waist Hip Ratio	0.917 (0.09)	0.908 (0.09)	0.929 (0.09)	1,73E-01
HbA1c (mmol/mol)	37.6 (5.5)	37.2 (5.5)	38.0 (5.6)	3,35E-01
Glucose (mmol/l)	5.57 (0.82)	5.38 (0.70)	5.73 (0.89)	4,10E-03
Triglycerides (mmol/l)	1.38 (1.02)	1.05 (0.76)	1.67 (1.12)	3,50E-07
LDL-Cholesterol (mmol/l)	3.19 (0.94)	3.20 (0.79)	3.18 (1.06)	8,93E-01
HDL-Cholesterol (mmol/l)	1.39 (0.42)	1.57 (0.43)	1.22 (0.34)	3,81E-08
Smoking (n)	6	0	6	
Dietary intake				
Energy (kcal/day)	1767 (1189)	2237 (676)	1550 (1311)	9,32E-04
Protein (g/day)	84.5 (134.3)	82.9 (23.8)	85.2 (162)	9,12E-01
Carbohydrates (g/day)	175.3 (78.8)	233.9 (87.6)	148.1 (57.2)	1,12E-05
Fats (g/day)	46.7 (60.1)	30.6 (24.1)	54.2 (69.7)	1,56E-02
Fiber (g/day)	15.7 (8.5)	22.2 (6.3)	12.8 (7.7)	1,866E-08

Table 1. Baseline characteristics of 177 subjects and comparison between 95 subjects with obesity versus 82 subjects without obesity. Data is expressed as mean (SD). LDL: low density lipoprotein, HDL: high density lipoprotein.

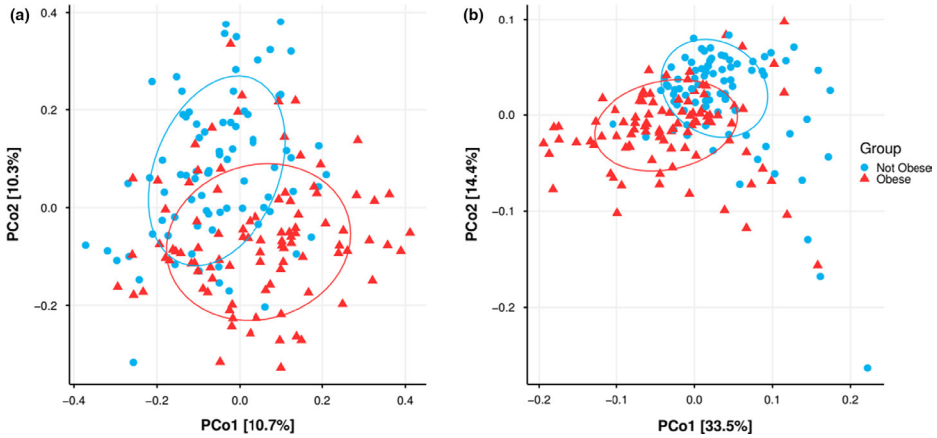


Figure 1. PCoA plots based on between-subjects Bray-Curtis distance of A) microbial composition at species level, and **B)** microbial functional pathways. Subjects are colored according to obesity status. Ellipses define an area 1 standard deviation away from the group centroids. The percent of variance explained by each principal coordinate is shown between brackets.

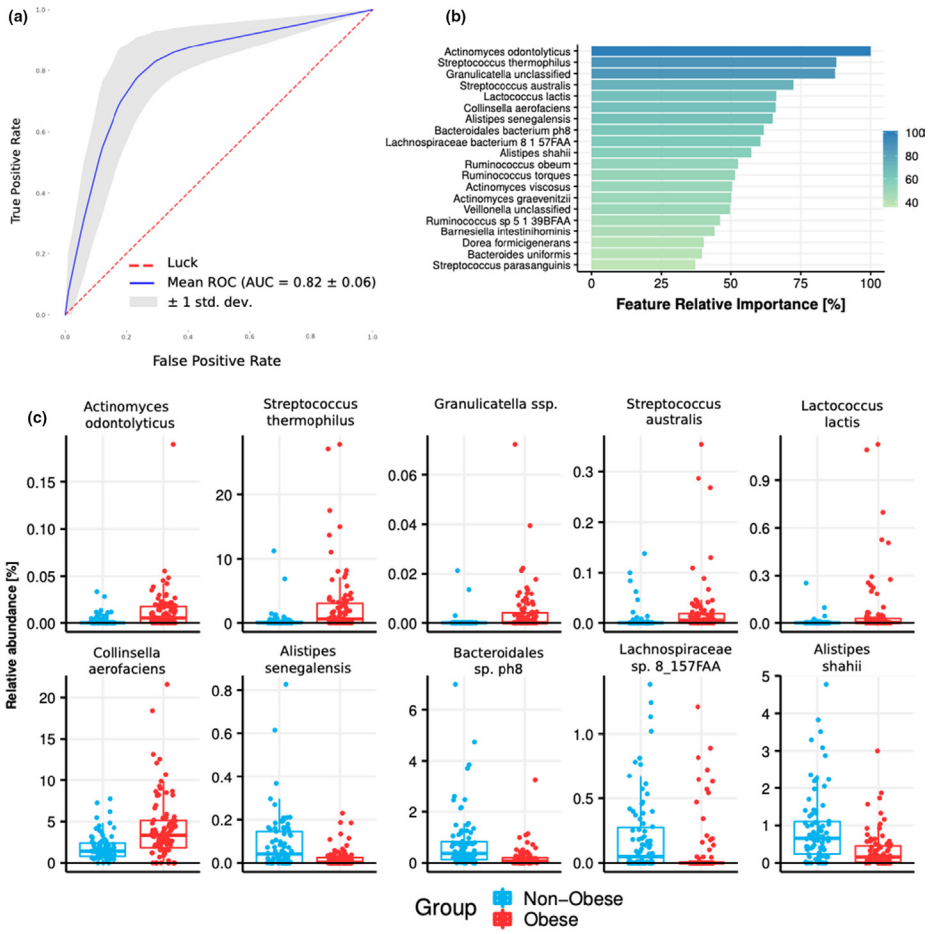


Figure 2. Results of ML models differentiating obese from non-obese individuals based on microbial composition. **A)** ROC-AUC; **B)** top 20 most predictive species; and **C)** boxplots stratified by obesity status showing relative abundance [%] of the top 10 microbial species predictors of obesity.

Since there were many significant differences in gut microbial species between subjects with a BMI over than 30 kg/m^2 compared to subjects with a BMI lower than 30 kg/m^2, we next investigated whether these significant differences could also be observed when we compared subjects with an intermediate BMI (28-35 kg/m^2) with subjects with a high BMI (>35 kg/m^2). Using the Mann Whitney U tests (p-values adjusted for multiple comparisons using Benjamini-Hochberg FDR corrections), we observed that 15 microbial species were significantly different between the 35 subjects with intermediate BMI and the 75 subjects with high BMI. Among the 15 microbial species were 5 of the top 6 most predictive species in the obese (BMI > 30) versus non-obese (BMI < 30) classification models, including all the

top 3 (*A. odontolyticus*, *S. thermophilus*, *Granulicatella* unclassified, *Lactococcus lactis*, and C. *aerofaciens*) (supplementary Table 2).

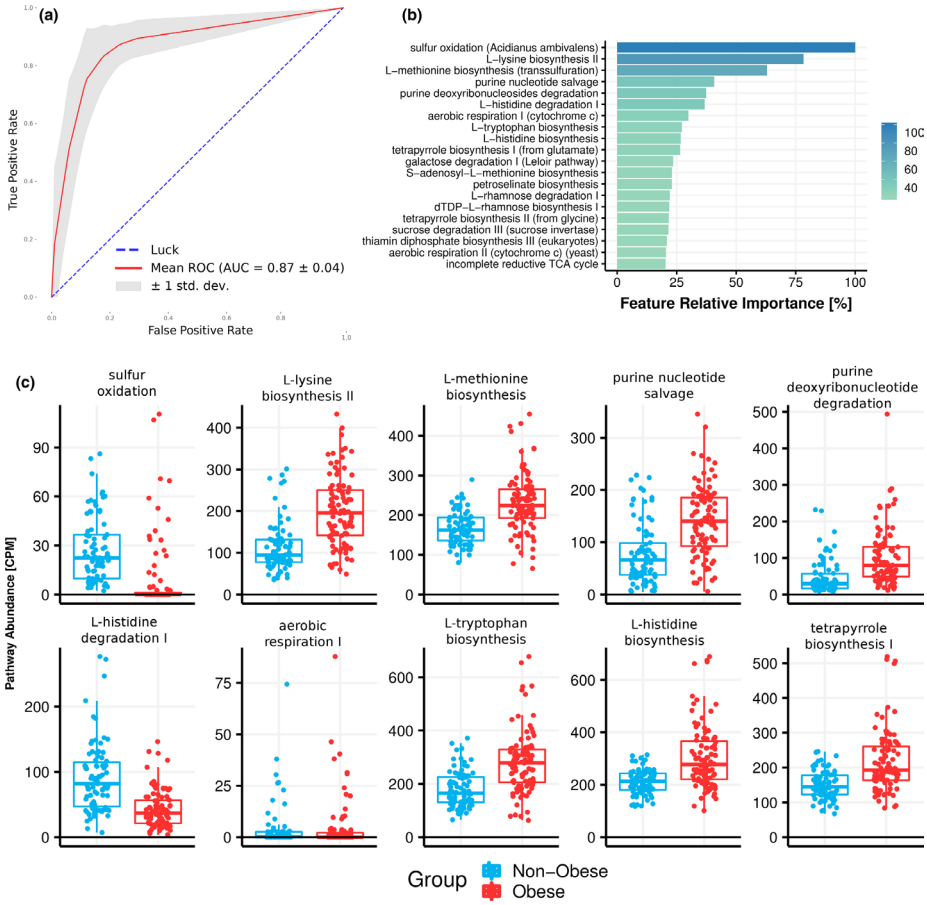


Figure 3. Results of ML models differentiating obese from non-obese individuals based on microbial metagenomic functional pathways. A) ROC-AUC; **B)** top 20 most predictive species; and **C)** boxplots stratified by obesity status showing relative abundance [counts per million - CPM] of the top 10 microbial functional pathways predictors of obesity.

Metabolic pathways associated with the obese and non-obese microbiome

We next investigated the role of fecal microbial metabolic pathways for differentiating subjects with and without obesity and observed a ROC AUC of 0.87 for this classification (Figure 3A). We found that metabolic pathways including the superpathway of sulfur oxidation, L-lysine biosynthesis, superpathway of L-methionine biosynthesis, the superpathway of purine nucleotide salvage, and L-histidine degradation were the best predictors of obesity status

(Figure 3B and 3C). The superpathway of sulfur oxidation was depleted in subjects with obesity compared to subjects without obesity (Fig 3C). In addition, histidine degradation and tRNA charging pathways were depleted in obese subjects compared to subjects without obesity (Figure 3C). Also, several amino acid biosynthesis pathways, including those for L-lysine, L-histidine, L-tryptophan, L-arginine, L-methionine, S-adenosyl-methionine and homoserine, were highly enriched in subjects with obesity (Figure 3C). Furthermore, metabolic pathways involved in galactose metabolism and purine degradation were also enriched in subjects with obesity compared to subjects without obesity. Correlation analysis revealed multiple significant correlations between metabolic pathways and BMI (Figure 4). We found that metabolic pathways including the superpathway of sulfur oxidation, L-lysine biosynthesis, superpathway of L-methionine biosynthesis, the superpathway of purine nucleotide salvage and L-histidine degradation were the best predictors of obesity status (Figure 3B).

Metabolic pathways association with gut microbial species

By using a machine learning approach, we next investigated to what extent the alterations in the gut microbiome in subjects with obesity were linked to fecal metabolic pathways. After adjustment for multiple comparisons (Bonferroni), several significant associations still remained (Figure 4), including positive correlations between *R. torques, Ruminococcus obeum*, and *C. aerofaciens* with microbial pathways involved in carbohydrate metabolism for sucrose, rhamnose, and galactose metabolism, as well as negative correlations between *A. shahii* and *A. senegalensis* with these pathways (Fig 4). In addition, *R. torques* and *S. thermophilius* were negatively correlated with histidine degradation, whereas *A. shahii, A. senegalensis, Bacteroidales bacterium,* and *Eubacterium eligens* were positively correlated with this pathway. In addition, the histidine biosynthesis pathway was negatively correlated with *A. shahii* and *E. eligens*, whereas a strong correlation between this pathway and *C. aerofaciens,* L. *lactis,* and *S. thermophilius* was observed. Comparing the gut microbial metabolic pathways between subjects with an intermediate BMI (28-35 kg/m^2) and subjects with a high BMI (>35 kg/m^2) revealed that 131 metabolic pathways were significantly different. Of the 131 microbial pathways significantly different between subjects with low vs intermediate BMI, 96 were also among the pathways significantly different between obese vs non-obese (after p-value adjustment for multiple comparisons), including almost all of the top 20 most predictive microbial pathways in the obese (BMI > 30 kg/m^2) versus non-obese (BMI < 30 kg/m^2) classification models (supplementary Table 3).

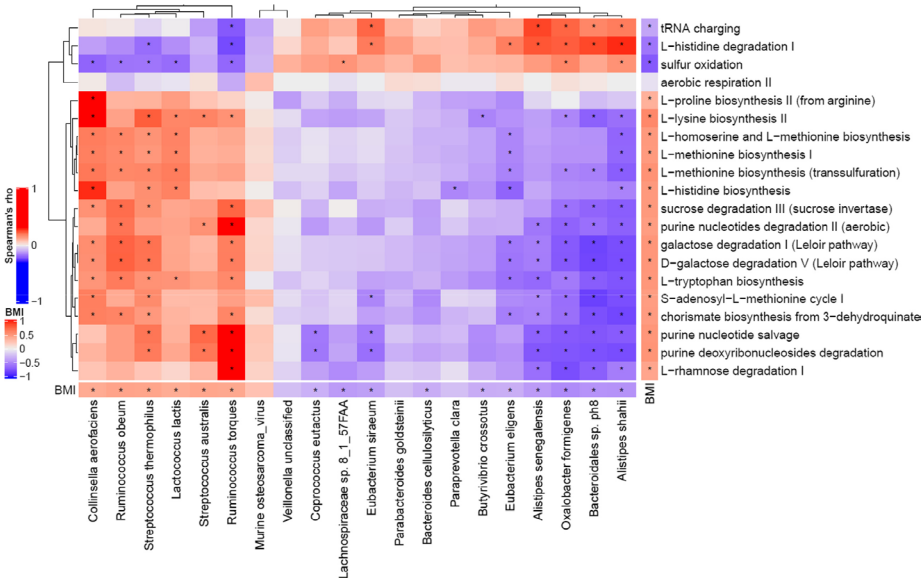


Figure 4. Correlation (Spearman's rho) between the top 20 microbial and the top 20 microbial functional pathways that were the best predictors of obesity status. Bottom and right-side annotations show the correlations between these microbes/pathways and BMI. All p-values are adjusted for multiple comparisons using the Bonferroni correction.

Microbiome-phenotypes associations

Recently, it has been shown that the gut microbiome can be associated with host characteristics using metagenomics data obtained in as little as a hundred subjects[14]. In the present study, machine learning models revealed that 45.5% of variance in BMI can be explained by the composition of the gut microbiota and 60% by gut microbial metabolic pathways (Figure 5A, 5B). Moreover, the gut microbiome and gut microbial metabolic pathways could explain the percentage body (measured by BIA), for respectively 48.5% and 65.5%. The top three bacteria that were positively associated with BMI and body fat percentage were *Ruminococcus* species (rho=0.37, p<0.001), *C. aerofaciens* (rho=0.35, p<0.001) and *Dorea formicigenerans* (rho = 0.33, p<0.001). Among the top three bacteria that were negatively associated with BMI were *A. shahii* (rho=-0.35, p<0.001), *A. senegalensis* (rho=-0.35, p<0.001) and *B. bacterium* (rho=-0.36, p<0.001). In addition, the top three pathways that were positively associated with BMI included L-histidine biosynthesis (rho=0.45, p<0.001), L-lysine biosynthesis (rho=0.5, p<0.001) and galactose degradation (rho=0.47, p<0.001). Pathways that were negatively correlated with BMI and body fat percentage were: L-histidine degradation (rho=-0.51, p<0.001) and tRNA charging (rho=-0.3, p<0.001). We next investigated whether the variance in frequently used clinical phenotypes could be explained by the gut microbiome composition and microbial

pathways. In addition, nearly 50% of the variance in waist circumference could be explained by the gut microbiome whereas the variance in triglycerides and high-density lipoprotein (HDL) could be explained for respectively 6% and 9% (Figure 5). In contrast, the variance in low-density lipoprotein, ycated hemoglobin and fasting glucose could not be explained by the gut microbiome.

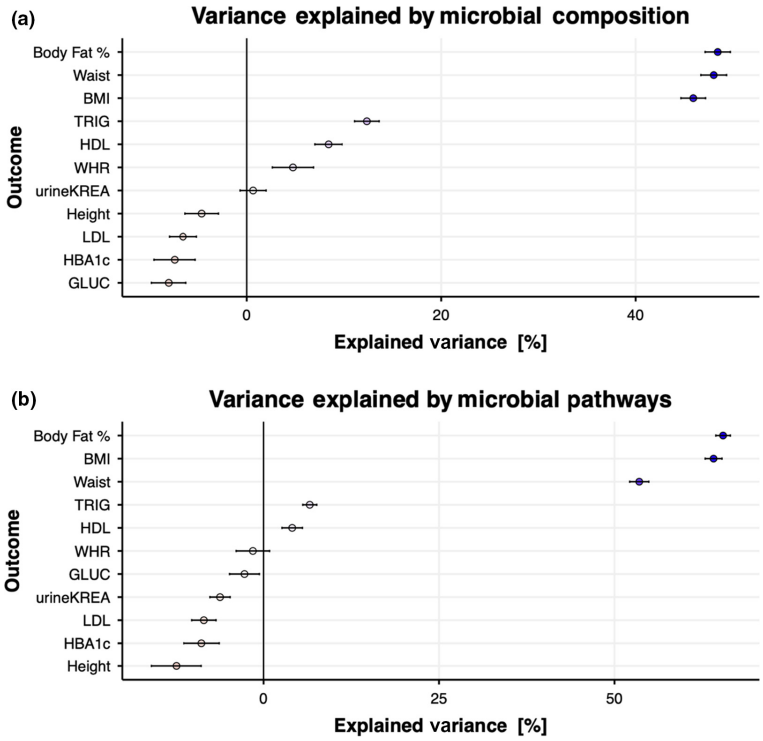


Figure 5. Explained variance by nested cross-validation ML models in human clinical phenotypes by A) microbes; and **B)** microbial functional pathways.

DISCUSSION

In this study, we revealed microbial metagenomic signatures of obesity by using shotgun metagenomics for subjects with a wide variety in BMI (i.e., from lean to morbid obesity). We could accurately separate subjects with and without obesity based on fecal metagenomics data. In total, 52 gut microbial species differed significantly between subjects with and without obesity, which enabled a separation of the groups with an accuracy of 0.82. In line with previous studies [8,15,16], species belonging to the phylum Bacteroidetes (*Alistipes shahii*

and *Alistipess senegalensis*) were depleted and species belonging to the phylum Firmicutes (*Ruminococcus torques, Ruminococcus obeum, and Dorea tormicigenerans*) were enriched in subjects with obesity compared to subjects without obesity. In contrast, one of the first papers on microbiota in severe obesity showed that a cluster of microbial species including *Lachnospiraceae*, *Ruminococcaeceae*, and other Firmicutes indeed decreased with increasing trunk fat, whereas *Ruminococcus* was negatively associated with BMI[17]. Bearing in mind that these bacterial species have previously been linked to moderate and even morbid obesity[8], our results indicate a possible robust consensus signature for the obese microbiome.

Despite contrasting observations at the species level, a Firmicutes-dominated microbiome has been linked to increased genetic potential for nutrient transporters[18]. In addition, this has been associated with increased expression of enzymes involved in complex carbohydrate degradation and fermentation[19]. The latter has been independently linked to increased energy harvest from diet[20]. Consistent with previous reports[8,17], we found that metabolic pathways involved in carbohydrate metabolism, including sucrose, galactose, and rhamnose metabolism, were highly enriched in subjects with obesity compared to subjects without obesity. By using a machine learning approach, we were able to link fecal metabolic pathways to gut microbial species. *C. aerofaciencs, R. torques,* and *R. obeum,* correlated strongly with microbial pathways involved in carbohydrate metabolism, whereas *A. shahii* and *A. senegalensis* were negatively correlated with these pathways. Previously we have shown that some of these bacterial species (e.g., *R. torques*) were altered upon lean donor FMT, which was associated with improvements in insulin sensitivity upon lean donor FMT underscoring potential causality[21]. Moreover, our findings are in line with previous observations and suggest that the gut microbiome of subjects with obesity may have a higher capacity for carbohydrate utilization and derive more energy from diet compared to subjects without obesity[22]. However, to what extent these specific microbial species play a causal role in increased carbohydrate utilization in humans needs to be further investigated.

We next investigated differences in fecal metabolic pathways between subjects with and without obesity. Pathways involving biosynthesis of several amino acids, including L-histidine, L-lysine, and L-tryptophan were highly enriched in subjects with obesity, whereas pathways involved in the degradation of these amino acids, in particular L-histidine, were depleted. Interestingly, these are all essential amino acids, that cannot be synthesized by the human body, suggesting that alterations in dietary intake is responsible for these differences. Nevertheless, protein intake was only slightly higher in subjects with obesity compared to subjects without obesity. This suggests that the microbiome of subjects with obesity has a higher potential to produce several amino acids compared to subjects without obesity but have reduced capacity of catabolizing specific amino acids. Although, previous reports have linked amino acid metabolism to obesity[8,17,23], we show very strong associations (i.e.,

rho up to 0.5) between BMI and fecal microbiota metabolic pathways related to amino acid biosynthesis. Interestingly, histidine metabolism has recently been identified as a direct contributor to the development of insulin resistance via microbial conversion of this dietary amino acid into imidazole propionate[24]. To further investigate which microbial species are linked to fecal metabolic pathways, we used a machine learning approach in and found several strong associations between microbial species and microbial amino acid metabolism. However, despite these strong associations, to what extent these microbial species are capable of increased histidine production or degradation remains to be demonstrated. Thus, more insight is needed to clarify the role of these specific microbial species and subsequent alterations in histidine catabolism in the development of obesity.

Fecal metagenomic data was recently shown to improve prediction accuracy for clinical features of obesity, including glucose, BMI, and lipid levels, compared to models that use only host genetic and environmental data[9]. In our cohort, 50% of the variance in BMI and body fat percentage could be explained by gut microbiota composition, whereas metabolic pathways could explain almost 60% of this variance. The explained variance in BMI in our study is almost twice as high compared to the study of Rothschild and colleagues[9]. This discrepancy is likely to be explained by differences in inclusion criteria between studies. In our cohort, subjects with a wide range in BMI (18-60 kg/m^2) were included, whereas the Rothschild study included subject with a narrower mean BMI (26 ± 5.2 kg/m^2). In addition, the gut microbiome was able to explain the variance in triglycerides and HDL for respectively 6% and 9%, which is in line with a previous observation that showed that the gut microbiome could explain the variance in triglycerides by 6% and HDL by 4%[25]. Nevertheless, we could not explain the variance in glycated hemoglobin and fasting glucose by the gut microbiome composition. This observation is in line with a recent study that showed that neither taxonomic profiles nor functional capacity is significantly associated with fasting glucose levels[26].

Our study has limitations. We note that results regarding gut microbiome alterations may be confounded by unmeasured environmental factors. Moreover, the prediction models cannot be used to determine causality since gut microbiota composition can both affect and be affected by host phenotypes. Another limitation of our study is that we included only Caucasian subjects, therefore it is not possible to extrapolate these findings to other ethnicities. Nevertheless, our results are in line with a study conducted in China[8]. Despite the strong link between alterations in amino acid metabolism and obesity, we could not support these findings with alterations in plasma metabolites since this data was not available. Another limitation is the relatively small sample size. However, only recently it has been demonstrated that clinical features of obesity can be accurately predicted using fecal metagenomics data of only a few hundred subjects. Nevertheless, cohorts including subjects

with BMI over 60 kg/m^2 and available shot gut microbiota data are scarce. We note that according to the dietary questionnaires, subjects with obesity had a lower energy, fiber and carbohydrate intake compared to subjects without obesity. Overall, macronutrient intake was in line with previous findings of our group[27]. Nevertheless, it is important to consider that subjects with overweight and obesity underreport their dietary intake ranging from 20-50%, making it therefore a frequently recognized problem[28]. Another limitation regarding food intake is that we did not obtained any data on the use of food additives such as artificial sweeteners. It is well known that next to macronutrients, artificial sweeteners are capable of inducing both compositional and functional changes in the gut microbiome[29,30].

In conclusion, we could separate subjects with and without obesity based on their microbiome composition and we identified 52 gut microbial species with differential abundance in these subjects. We found that gut microbiome composition could explain more than 50% of the variance in clinical features of obesity such as BMI, waist, HDL and triglycerides. In addition, we observed differences in fecal microbial metabolic pathways in people with and without obesity, and particularly strong associations between amino acid metabolism, microbial species and obesity. These results identify new bacterial species that are altered in subjects with morbid obesity and also contribute to a possible robust consensus signature for the obese and lean microbiome.

Competing interests

M. N. is on the Scientific Advisory Board of Caelus Health, The Netherlands. F.B. is on board of directors of Metabogen, Sweden. However, none of these possible conflicts of interest bear direct relation to the outcomes of this specific study. All other authors declare that they have no competing interests.

Authors' contributions

AM, OA, YA, SB performed the patient visits and were involved in the recruitment of subjects. AM, AP, EL, performed the data analyses. VT, HH were responsible for sample processing. AG, FB, VG, HH, MN designed the study and supervised all parts of the project. AM, AG, HH, MN drafted the manuscript. All authors provided support and constructive criticism throughout the project and approved the final version of the paper.

Acknowledgments

We thank the nursing staff and all participating subjects for making this study possible. We also thank Manuela Krämer and Robert Jakubowicz for the preparation of fecal genomic DNA. This study was funded by Le Ducq consortium grant 17CVD01 and Novo Nordisk Foundation Gut-MMM grant.

REFERENCES

1. Ng, M. *et al.* Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* (2014) doi:10.1016/S0140-6736(14)60460-8.
2. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
3. Schwartz, M. W. *et al.* Obesity Pathogenesis: An Endocrine Society Scientific Statement. *Endocrine Reviews* **38**, 267–296 (2017).
4. Meijnikman, A. S., Gerdes, V. E., Nieuwdorp, M. & Herrema, H. Evaluating causality of gut microbiota in obesity and diabetes in humans. *Endocrine Reviews* Preprint at https://doi.org/10.1210/er.2017-00192 (2018).
5. Vrieze, A. *et al.* Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, (2012).
6. Ley, R., Turnbaugh, P., Klein, S. & Gordon, J. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–3 (2006).
7. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–31 (2006).
8. Liu, R. *et al.* Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nature Medicine* **23**, 859–868 (2017).
9. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* (2018) doi:10.1038/nature25973.
10. Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nature Methods* **15**, 962–968 (2018).
11. Truong, D. T. *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature methods* **12**, 902–3 (2015).
12. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. doi:10.1145/2939672.2939785.
13. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* vol. 57 289–300 Preprint at https://doi.org/10.2307/2346101 (1995).
14. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* (2018) doi:10.1038/nature25973.
15. Ley, R., Turnbaugh, P., Klein, S. & Gordon, J. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–3 (2006).
16. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* **352**, 565–569 (2016).
17. Aron-Wisnewsky, J. *et al.* Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut* **68**, 70–82 (2019).
18. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457**, 480–484 (2009).
19. Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science translational medicine* **1**, 6ra14 (2009).
20. Jumpertz, R. *et al.* Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *The American journal of clinical nutrition* **94**, 58–65 (2011).
21. Kootte, R. S. *et al.* Improvement of Insulin Sensitivity after Lean Donor Feces in Metabolic Syndrome Is Driven by Baseline Intestinal Microbiota Composition. *Cell metabolism* **26**, 611-619.e6 (2017).
22. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–31 (2006).
23. Pedersen, H. K. *et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* (2016) doi:10.1038/nature18646.

24. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).

25. Fu, J. *et al.* The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids. *Circulation research* **117**, 817–24 (2015).

26. Thingholm, L. B. *et al.* Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell Host and Microbe* (2019) doi:10.1016/j.chom.2019.07.004.

27. De Groot, P. *et al.* Donor metabolic characteristics drive effects of faecal microbiota transplantation on recipient insulin sensitivity, energy expenditure and intestinal transit time. *Gut* (2019) doi:10.1136/gutjnl-2019-318320.

28. Park, H. A., Lee, J. S. & Kuller, L. H. Underreporting of dietary intake by body mass index in premenopausal women participating in the Healthy Women Study. *Nutrition Research and Practice* (2007) doi:10.4162/nrp.2007.1.3.231.

29. Chassaing, B. *et al.* Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* **519**, 92–96 (2015).

30. Suez, J. *et al.* Artificial sweeteners induce glucose intolerance by altering the gut microbiota. *Nature* **514**, 181–186 (2014).

Supplementary data available at:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fjoim.13137&file=joim13137-sup-0001-Supinfo.docx

4

A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study

Van Olden CC#, Van de Laar AW#, Meijnikman AS, Aydin O, Van Olst N, Hoozemans J.B., De Brauw LM, Bruin SC, Acherman YIZ, Verheij J, Pyykkö JE, Hagedoorn M, Sanderman R, Bosma NC, Tremaroli V, Lundqvist A, Olofsson LE, Herrema H, Lappa D, Hjorth S, Nielsen J, Schwartz T, Groen AK, Nieuwdorp M, Bäckhed F, Gerdes VEA.

#) Equal contribution

Journal of Internal Medicine, 2021 Mar;289(3):340-354

ABSTRACT

Prevalence of obesity and associated diseases, including type 2 diabetes mellitus, dyslipidaemia and non-alcoholic fatty liver disease (NAFLD), are increasing. Underlying mechanisms, especially in humans, are unclear. Bariatric surgery provides the unique opportunity to obtain biopsies and portal vein blood-samples. The BARIA Study aims to assess how microbiota and their metabolites affect transcription in key tissues and clinical outcome in obese subjects and how baseline anthropometric and metabolic characteristics determine weight loss and glucose homeostasis after bariatric surgery.

We phenotype patients undergoing bariatric surgery (predominantly laparoscopic Roux-en-Y gastric bypass), before weight loss, with biometrics, dietary and psychological questionnaires, mixed meal test (MMT) and collect fecal-samples and intra-operative biopsies from liver, adipose tissues and jejunum. We aim to include 1500 patients. A subset (approximately 25%) will undergo intra-operative portal vein blood-sampling. Fecal-samples are analyzed with shotgun metagenomics and targeted metabolomics, fasted and postprandial plasma-samples are subjected to metabolomics, and RNA is extracted from the tissues for RNAseq-analyses. Data will be integrated using state-of-the-art neuronal networks and metabolic modeling. Patient follow-up will be ten years. Preoperative MMT of 170 patients were analysed and clear differences were observed in glucose homeostasis between individuals. Repeated MMT in 10 patients showed satisfactory intra-individual reproducibility, with differences in plasma glucose, insulin, and triglycerides within 20% of the mean difference.

The BARIA study can add more understanding in how gut-microbiota affect metabolism, especially with regard to obesity, glucose metabolism and NAFLD. Identification of key factors may provide diagnostic and therapeutic leads to control the obesity-associated disease epidemic.

INTRODUCTION

Obesity is on the rise. At the current pace, more than one billion adults will be obese by 2030[1]. An increase in obesity-associated diseases will follow in its wake, including type 2 diabetes mellitus (T2DM), dyslipidaemia, non-alcoholic fatty liver disease (NAFLD) and cardiovascular disease. However, it has been challenging to identify underlying molecular mechanisms contributing to cardiometabolic diseases, in part because T2DM has several subclasses[2]. Several pathways have been suggested to contribute to obesity and impaired glucose control, such as the immune system and gut microbiota[3–5]. They include short chain fatty acids, bile acids, amino acids derived metabolites, neural pathways, and lymphoid cells. Interestingly, these have also been shown to be involved in glucose metabolism and the development of NAFLD, which illustrates the interconnectivity of cardiometabolic diseases. Moreover, a chronic low-grade inflammation can be measured in individuals with obesity, possibly caused by a disturbance in the intestinal microbiota composition. Faecal microbiota transplantation (FMT) from human subjects to mice transferred adiposity phenotype suggesting that, in mice, the microbiota may be a contributing factor[6]. In humans, the effect of FMT is less significant, yet insulin sensitivity can improve for a short while in individuals with metabolic syndrome after infusion of intestinal microbiota from lean donors[7].

The relative contribution of different organs (liver, adipose tissue and gut) to whole body metabolism as well as immunological tone on weight loss in relation to improvement of insulin sensitivity is not known. Neither are the mechanisms that trigger the innate and adaptive (intestinal) immune system by altered intestinal microbiota, or their effects on metabolism. Most interventions aimed at losing weight in individuals with morbid obesity have little effect, except for bariatric surgery[8]. Bariatric surgery is also the most effective intervention to reduce obesity-related morbidity and mortality[9]. In this regard, one of the most common and well-studied bariatric procedures is laparoscopic Roux-en-Y gastric bypass (LRYGB). The increased insulin sensitivity found shortly after LRYGB, even before significant weight loss is obtained, suggests immediate systemic changes in metabolism upon surgery, which are long standing, as even ten years after surgery beneficial effects on glucose metabolism, lipids and blood pressure can be seen[10,11]. Although being an important treatment for over forty years, the mechanisms behind the beneficial effect of bariatric surgery have been elusive. They may include bile flow alteration, reduction of gastric size, anatomical rearrangement and altered flow of nutrients, vagal manipulation and enteric gut hormone modulation[12]. Although some studies have demonstrated that intestinal microbiota are altered after bariatric surgery as well, the prospective value of (baseline) intestinal microbiota composition and the relation with the (diet derived) metabolites that these bacteria produce has never been investigated at a larger scale[13,14].

Significant differences in the response to bariatric surgery can be observed, both in weight loss, obesity related morbidities and psychological factors, including self-esteem, risk of addiction and quality of life[15–18]. Despite some methodological limitations, psychological studies have shown improvements in psychopathology, eating disorders, depressive symptoms, body image and social functioning after bariatric surgery[19]. Systems biology models can provide an advanced reconstruction of individuals' metabolism at different organ levels in patients with morbid obesity. They could provide a valuable tool in predicting individuals' outcomes of bariatric surgery and hereby develop a personalized medicine approach for this disease. First steps in utilising this technique to study altered metabolism in obesity related diseases have produced interesting results[20–23].

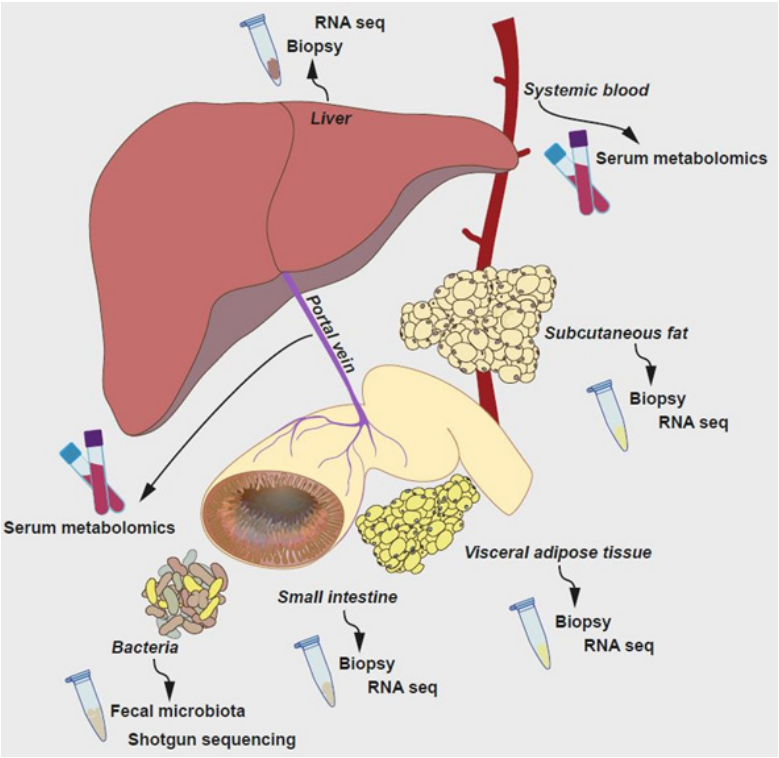


Figure 1. A systems biology approach, identifying gut microbial, immunological and metabolic markers in a large and well phenotyped bariatric surgery cohort: the BARIA study.

We aim to perform a systems biology approach, as schematically depicted in figure 1, identifying gut microbial, immunological and metabolic markers in a large and well phenotyped bariatric surgery cohort (BARIA study) to identify signalling pathways that can affect

metabolic circuits in humans. Our study aims to identify novel pathways in the pathogenesis of obesity, T2DM and NAFLD, taking the gut-brain axis into account as well, which may be targets for drug development. Finally, we will follow the patients prospectively in an attempt to identify mechanisms affecting the surgical outcome.

METHODS

Study design

We include subjects that are patients with morbid obesity scheduled for bariatric surgery. From September 2016 until the end of 2018 the study was performed at the former MC Slotervaart (Amsterdam) and is now continued, after closure of that hospital, by the same surgical group and research team at the Spaarne Gasthuis hospital (Hoofddorp) in the Netherlands. The study protocols were approved by the Ethical Review Board of the Academic Medical Center, Amsterdam, (approval code: NL55755.018.15) and all patients that have been (and will be) included provided informed consent. Preoperative screening, surgery and follow up are performed following institutional procedure protocols. All patients are screened preoperatively by a bariatric surgeon, an internist, a dietician, and a psychologist. We aim to include predominantly LRYGB procedures. In a shared-decision making process, surgeon and patient decide for the bariatric procedure type: LRYGB, laparoscopic omega-loop gastric bypass (LOGB) or laparoscopic sleeve gastrectomy (LSG), which, in our bariatric surgery centre, has resulted in more than 90% LRYGB of all surgeries in the past ten years. All LRYGB are standardized, with approximated measurements of 4x8 cm gastric pouch, 50 cm biliopancreatic limb, 150 cm alimentary limb[24]. The LOGB is made with a longer gastric pouch and a longer biliopancreatic limb of approximately 200 cm. The LSG is calibrated with a 34 Charrière bougie with the staple line starting at approximately 2 cm from the pylorus.

Study Population

Patients are screened at the outpatient clinic (MC Slotervaart hospital, Spaarne Gasthuis hospital) after being approved for bariatric surgery. Screening started in September 2016. We aim to include 1500 patients. Subjects are considered eligible for participation if they meet following criteria:

Inclusion

- Male and female patients scheduled for primary bariatric surgery recruited from an experienced Dutch bariatric surgery clinic.

- Body mass index (BMI) ≥40 kg/m^2, or: BMI ≥35 kg/m^2 with obesity related co-morbidity.
- Recent history of supervised attempts to lose weight.
- Age 18 to 65 years.
- Ability to provide informed consent.

Exclusion

- Primary lipid disorder.
- Known genetic basis for insulin resistance or glucose intolerance.
- Psychiatric conditions.
- Coagulation disorders (patient reported, or prolonged prothrombin time or activated partial thromboplastin time).
- Uncontrolled hypertension (blood pressure > 150/95 mmHg).
- Renal insufficiency (creatinine > 150 umol/L).
- Excessive alcohol intake (> 14 units/week, patient reported).
- Pregnancy, breastfeeding.

Outcome measures

For the characterization of subjects before surgery we have chosen variables that are linked to obesity and obesity associated diseases. For clinical follow-up we chose variables that can be tested minimally invasive (only venepuncture) and which can be easily reproduced, at low cost, without extensive training in a Western hospital. The reason for this is twofold. First, we aimed to minimize the demand of our study subjects. Second, our results need to be reproducible and applicable in other settings without the need for major investments in equipment or logistics. That way our project can benefit the greatest number of people while still remain ambitious in aiming to discover new mechanisms.

The included patients undergo the repetitive measurements detailed in table 1.

Visit	Type of measurement	Specific values	Biological samples stored in biobank
- baseline - 1 year - 2 years - 5 years* - 10 years*	demographic	age, sex, medical history, medication use, history of obesity, history of smoking and alcohol, education level, employment status, anticonception use, physical activity	
	biometric	height, weight, waist- and hip circumference, temperature, blood pressure, pulse, non-invasive haemodynamics (stroke volume, cardiac output, systemic vascular resistance), bioelectrical impedance measurement, electrocardiogram	
	blood	haemoglobin, CRP, leukocytes, platelets, HbA1c, glucose, electrolytes, kidney function, lipid profile, iron, hepatic enzymes, thyroid profile, plasma metabolites	stored plasma and DNA samples (-80°c)
	mixed meal test	glucose, insulin, triglycerides	stored plasma samples (-80°c)
	dietary questionnaire	Satiety (visual analogue scale)[25], dietary intake last 3 days prior to 24h faeces collection	
	psychological questionnaire	see table 2.	
	morning faecal samples 24h faeces	gut microbiota composition and faecal metabolites (scfa), bile acids and caloric bomb	stored samples (-80°c)
	gingival swab	oral microbiota	stored samples (-80°c)
	urine	albumin and creatinine, metabolites	stored samples (-80°c)
- primary operation - re-surgery	liver biopsy	snap frozen (liquid N_2) and formaldehyde	stored samples (-80°c) and paraffin
	subcutaneous adipose tissue	snap frozen (liquid N_2) and formaldehyde	stored samples (-80°c) and paraffin
	visceral adipose tissue	snap frozen (liquid N_2) and formaldehyde	stored samples (-80°c) and paraffin
	omental adipose tissue	snap frozen (liquid N_2) and formaldehyde	stored samples (-80°c) and paraffin
	portal vein blood (subset)	plasma metabolites and proteomics	stored plasma samples (-80°c)
	small intestine biopsy (LRYGB only)	snap frozen (liquid N_2) and formaldehyde	stored samples (-80°c) and paraffin
- 6 weeks - 6 months	biometric	weight, waist- and hip circumference, blood pressure and pulse	
	blood	haemoglobin, CRP, leukocytes, platelets, HbA1c, glucose, electrolytes, kidney function, lipid profile, iron, hepatic enzymes, thyroid profile, plasma metabolites	stored plasma samples (-80°c)
- 2 weeks - 6 weeks - 6 months	morning faeces	gut microbiota composition and faecal metabolites (scfa)	stored samples (-80°c)
	urine	albumin and creatinine, metabolites	stored samples (-80°c)

Table 1. Overview of visits and measurements. BARIA longitudinal cohort study. * at 5 and 10 years no mixed meal test will be performed

For the physicians and researchers, we made a standard operating procedure. The psychological measures were assessed with Dutch versions of validated questionnaires, presented in table 2. Tissue biopsies are obtained during operation of three adipose tissue compartments: subcutaneous (from one of the laparoscopic incisions in the upper abdomen), greater omentum, and visceral fat (omental appendices of the transverse colon); from the diaphragmatic surface of segment three or five of the liver; and from the jejunum at the site of the jejuno-jejunostomy, approximately 50 cm from the Treitz ligament. The jejunum biopsy cannot be obtained during LOGB or LSG, as those operation techniques, unlike LRYGB, do not involve a jejuno-jejunostomy. Blood sample of the portal vein is taken at the beginning of the surgery, only if considered safe by the surgeon, mainly depending on the amount of fatty tissue surrounding the hepatoduodenal ligament. Biopsies are assessed for histology (paraffin embedded), gene regulation (RNA-sequencing) and protein expression (immunoblotting). NAFLD status is determined in histology of liver biopsies and individually scored by members of Dutch Liver Pathology Panel, after training sessions, while difficult or borderline cases are discussed during panel meetings for consensus. SAF scores are determined, separately assessing steatosis (S), activity (A, the sum of hepatocyte ballooning and lobular inflammation), and fibrosis (F)[26]. From the beginning of 2019 we added routine preoperative ultrasound of the gallbladder. Hollow needle subcutaneous fat aspirate biopsy under local anesthesia (peri-umbilical region) is optional at follow up. Of note, the tissues collected during surgery comprise tissue that is thought to play a crucial role in glucose metabolism and can be biopsied with minimal risk to the patient being small intestine, adipose tissue and liver samples. We assess all liver biopsies for NAFLD/NASH, as it is the gold standard for diagnosing liver disease.

Plasma metabolites are studied in portal vein blood (fasted) and in both fasted and two hours after mixed meal test (MMT) peripheral blood samples. Intestinal immunological cells are looked for in GALT tissue (Peyer's patches), visceral and subcutaneous adipose tissue, liver in relation to inflammation gene expression (IL -1β, IL-6, IL-8, IL-18, CXCR2 TNF-α and TLR 1, 2, 4, 5 and 6 and IRX 3 and 5 and RNA-sequencing) and in specific innate lymphoid cells (ILC), macrophages, T/B-cells and dendritic cells, and peripheral blood. Immunological parameters assessed in small-intestinal tissue and adipose tissue were selected for those that are linked in literature to have an effect on glucose metabolism and with which we have experience in the analysis. Morning faecal samples obtained at several time points will be analysed by shotgun sequencing (NovaSeq). Buffycoat samples of peripheral blood are taken at baseline for genomic DNA analyses. Cardiac output and peripheral resistance are assessed using the Nexfin system, measuring blood pressure beat-to-beat with a small cuff around the index finger[27].

Questionnaire	#
Sociodemographic information: place of birth patient, father, mother; number of children; marital status; education; occupation.	7
Quality of life (WHO HIV QOL)	2
Change in life	1
Professional support	5
Self-management after Bariatric surgery (BSSQ)	8
TFEQ- hunger scale	9
Center for Epidemiology Studies Depression Scale Revised (CES-D)	20
Impact of Weight on Quality of Life (IWQOL-Lite)	31
Body Image Scale	10
De Jong-Gierveld Loneliness Scale	11
Social Participation Scale	3
SCI Exercise Self-Efficacy	10
Stanford Exercise behaviour	6
Weight Efficacy Lifestyle Questionnaire (WEL-Q)	20
G-Food Craving Questionnaire-Trait (FCQ-T)	21
Quality of Relationship and Relationship ladder	2
Experience in Close Relationships Scale (ECRR-SF)	16
Social Support (SSQSR)	12
Social Support And Diet	10
Social Support And Exercise	13
Personality NEO-FFI (neuroticism and conscientiousness subscales)	12+12
Self-compassion Scale Short Form	12
Rosenberg Self-esteem Questionnaire	10
Chronotype working day	8
Chronotype free day	8

Table 2. Psychological questionnaires. BARIA longitudinal cohort study.

In the case of a non-acute operation more than one month after primary surgery, for example for laparoscopic cholecystectomy, new liver and adipose tissue biopsies can be obtained, as well as gallbladder and bile from cholecystectomy patients. Gallbladder tissue will be assessed for bile acid composition, histology, gene expression (RNA-sequencing) and protein expression.

The two hour, seven sample oral MMT, as described by Dalla Man et al. is repeated several times over 2 years follow up[28]. It consists of two Nutridrink compact 125ml (Nutricia®), containing 23.3 grams fat, 74.3 grams carbohydrates (of which 38.5 grams sugar) and 24.0 grams protein. The patients receive this meal after fasting for a minimum of nine hours. Time point zero is the moment the patient fully consumed the meal. Blood samples are drawn via intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes and analysed for insulin sensitivity / insulin resistance, plasma metabolites and bile acids.

Data handling and analysis

Data are collected on data collecting forms and entered after validation in a computer system for subsequent tabulation and statistical analysis. All research and medical data is kept strictly confidential and registered under a unique study code. Only the researchers that are involved in this study are able to see the data and to identify a participant. Study material will be stored for a period of 20 years after study completion. Data from the first approximately 100 patients is analyzed to check data quality and logistics (first data-freeze). A first interim analysis will be performed on data of the first approximately 300 patients and the primary analysis will be performed on data of 500 patients (second and third data-freeze). We intend to continue inclusions till 1500 for additional analyses and validation of primary findings. The data are analyzed using a range of different techniques, including being used as input for metabolic modeling and for phenotyping the patients using machine learning algorithms.

Study integrity, monitoring, safety

The BARIA study is conducted according to the principles of the Declaration of Helsinki (October 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO). All adverse events reported by the patients or observed by the investigator or staff will be recorded. All adverse events will be followed until they have abated, or until a stable situation has been reached. Depending on the event, follow up may require additional tests or medical procedures as indicated.

Validation of the mixed meal test

Next to the elaborate analysis of data focussing on the aims of the BARIA study, we used the results of the preoperative MMT of the patients included and operated in the first two years of the study to validate the reproducibility of the MMT-stimulated postprandial glucose, triglycerides and insulin curves. We therefore stratified these results by classifications of glycaemic control as formulated in the American Diabetes Association (ADA) criteria: normoglycemia (fasting glucose (FG) <100mg/dL; <5.6mmol/L), impaired FG (100-125mg/dL; 5.6-6.9mmol/L) and / or increased haemoglobin A1c (5.7-6.4%; 39-47mmol/mol) and diabetes mellitus (FG ≥126mg/dL; ≥7.0mmol/L)[29]. Of all measurements during MMT in these patients there were 2.1% missing values for glucose, 5.5% for insulin and 1.8% for triglycerides. We repeated the preoperative MMT after one week in ten randomly selected patients. Of all repeat measurements of those ten patients there were 2.9% missing values for glucose, 5.7% for insulin and none for triglycerides. For validating the MMT, imputation of predictive mean matching was performed for all missing values.

RESULTS

Inclusion of patients in the BARIA study began in September 2016. During the first two years of the BARIA study, portal vein sampling was performed in 32% of the surgeries. Types of procedure were 94% LRYGB, 6% LOGB and no LSG. No serious adverse events occurred.

Baseline characteristics and MMT results of the first 170 patients included in this two year period are presented in table 3.

	Healthy	IFG	IHbA1c	Comb	T2DM
n	57	21	19	26	47
age (years)	41.4 (11.1)	46.8 (11.7)	44.6 (9.5)	49.2 (9.2)	49.5 (10.2)
sex (female)	45 (78.9)	20 (95.2)	17 (89.5)	16 (61.5)	31 (66.0)
BMI	39.5 (3.9)	39.4 (3.1)	40.6 (7.1)	40.6 (3.6)	39.2 (4.5)
hypertension	8 (14.0)	5 (23.8)	3 (15.8)	8 (30.8)	25 (53.2)
Systolic BP (mmHg)	129.5 (16.6)	130.6 (13.6)	134.2 (15.8)	133.2 (12.0)	132.1 (13.7)
Diastolic BP (mmHg)	80.1 (11.3)	80.5 (8.2)	78.1 (13.2)	84.0 (7.9)	82.6 (9.4)
insulin use					10 (21.3)
glucose (mmol/l)	5.1 (0.4)	5.9 (0.2)	5.2 (0.2)	6.1 (0.4)	7.4 (1.5)
insulin (pmol/l)	84.8 (48.0)	89.4 (46.5)	79.2 (37.2)	111.2 (46.9)	180.2 (222.5)
HbA1c (%)	5.31 (0.23)	5.41 (0.19)	5.79 (0.09)	5.88 (0.17)	7.10 (1.14)
HOMA2 IR	1.60 (0.90)	1.71 (0.83)	1.48 (0.67)	2.14 (0.85)	2.44 (1.24)
HOMA2 Beta (%)	125.4 (50.9)	98.1 (37.2)	112.6 (33.9)	105.8 (38.7)	87.3 (37.2)
AUC glucose (mmol/l)	137.1 (109.5)	122.5 (85.9)	194.6 (112.9)	211.7 (105.0)	386.3 (193.7)
AUC insulin (mmol/l)	42.3 (30.4)	46.0 (29.4)	48.7 (21.4)	50.8 (20.8)	37.6 (31.5)
eGFR (MDRD ml/min/1,73m^2)	94.5 (18.0)	92.7 (19.8)	95.6 (21.7)	94.7 (19.7)	95.7 (17.6)
ASAT (U/l)	23.6 (4.9)	23.5 (6.5)	25.1 (5.5)	25.3 (4.9)	29.9 (14.0)
ALAT (U/l)	28.6 (13.4)	28.3 (14.7)	33.7 (18.5)	30.4 (10.1)	42.1 (25.8)
Cholesterol (mmol/l)	4.6 (1.0)	5.1 (1.2)	5.2 (1.0)	4.8 (1.1)	4.1 (0.9)
HDLc (mmol/l)	1.12 (0.29)	1.13 (0.23)	1.16 (0.16)	1.08 (0.29)	1.05 (0.23)
Triglycerides (mmol/L)	1.08 (0.44)	1.58 (0.91)	1.10 (0.42)	1.79 (1.17)	1.40 (0.62)

Table 3. Baseline characteristics and results of mixed meal test in 170 participants in the first two years of inclusion in the BARIA longitudinal cohort study, stratified by glycaemic classification, as formulated in the American Diabetes Association criteria: normoglycemic (Healthy), impaired fasting glucose (IFG), increased haemoglobin A1c (IHbA1c), combination of IFG and IHbA1c (Comb) and Type-2 Diabetes Mellitus (T2DM). Categorical variables are displayed as absolute numbers (percentage), continuous variables as means (SD).

MMT curves of ten patients assigned to the category diabetes mellitus were excluded because of insulin use. Results of the preoperative MMT of the remaining 160 patients are presented in figure 2.

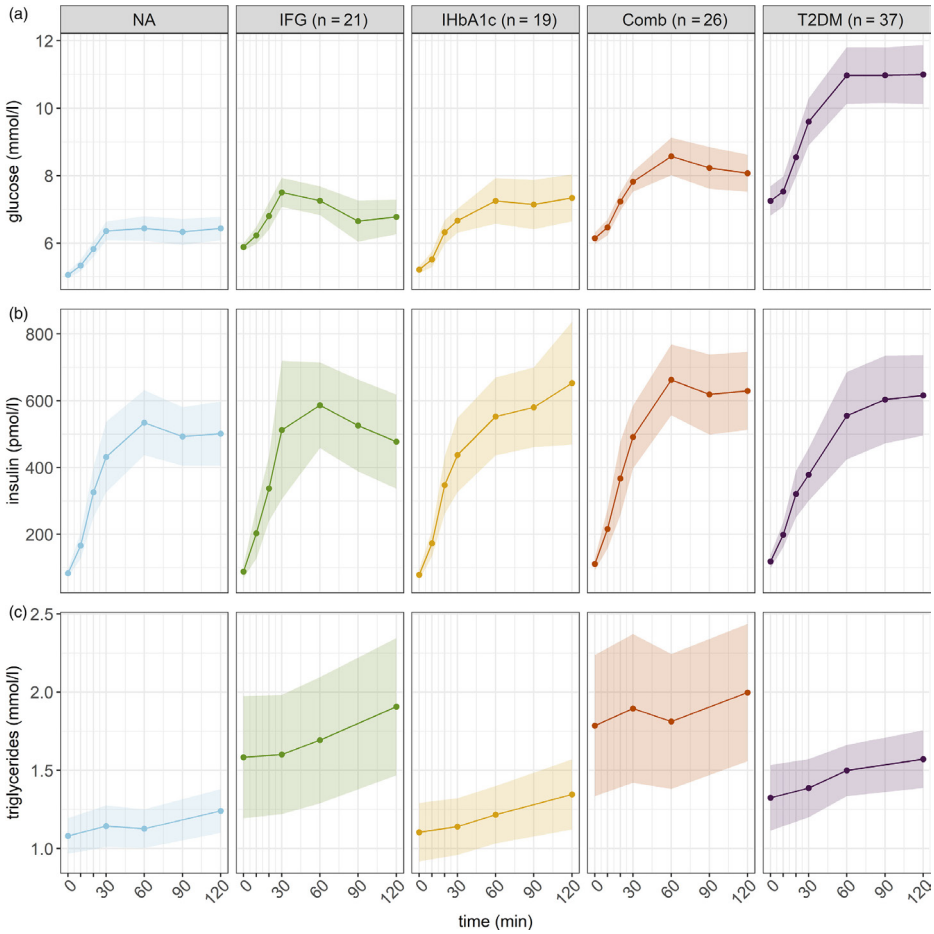


Figure 2. Glucose, insulin and triglycerides measurements during 2-hour 7-sample mixed meal test, stratified by glycaemic classification, as formulated in the American Diabetes Association criteria: normoglycemic (Healthy), impaired fasting glucose (IFG), increased haemoglobin A1c (IHbA1c), combination of IFG and IHbA1c (Comb) and Type-2 Diabetes Mellitus (T2DM). Values are presented as means with 95% confidence intervals. **A)** glucose curves; **B)** insulin curves; **C)** triglycerides curves.

Individuals with different classifications of glycaemic control showed markedly different profiles for MMT-stimulated plasma insulin, glucose and triglycerides. Triglycerides were clearly higher at baseline and all following time points in patients with IFG, with or without increased Hba1c. HOMA2-IR and HOMA2-B values and correlations with postprandial glucose and insulin curves are presented in figure 3. The HOMA2-IR and HOMA2-B values showed a good correlation with the AUC postprandial insulin, but not with the AUC postprandial glucose.

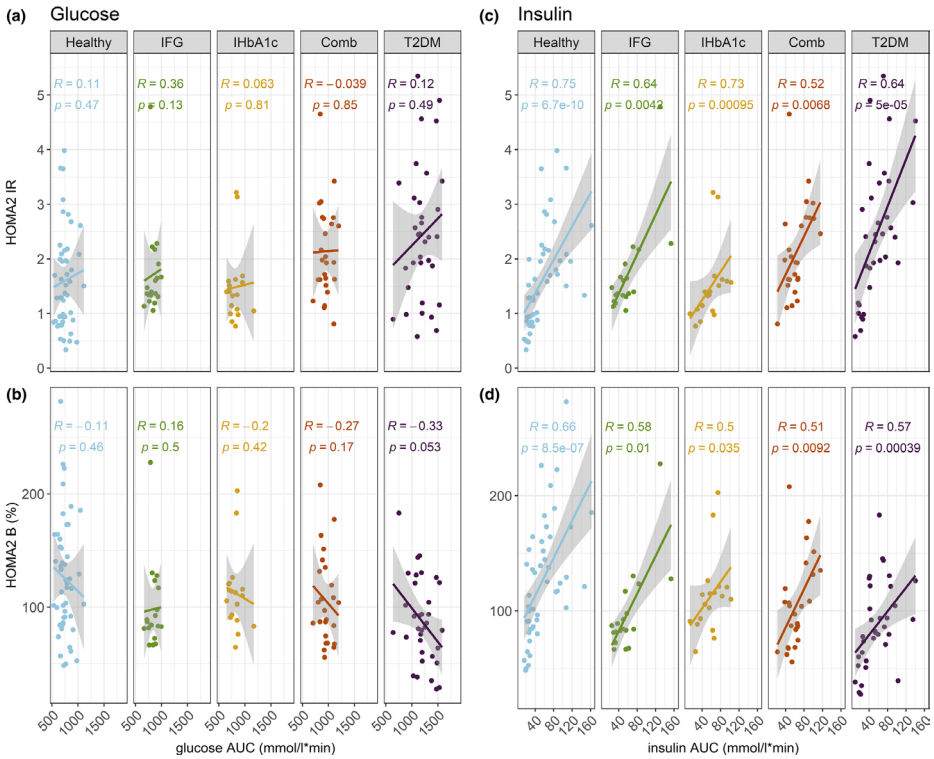


Figure 3. Area under the curve (AUC) of insulin and glucose during mixed meal test and HOMA2 insulin resistance (IR) and beta cell function (B), stratified by glycaemic classification, as formulated in the American Diabetes Association criteria: normoglycemic (Healthy), impaired fasting glucose (IFG), increased haemoglobin A1c (IHbA1c), combination of IFG and IHbA1c (Comb) and Type-2 Diabetes Mellitus (T2DM). Points are individual values, solid line represents linear regression, banded area is 95% confidence interval. A) Glucose AUC and HOMA2 IR. **B)** Glucose AUC and HOMA2 B (%). **C)** Insulin AUC and HOMA2 IR. **D)** Insulin AUC and HOMA2 B. Correlation coefficient (R) and p-values calculated with Spearman's rank correlation test.

Results of the ten patients that underwent repeated (1 week interval) preoperative MMT are presented in figure 4. We found a good coefficient of variance (figure 4, blue lines) with a mean average of difference between two MMT measurements of 6.3% for area under the curve (AUC) postprandial glucose, 13.9% for AUC postprandial insulin and 7.4% for AUC postprandial triglycerides, while most of the differences between the two measurements were well within the 20% range of the average mean difference underscoring reasonably good intra-individual reproducibility.

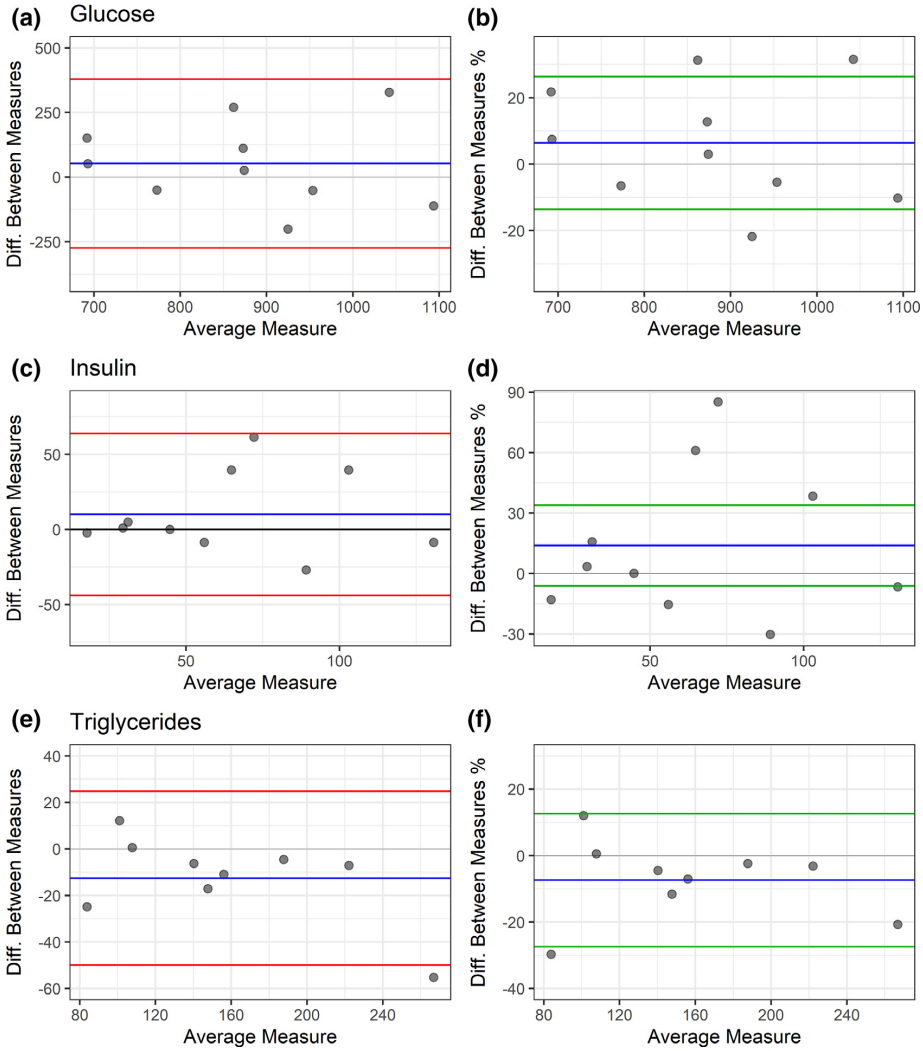


Figure 4. Reproducibility of mixed meal test (MMT). Bland Altman plots of MMT (repeated within 1 week) for glucose, insulin and triglycerides. Blue line is mean of difference between measurements, red line is +/- 1.96*SD of mean difference, green line is +/- 20% of mean difference. **A)** Glucose area under the curve (AUC) in mmol/L*time. **B)** Glucose AUC percent change. **C)** Insulin AUC in mmol/L*time. **D)** Insulin AUC percent change. **E)** Triglycerides AUC in mmol/L*time. **F)** Triglycerides AUC percent change.

DISCUSSION

The BARIA cohort study will generate a large phenomic database on the systems biology of subjects with morbid obesity, both before and after bariatric surgery. Advanced data science, including application of machine learning and artificial neural networks data analysis is used to select microbiome-produced metabolites and identify their receptors in target tissue. It will be the first large bariatric cohort study to include portal vein blood sampling in a considerable subset of patients for untargeted metabolites, which, when also studying peripheral metabolites, will enable to study the gradient of metabolites filtered by the liver. We aim to include 1500 patients undergoing primary laparoscopic bariatric surgery (gastric bypass or sleeve gastrectomy). Before surgery, they are subjected to MMT, blood and fecal sampling, and questionnaires, including psychology and VAS lists taken at the start of the MMT in all patients at all timepoints to minimize variation. During surgery, biopsies are obtained from three fat depots, jejunum, liver, and samples from portal and peripheral venous blood. Thereafter, further sampling (MMT, blood and fecal samples) is performed. In the event of another surgery (revisional surgery, cholecystectomy) further biopsies can be obtained, which is included in the ethical protocol. We process tissues for RNA-sequencing, analyze intestinal microbiota, and perform untargeted (postprandial) plasma metabolomics on both fasting and postprandial (MMT) plasma samples. These metabolites will be investigated further *in vitro* and *in vivo* to determine causality and identify receptors. After the primary analysis, the generated database will also allow for additional secondary analyses.

The bariatric patient scheduled for primary bariatric surgery is an interesting model for several reasons. All patients suffer from morbid obesity and generally expect to undergo examinations, measurements and interviews both prior to surgery and in follow up. The laparoscopic procedures give proper access to different adipose compartments, as well as liver and intestine for biopsy and, if the hepatoduodenal ligament is not too much embedded in fatty tissue, to the portal vein for fine needle blood sample as well. Any haemorrhages can readily be detected and addressed surgically, minimizing the expected adverse events. In the hands of our surgical team, mortality of routine LRYGB is low (0.03%) and two-year follow-up is high (71%)[24]. During the first two years of inclusion, portal vein sampling could be performed safely in about one out of three cases. Other studies with similar bariatric surgery cohorts with invasive assessments showed that a majority of patients remains interested in participating during two years of follow-up[30]. Furthermore, up to 10% of bariatric surgery patients need additional surgery within two years after primary procedure (for example revision surgery or cholecystectomy), which opens up the possibility for renewed biopsies and blood sampling[31].

However, studying bariatric patients has some limitations intrinsic to the surgical procedure. Biopsies and portal vein blood are taken under general anaesthesia and therefore potentially influenced by anaesthesia medication. For example, these drugs will be found in portal vein plasma and might accumulate in fatty tissues during surgery and, most importantly, will be metabolized by the liver. Furthermore, patients are routinely urged by their bariatric surgeon to lose as much weight as possible before the operation to reduce the surgical risk. It can be expected that such forced weight loss will influence metabolism, gene expression and gut microbiota. Although no standardized diet is prescribed, we nevertheless choose to exclude those patients that lose more than 5% in six months (or more than 3% in one month) prior to surgery. Another limitation is the fact that many patients using medication for obesity related diseases will need less or even no medication after bariatric surgery, which might be a confounder for outcome measurements.

In a separate analysis of the MMT results in a subset of included patients we showed that the preoperative MMT has a good intraindividual reproducibility, which makes it a better estimate for glycaemic regulation than the oral glucose tolerance test[32]. We also showed that the MMT is able to represent the underlying metabolic dysregulation well, evident in the different curves and the steady state model assessment. The differences observed in the curves correspond well with the pathophysiology. First, impaired fasting glucose (IFG) is consistent with hepatic insulin resistance as is evident, apart from the increased glucose, by increased baseline insulin and a decreased suppression of apo B production, resulting in increased triglycerides. An initial quick rise in glucose is followed by a steady decline of both glucose and insulin, as peripheral insulin resistance remains largely normal[33]. Second, the increased haemoglobin A1c (IHbA1c) group corresponds with peripheral insulin resistance, represented by a steady increase until the 2-hour time point of both glucose and insulin with relatively normal triglyceride levels. Finally, the group with a combination of IHbA1c and IFG (Comb) and the T2DM group show both characteristics, with the T2DM group reaching higher glucose levels. The HOMA2-IR and HOMA2-B values showed a good correlation with the AUC postprandial insulin, but not with the AUC postprandial glucose, which reiterates the suggestion that they are used best in combination with other clinical parameters[34].

With regard to the use of the MMT in postoperative follow up, it must be noted that the anatomical changes affecting gastric emptying and resorption might impede the comparison of the MMT before and after surgery. However, the MMT is biologically a more relevant test than the glucose tolerance test, as one is rarely solely exposed to glucose without fat and proteins. Studies with a similar follow up using intravenous glucose tolerance test and euglycaemic- hyperinsulinaemic clamp showed an improvement in insulin sensitivity in all patients, with least improvement for TDM2 patients[35,36]. One other study assessing meal response after a follow up of more than one year was cross sectional, but with smaller

numbers[37]. Outcome of the MMT in our BARIA study can provide further insight in the metabolic response following a meal after bariatric surgery. Another limitation of the MMT in bariatric patients is that the test can provoke early dumping, a well-known side effect of LRYGB and LOGB due to loss of pyloric regulation, which makes a heavy caloric MMT hard to endure for some patients in the first years of their follow up.

We believe that the different subclasses of T2D are different paths of progression to the disease, with, in some individuals, a simultaneous existence of several pathways[2]. The underlying molecular mechanisms that lead to these different trajectories are probably different. Similarly, the reversibility and the therapeutic intervention that has the greatest effect on their progression may vary. To the best of our knowledge there are no successful therapeutic modalities specifically aimed at targeting short chain fatty acids (SCFA's), bile acids, amino acid derived metabolites, neural pathways and lymphoid cells with the aim of improving glucose metabolism. There have been several trials using specific SCFA as supplements to improve glucose metabolism and weight loss[38,39]. The effects of the intervention in these studies as well as in faecal microbiota transplantation studies are usually limited with only few showing great improvement[7] where other groups found less efficacy of donor FMT (but were also using different FMT applications), but did observe the similar relation between FMT efficacy and decreased faecal microbiota diversity at baseline[40]. A better understanding of which molecular mechanisms need to be targeted in which patients will lead to a better personalized treatment.

With the comprehensive systems approach of the BARIA longitudinal cohort study, we aim to provide more understanding in to how the (small) intestinal microbiota affects our metabolism, especially with regard to NAFLD and T2DM. Moreover, we aim to identify leads that drive weight loss and psychological improvement upon surgery, thus identifying the causal factors connecting beneficial changes in metabolism, microbiota and immunological tone that will be of value to find new diagnostic and therapeutic leads to control the obesity associated disease epidemic.

Lessons learned so far

During our study we encountered a few learning points, which, we hope, future researches planning similar research, can benefit from and not run into the same problems. We based the feasibility of our protocol on previous studies detailing MMT after RYGB surgery[41,42]. In our study so far, a relatively large number (38 out of 134 participants) of participants exhibited adverse effects during the MMT at the one year after bariatric surgery (nausea, diarrhoea, dizziness and weakness). We suspect these adverse effects to be related to dumping syndrome. The symptoms were not of a severity that we found a need for extra diagnostic tests. None of the subjects experienced loss of consciousness and there was no need for

extended stay in the hospital beyond the normal testing time. Another valuable learning point was related to subject follow-up. In order to achieve a dropout rate of <20%, extensive contact with participants had to be maintained. Many participants needed to be contacted via telephone several times, for reminders to schedule every visit. The amount of manpower and time necessary for that was greater than we anticipated.

At the 6 weeks and 6 month collection timepoints we collect blood for fasting glucose measurement, as well as anthropometric measurements and changes in medication. Our initial aim was to also collect blood at the 2 weeks timepoint. During our try-out phase we discovered that having these measurements was too demanding for our patients during this initial recovery period at 2 weeks. Nutrition questionnaires were also reported as stressful by our patients and we have chosen to only include these in the large (1 and 2 year) collection timepoints.

Acknowledgements / funding

The BARIA study is funded by the Novo Nordisk Foundation (NNF15OC0016798). The BARIA study is a Scandinavian-Dutch collaboration. We thank Anna Hallén for figure 1.

REFERENCES

1. Kelly, T., Yang, W., Chen, C.-S., Reynolds, K. & He, J. Global burden of obesity in 2005 and projections to 2030. *International journal of obesity (2005)* **32**, 1431–7 (2008).
2. Ahlqvist, E. *et al.* Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes and Endocrinology* **6**, 361–369 (2018).
3. Park, H. S., Park, J. Y. & Yu, R. Relationship of obesity and visceral adiposity with serum concentrations of CRP, TNF-alpha and IL-6. *Diabetes research and clinical practice* **69**, 29–35 (2005).
4. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–3 (2006).
5. Tremaroli, V. & Bäckhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **489**, 242–9 (2012).
6. Turnbaugh, P. J., Bäckhed, F., Fulton, L. & Gordon, J. I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell host & microbe* **3**, 213–23 (2008).
7. Kootte, R. S. *et al.* Improvement of Insulin Sensitivity after Lean Donor Feces in Metabolic Syndrome Is Driven by Baseline Intestinal Microbiota Composition. *Cell Metabolism* **26**, 611-619.e6 (2017).
8. Fildes, A. *et al.* Probability of an Obese Person Attaining Normal Body Weight: Cohort Study Using Electronic Health Records. *American journal of public health* e1–e6 (2015) doi:10.2105/AJPH.2015.302773.
9. Sjöström, L. *et al.* Effects of bariatric surgery on mortality in Swedish obese subjects. *The New England journal of medicine* **357**, 741–52 (2007).
10. Sjöström, L. *et al.* Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *The New England journal of medicine* **351**, 2683–93 (2004).
11. Adams, T. D. *et al.* Weight and Metabolic Outcomes 12 Years after Gastric Bypass. *New England Journal of Medicine* **377**, 1143–1155 (2017).
12. Ashrafian, H. *et al.* Diabetes resolution and hyperinsulinaemia after metabolic Roux-en-Y gastric bypass. *Obesity Reviews* **12**, (2011).
13. Tremaroli, V. *et al.* Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation. *Cell Metabolism* **22**, 228–238 (2015).
14. Davies, N. K., O'Sullivan, J. M., Plank, L. D. & Murphy, R. Altered gut microbiome after bariatric surgery and its association with metabolic benefits: A systematic review. *Surgery for Obesity and Related Diseases* vol. 15 656–665 Preprint at https://doi.org/10.1016/j.soard.2019.01.033 (2019).
15. van de Laar, A. W. *et al.* The Dutch bariatric weight loss chart: A multicenter tool to assess weight outcome up to 7 years after sleeve gastrectomy and laparoscopic Roux-en-Y gastric bypass. *Surgery for Obesity and Related Diseases* **15**, 200–210 (2019).
16. Buchwald, H. *et al.* Bariatric surgery: A systematic review and meta-analysis. *Journal of the American Medical Association* vol. 292 1724–1737 Preprint at https://doi.org/10.1001/jama.292.14.1724 (2004).
17. Raaijmakers, L. C. H., Pouwels, S., Thomassen, S. E. M. & Nienhuijs, S. W. Quality of life and bariatric surgery: A systematic review of short- and long-term results and comparison with community norms. *European Journal of Clinical Nutrition* vol. 71 441–449 Preprint at https://doi.org/10.1038/ejcn.2016.198 (2017).
18. Aarts, F. *et al.* Attachment Anxiety Predicts Poor Adherence to Dietary Recommendations: an Indirect Effect on Weight Change 1 Year After Gastric Bypass Surgery. *Obesity Surgery* **25**, 666–672 (2015).
19. Van Hout, G. C. M., Boekestein, P., Fortuin, F. A. M., Pelle, A. J. M. & Van Heck, G. L. Psychosocial functioning following bariatric surgery. *Obesity Surgery* **16**, 787–794 (2006).
20. Mardinoglu, A., Gatto, F. & Nielsen, J. Genome-scale modeling of human metabolism - a systems biology approach. *Biotechnology journal* **8**, 985–96 (2013).
21. Väremo, L., Nielsen, J. & Nookaew, I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic acids research* **41**, 4378–91 (2013).

22. Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1777–82 (2007).

23. Mutch, D. M. *et al.* A distinct adipose tissue gene expression response to caloric restriction predicts 6-mo weight maintenance in obese subjects. *The American journal of clinical nutrition* **94**, 1399–409 (2011).

24. van de Laar, A. W., de Brauw, M., Bruin, S. C. & Acherman, Y. I. Weight-Independent Percentile Chart of 2880 Gastric Bypass Patients: a New Look at Bariatric Weight Loss Results. *Obesity Surgery* **26**, 2891–2898 (2016).

25. Flint, A., Raben, A., Blundell, J. E. & Astrup, A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *International Journal of Obesity* **24**, 38–48 (2000).

26. Bedossa, P. *et al.* Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology* **56**, 1751–1759 (2012).

27. Eeftinck Schattenkerk, D. W. *et al.* Nexfin noninvasive continuous blood pressure validated against Riva-Rocci/Korotkoff. *American Journal of Hypertension* **22**, 378–383 (2009).

28. Dalla Man, C. *et al.* Two-Hour Seven-Sample Oral Glucose Tolerance Test and Meal Protocol: Minimal Model Assessment of -Cell Responsivity and Insulin Sensitivity in Nondiabetic Individuals. *Diabetes* **54**, 3265–3273 (2005).

29. American Diabetes Association, A. D. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2018. *Diabetes care* **41**, S13–S27 (2018).

30. Gourash, W. F., Lockhart, J. S., Kalarchian, M. A., Courcoulas, A. P. & Nolfi, D. Retention and attrition in bariatric surgery research: an integrative review of the literature. *Surgery for obesity and related diseases : official journal of the American Society for Bariatric Surgery* **12**, 199–209 (2016).

31. Li, V. K. M. *et al.* Predictors of gallstone formation after bariatric surgery: a multivariate analysis of risk factors comparing gastric bypass, gastric banding, and sleeve gastrectomy. *Surgical endoscopy* **23**, 1640–4 (2009).

32. Mooy, J. M. *et al.* Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. *Diabetologia* **39**, 298–305 (1996).

33. Abdul-Ghani, M. A. *et al.* The relative contributions of insulin resistance and beta cell failure to the transition from normal to impaired glucose tolerance varies in different ethnic groups. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **1**, 105–112 (2007).

34. Wallace, T. M., Levy, J. C. & Matthews, D. R. Use and Abuse of HOMA Modeling. *Diabetes Care* **27**, 1487–1495 (2004).

35. Mari, A. *et al.* Restoration of normal glucose tolerance in severely obese patients after bilio-pancreatic diversion: Role of insulin sensitivity and beta cell function. *Diabetologia* **49**, 2136–2143 (2006).

36. Polyzogopoulou, E. V, Kalfarentzos, F., Vagenakis, A. G. & Alexandrides, T. K. Restoration of euglycemia and normal acute insulin response to glucose in obese subjects with type 2 diabetes following bariatric surgery. *Diabetes* **52**, 1098–103 (2003).

37. Goldfine, A. B. *et al.* Patients with neuroglycopenia after gastric bypass surgery have exaggerated incretin and insulin secretory responses to a mixed meal. *Journal of Clinical Endocrinology and Metabolism* **92**, 4678–4685 (2007).

38. Bouter, K. E. *et al.* Differential metabolic effects of oral butyrate treatment in lean versus metabolic syndrome subjects. *Clinical and Translational Gastroenterology* **9**, e155 (2018).

39. Chambers, E. S. *et al.* Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. *Gut* **64**, 1744–1754 (2015).

40. Yu, E. W. *et al.* Fecal microbiota transplantation for the improvement of metabolism in obesity: The FMT-TRIM double-blind placebo-controlled pilot trial. *PLOS Medicine* **17**, e1003051 (2020).

41. ter Horst, K. W. *et al.* The FGF21 response to fructose predicts metabolic health and persists after bariatric surgery in obese humans. *Molecular Metabolism* **6**, 1493–1502 (2017).

42. Grong, E. *et al.* Gastrin Secretion After Bariatric Surgery—Response to a Protein-Rich Mixed Meal Following Roux-En-Y Gastric Bypass and Sleeve Gastrectomy: a Pilot Study in Normoglycemic Women. *Obesity Surgery* **26**, 1448–1456 (2016).

5

Self-organized metabotyping of obese individuals identifies clusters responding differently to bariatric surgery

Dimitra Lappa, Abraham S. Meijnikman, Kimberly A. Krautkramer, Lisa M. Olsson, Ömrüm Aydin, Anne-Sophie Van Rijswijk, Yair I.Z. Acherman, Maurits L. De Brauw, Valentina Tremaroli, Louise E. Olofsson, Annika Lundqvist, Siv A. Hjorth, Boyang Ji, Victor E. A. Gerdes, Albert K. Groen, Thue W. Schwartz, Max Nieuwdorp, Fredrik Bäckhed and Jens Nielsen

Plos one, epub ahead of print

ABSTRACT

Weight loss through bariatric surgery is efficient for treatment or prevention of obesity related diseases such as type 2 diabetes and cardiovascular disease. Long term weight loss response does, however, vary among patients undergoing surgery. Thus, it is difficult to identify predictive markers while most obese individuals have one or more comorbidities. To overcome such challenges, an in-depth multi-omics analyses including fasting peripheral plasma metabolome, fecal metagenome as well as liver, jejunum, and adipose tissue transcriptome were performed for 106 individuals undergoing bariatric surgery. Machine leaning was applied to explore the metabolic differences in individuals and evaluate if metabolism-based patients' stratification is related to their weight loss responses to bariatric surgery. Using Self-Organizing Maps (SOMs) to analyze the plasma metabolome, we identified five distinct metabotypes, which were differentially enriched for KEGG pathways related to immune functions, fatty acid metabolism, protein-signaling, and obesity pathogenesis. The gut metagenome of the most heavily medicated metabotypes, treated simultaneously for multiple cardiometabolic comorbidities, was significantly enriched in *Prevotella* and *Lactobacillus* species. This unbiased stratification into SOM-defined metabotypes identified signatures for each metabolic phenotype and we found that the different metabotypes respond differently to bariatric surgery in terms of weight loss after 12 months. An integrative framework that utilizes SOMs and omics integration was developed for stratifying a heterogeneous bariatric surgery cohort. The multi-omics datasets described in this study reveal that the metabotypes are characterized by a concrete metabolic status and different responses in weight loss and adipose tissue reduction over time. Our study thus opens a path to enable patient stratification and hereby allow for improved clinical treatments.

INTRODUCTION

Obesity is generally associated with several different comorbidities, with type 2 diabetes (T2D) and cardiovascular diseases among the most common, and cross interaction of metabolic responses from these co-morbidities makes it difficult to study metabolic alterations associated with obesity. Thus, there is an increasing interest to study heterogeneous diseases like obesity through the collection of multi-omics data from various cohorts[1–3]. Due to the heterogeneity of phenotypes within obese individuals it is, however, generally difficult to stratify cohorts into groups, e.g., individuals with or without the metabolic syndrome, that can be compared using traditional statistical methods when omics data are to be analyzed. The use of machine learning methods is therefore gaining more attention for understanding and deconvoluting multifactorial disease[4,5], in particular as it enables stratification of individuals in a given cohort, without a priori knowledge of clinical labels.

Obesity is a growing worldwide epidemic, with an estimated 1.9 billion adults being overweight and another 650 million being obese[6–8], and it is associated with increased risk of multiple comorbidities including T2D, hypertension, dyslipidemia, non-alcoholic fatty liver disease and various types of cancers[9,10]. Numerous clinical approaches have been proposed to model obesity and predict bariatric surgery outcomes, by using clinical parameters, artificial intelligence and comparing predefined patient groups[11–14]. Another clinical definition for describing individuals with multiple dysmetabolic morbidities, including obesity, is the metabolic syndrome, where obese individuals fulfill two out of these four criteria: 1) fasting glucose >100 mg/dl; 2) triacylglycerol > 150 mg/dl; 3) high-density lipoprotein (HDL) cholesterol <40 mg/dl for males and <50 mg/dl for females; 4) blood pressure above 130 systolic or 85 diastolic[15]. The multitude of co-existing metabolic perturbations may also mask associations between metabolic activities in different tissues, including the gut microbiota, hence posing a challenge in systematically studying obesity, its' implications, and the outcome of surgical intervention with higher resolution. A systems biology approach on the other hand could offer detailed phenotypic profiling possibilities using omics analysis. Metabolomics has recently been proposed as an approach to better comprehend obesity and linked comorbidities[16–18] and identify optimal candidate groups for further interventions[19,20]. The gut metagenome is a contributing factor to the complexity of obesity[21–25], although it's causal role has yet to be established[26]. Recent studies have pinpointed that the production and regulation of metabolites of bacterial origin in humans, play an important role in metabolic diseases[24,27–30]. Given these interactions, there is a clear need to propose a systems biology framework to obesity population-based studies, to improve the identification of distinct sub-populations but also drive the development of personalized interventions[31].

With the objective of getting novel insight into how metabolism in different tissues varies in obese individuals and evaluate if grouping of patients according to metabolism is related to their weight loss response to bariatric surgery, we generated multi-omics datasets from 106 individuals undergoing bariatric surgery. Specifically, we wanted to evaluate if the heterogeneity of a bariatric surgery population can be stratified phenotypically using metabotyping, i.e., grouping according to the individuals fasting plasma metabolome, that captures the functional output of a complex multi-organ system, human hosts and their microbes rather than by traditional clinical classifiers, e.g. the metabolic syndrome. For this we established a novel workflow that first utilizes metabolomics for unlabeled stratification of individuals with several comorbidities and different pharmacological treatment regimens. We then analyzed transcriptome data from liver, jejunum, mesenteric and subcutaneous adipose tissues along with shotgun metagenomic sequencing from fecal samples to produce a discriminatory multi-marker signature of underlying metabolic phenotypes within obesity. The framework is solely based on omics data types of representative of various biological molecule classes (metabolome, transcriptome, metagenome) and machine learning, instead of comorbidities, medications, and disease-specific classifiers, thus making it suitable for studying multifactorial metabolic conditions, besides obesity.

METHODS

BARIA Cohort

The recruitment of participants was conducted from the BARIA study[32] with a total of 106 individuals included. The baseline characteristics of BARIA participants in the Self-Organizing Map (SOM)-defined metabotypes are described in Table 1.

Individuals underwent a complete metabolic work-up at the start of their bariatric surgery trajectory. Anthropometric measurements including height, weight and waist and hip circumference were taken. In addition, body fat percentage using bioelectrical impedance and blood pressure were measured. Fasting blood samples were used for the determination of hemoglobin, HbA1c, glucose, lipid profile, alanine aminotransferase, aspartate amino-transferase, insulin, and creatinine levels. Within three months before surgery, a 2-hour mixed meal tolerance test was performed to assess insulin resistance and investigate dynamic alterations in circulating metabolites. Within three months before surgery, a 2-hour mixed meal tolerance test (MMT) was performed to assess insulin resistance and investigate dynamic alterations in circulating metabolites. The MMT consisted of a compact 125ml drink (Nutricia®) containing in total 23.3 grams fat, 74.3 grams carbohydrates (of which 38.5 grams sugar) and 24.0 grams protein. The participants received this meal after

fasting for a minimum of nine hours. Time point zero refers to the moment at which the participant had fully consumed the meal. Blood samples were drawn *via* an intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes. All samples were stored at -80°C until further processing.

Clinical Metadata	SOM Cluster 1	SOM Cluster 2	SOM Cluster 3	SOM Cluster 4	SOM Cluster 5
Demographic					
Participants (%)	17(16%)	29(27.4%)	25(23.6%)	18(17%)	17(16%)
Female *(% Total Participants, % of SOM Cluster)*	13 (12.2%, 76.5%)	25 (23.6%, 86.2%)	18 (17%, 72%)	14 (13.2%, 77.8%)	14 (13.2%, 82.4%)
Male *(% Total Participants, % of SOM Cluster)*	4 (3.8%, 23.5%)	4 (3.8%, 13.78%)	7 (6.6%, 28%)	4 (3.8%, 22.22%)	3 (2.8%, 17.6%)
Anthropometric					
Age *(years)*	48(29-60)*	40(20-57)*	53(26-64)*	56(39-64)*	44(22-62)*
BMI *(kg/m²)*	39.5(34-45.4)	38.2(32.9-60.9)	39.8(33-57.5)	38.3(33.8-47.1)	39.8(34.7-46.4)
Waist circumference *(cm)*	125.3 ± 12.6	122.6 ± 12.3	123.7 ± 11.5	125.8 ± 12.2	123 ± 9.9
Upper thigh circumference *(cm)*	135(120-149)	133(116-147)	130(103-165)	133(115-139)	136(123-144)
Total Body Fat *(%)*	53.6(41.6-64.7)	54.1(31.7-94.9)	51.8(39.3-104.8)	56.5(40.6-78.9)	57.6(44-64.5)
Fat Free Mass *(kg)*	60.9(54.1-93.8)	59.6(50.3-90.6)	59.1(47.5-90.2)	59.8(49.5-85.1)	60.8(54-83.5)
Systolic blood Pressure *(mmHg)*	131.5(116-156)	132(102-155)	133(108-161)	136(115-193)	135(115-157)
Diastolic blood Pressure *(mmHg)*	84.5(59-91)	81(54-99)	82(67-105)	80(45-121)	82(65-94)
Clinical lab values					
Fasting glucose *(mmol/l)*	5.8(4.8-11.4)	5.9(4.6-14.8)	5.7(5-13.8)	5.8(4.6-6.8)	5.6(4.5-8.7)
HbA1c *(mmol/mol)*	5.7(5.3-9.1)	5.7(4.6-9.8)	5.6(5-9.3)	5.8(5.2-6.9)	5.5(5.2-8.3)
HOMA-IR	1.7(0.6-3.4)	1.6(0.5-6.9)	2.2(0.5-4.7)	1.3(0.8-4.8)	1.5(0.8-4.8)
HOMA2-β	108.7 (38.3-183.2)	87.9 (29.1-227.8)	112 (52.7-226.2)	92.1 (52.4-357.8)	104.2 (50.8-185.5)
Total Cholesterol *(mmol/l)*	5.4 ± 1.1	4.6 ± 1	4.9 ± 1.1	5.3 ± 1.2	4.3 ± 0.9
Triglycerides *(mmol/l)*	1.5(0.8-3.5)	1.3(0.6-5.8)	1.4(0.8-6)	1.4(0.8-5.9)	1.2(0.6-1.9)
HDL Cholesterol *(mmol/l)*	1.2(0.8-1.8)*	1.1(0.6-1.9)*	1.1(0.7-2.5)*	1.2(0.7-2.1)*	1.2(1-2.7)*
LDL Cholesterol *(mmol/l)*	3.6 ± 1.1	2.9 ± 0.9	3.6 ± 0.9	3.4 ± 1.7	2.6 ± 0.8
Creatinine *(mmol/l)*	68(55-96)	63(46-83)	66(47-112)	75(56-172)	65(58-99)
Glomerular Filtration Rate *(kl/1.73m2)*	85(70-91)*	90(71-91)*	86(62-91)*	78(26-90)*	89(66-91)*

Table 1. Baseline characteristics of 106 BARIA participants included. Data is expressed as mean ± standard deviation. Categorical variables are presented as numbers and percentages. Non-normally distributed variables are presented as median with interquartile range. For comparison among metabotypes Kruskal-Wallis test (extended Mann-Whitney U-test for multiple groups) was used. '*' denotes differentially significant variables among the five metabotypes clusters ($P<0.05$). BMI: Body Mass Index, HbA1c: Hemoglobin A1c, HOMA-IR: Homeostatic Model Assessment of Insulin Resistance, HOMA-β: Homeostatic Model Assessment of beta-cell function, LDL: Low-Density Lipoprotein, HDL: High-Density Lipoprotein.

Metabolome Analysis

EDTA plasma samples under fasting conditions were collected from 106 BARIA participants. Samples were shipped to METABOLON (Morisville, NC, USA) for performing analysis using ultra high-performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) untargeted metabolomics, as previously described[27]. The metabolomic counts obtained, underwent significant curation via metabolites' pre-filtering, imputation for subsets of metabolites' missing values and data normalization, in order to minimize the effect of artifacts in the downstream analysis. Out of 1345 metabolites measured by METABOLON, 652 metabolites were fully detected across all samples, 640 metabolites were partially detected across all samples, and 53 metabolites were not detected or failed to reach detection limit and therefore had a missing value. Metabolomics prefiltering and imputation were performed by utilizing a variation of the Perseus platform[33]. Essentially, data has been pre-filtered to have a maximum of 25% missing values for a metabolite across all samples. This was followed by a log transformation of all the measured metabolites' raw intensities across the entire dataset. Then, we calculated the total data mean and standard deviation (by omitting missing values). Considering that the metabolite intensities distribution is approximately following normality, we chose a small distribution 2.5 standard deviations away from the original data mean towards the left tail of the original data distribution, with 0.5 standard deviations width. This new shrunken range corresponds to the actual lowest level of detection by the spectrometer. Here by drawing random values from this mini distribution, we filled the missing prefiltered data of choice.

Normalization was conducted to the total signal for each sample, since each sample is a separate injection on the mass spectrometer. Effective control for changes in sample matrix affects ionization efficiency, hence there can be inevitable differences in how much each sample is loaded onto the column with each injection, etc. Therefore, we summed up the total ion intensity (i.e. total signal) for each of the samples and identified the sample with the lowest total signal. After this we could proceed to calculating the correction factor for each sample by dividing the total signal with the lowest total signal. After imputation and normalization, we obtained 986 metabolites. All the calculations for imputing and normalizing the metabolomics dataset have been conducted with MATLAB_R2018b and the standard built-in packages.

Differential analysis was conducted among the five SOM-defined Clusters in R (version 3.6.3) and RStudio (version 1.2.5033). Statistical analysis has been performed for fasting peripheral plasma with two methods: ANOVA (Analysis of Variance) and Kruskal Wallis test, with the use of HybridMTest package[34]. HybridMTest performs hybrid multiple testing using empirical Bayes probability. The significance level and cut-off used for the dataset of fasting peripheral plasma was $P<0.05$ and was applied to metabolites that were significantly differential with both ANOVA and Kruskal Wallis methods.

Clustering Metabolome Profiles with Self- Organizing Maps

The fasting peripheral metabolomics were then input to the SOM toolbox[35] algorithmic setup in MATLAB_R2014b. SOMs conducted unsupervised competitive learning and produced low-dimensionality visualizations by employing vector quantization[36,37], a topology preserving projection. SOMs are essentially networks consisting of neurons in a lower dimensional space than the initial dataset, visually represented in a 2-dimensional grid. Each neuron has d-dimensions, equal to the number of features of the dataset and acts as a weight vector. During the SOM training phase, the weight vectors are gradually shifted in each iteration of learning, and the map gradually gets organized, so that neurons that are neighbors on the grid get similar weight vectors throughout the iterative training.

In our analysis, SOM took as input a set of prototype vectors representing the data. Every data item, here BARIA subject's fasting metabolome, was mapped into one point (node or neuron) in the map[38]. Mapping took place throughout the training phase of the SOM. The number of nodes was calculated internally by a heuristic formula, given the number of input vectors and their dimensionality, as ~, where is the number of data items and the number was slightly altered in order to fit hexagonal (instead of rectangular) nodes. The training method deployed in our study was batch training, where instead of taking each input vector separately and assigning a weight vector, the dataset was given to the SOM as a whole and the new weight vectors are weighted averages of the data vectors. In order to assign the prototype vector to the node, the Euclidean distances among prototype vectors and each neuron were calculated and set as the metric for the similarity measure. The "winner" node in the grid, was the one with the smallest Euclidean distance from the input vector. Once the assignment was complete, then the weights of the prototype vector along with the weights of the subset of its spatial neighbors in the array, got updated[39,40]. This entailed that all these local re-arrangements would be propagated along the grid, during the training epochs. As a result, after learning, more similar data items would be associated with nodes that are closer in the array, whereas less similar items would be situated gradually farther away in the array.

When having a very large number of SOM nodes, one cannot easily quantify the results, hence the need for further grouping with a partitive approach. The resulting map was then subjected into clustering, as a built-in function of the SOM toolbox, for obtaining a recommended partition of map nodes. An open question in this case was the number of clusters, since in general takes this as a predefined parameter. Since this is sensitive to initialization, we ran a cross validation simulation for 100 times for each (starting from ~, which corresponds to the number of nodes of the neural network to 1 with step of -1) for each with different random initializations. The best partitioning for each number of clusters was selected using error criterion and the minimization of the Davies-Bouldin cluster validity

index[41]. Davies-Bouldin index is a metric of the ratio of the within cluster scatter, to the between cluster separation. The index's value is essentially the average similarity between each cluster and its most similar one, averaged over all the clusters. This implies that the best clustering scheme minimizes the Davies-Bouldin index. Eventually, when all the iteration for the potential values of were concluded, the minimum overall Davies-Bouldin index was chosen, which resulted in the recommended partition of five clusters.

Transcriptome Analysis

Biopsies from liver (106 samples), jejunum (105 samples), mesenteric adipose fat (104 samples) and subcutaneous adipose fat (105 samples) were collected at the time of the bariatric surgery, as previously described[32]. RNA was extracted from biopsies using TriPure Isolation Reagent (Roche, Basel, Switzerland) and Lysing Matrix D, 2 mL tubes (MP Biomedical, Irvine, CA, USAs) in a FastPrep®-24 Instrument (MP Biomedical, Irvine, CA, USAs) with homogenization for 20 seconds at 4.0 m/sec, with repeated bursts until no tissue was visible; homogenates were kept on ice for 5 minutes between homogenization bursts if multiple cycles were needed. RNA was purified with chloroform (Merck, Darmstadt, Germany) in phase lock gel tubes (5PRIME) with centrifugations at 4°C, and further purified and concentrated using the RNeasy MinElute kit (Qiagen, Venlo, The Netherlands). The quality of RNA was analysed on a BioAnalyzer instrument (Agilent), with quantification on Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA). Due to degradation of the RNA, libraries for RNAseq sequencing were prepared by rRNA depletion; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end reads and 10G data/sample. The average read count per sample from liver and jejunum tissues were 42 ± 15 million. For mesenteric and subcutaneous fat, the average read count per sample were 43.2 ± 20 million.

The extracted fastq files were analyzed with nf-core/rnaseq[42], a bioinformatics analysis pipeline used for RNA sequencing data. The workflow processed raw data from FastQ inputs (FastQC, TrimGalore!), aligned the reads (STAR) with *Homo sapiens* GRCh38 as reference genome, generates gene counts (featureCounts, StringTie) and performed extensive quality-control on the results (RSeqQC, dupRadar, Preseq, edgeR, multiQC). The pipeline was built using Nextflow.

Differential gene expression analysis for five SOM defined cluster participants has been performed for liver, jejunum, subcutaneous adipose and mesenteric adipose tissues, respectively, in R (version 3.6.3) and RStudio (version 1.2.5033) with DESeq2[43] package. The statistical analysis method for calculating differential expression rates was the LRT test (log-ratio test). After FDR correction with multiple hypothesis testing with IHW[44] package, we analyzed genes with $P<0.05$ by DEGreport's[45] degPatterns function, so as to

identify subgroups of co-expressed genes across the SOM clusters and assign a *z score* to each metabotype. For these differentially significant co-expressed genes we performed gene enrichment with Enrichr platform[46] using KEGG (Kyoto encyclopedia of genes and genomes) metabolic pathways[47].

Microbiome Analysis

Fecal samples from 106 participants (108 fecal samples due to having two samples from two participants) were collected on the day of surgery and immediately frozen at -80C. Total fecal genomic DNA was extracted from 100 mg of feces using a modification of the IHMS DNA extraction protocol Q[48]. Briefly, fecal samples were extracted in Lysing Matrix E tubes (MP Biomedical, Irvine, CA, USA) containing ASL buffer (Qiagen), and lysis of cells was obtained, after homogenization by vortexing for 2 minutes, by two cycles of heating at 90°C for 10 minutes followed by three bursts of bead beating at 5.5 m/sec for 60 seconds in a FastPrep®-24 Instrument (MP Biomedical, Irvine, CA, USAs). After each bead-beating burst, samples were placed on ice for 5 minutes. The supernatants containing fecal DNA were collected after the two cycles by centrifugation at 4°C. Supernatants from the two centrifugations steps were pooled and a 600 µL aliquot from each sample was purified using the QIAamp DNA Mini kit (Qiagen, Venlo, The Netherlands) in the QIAcube (Qiagen, Venlo, The Netherlands) instrument using the procedure for human DNA analysis. Samples were eluted in 200 µL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0). Libraries for shotgun metagenomic sequencing were prepared using a PCR-free method; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end reads and 6G data/sample.

MEDUSA[49] pipeline was used for pre-processing of raw shotgun metagenomics sequence data. MEDUSA is an integrated pipeline for analysis of short metagenomic reads, which maps reads to reference databases, combines output from several sequencing runs and manipulates tables of read counts. The input number of total reads from the metagenome analysis were on average 23.4 ± 2.2 million reads per sample and the total aligned reads 16.6 ± 1.8 million reads per sample. The sequencing runs had high quality with almost 98% of the reads passing the quality cut-off (~(20 million reads per sample) . Out of the high-quality reads, on average 0.04% aligned to the human genome, although the data had been cleaned for human reads. Out of the high quality non-human reads, 78.4% aligned to MEDUSA's software gene catalogue. Quality filtered reads were mapped to a genome catalogue and gene catalogue using Bowtie2[50]. Statistical analysis was performed in R (version 3.6.3) and RStudio (version 1.2.5033) on rarefied count, (20 million reads per sample). The taxon ids were input to taxize[51] package, so as to get full taxonomic information and ranking for

the species. This dataset was input to DESeq2[43] and phyloseq[52] packages for conducting downstream differential statistical analysis. Similar to the BARIA transcriptomics counts, log normalization has been conducted based on gene counts geometric distribution. Statistical analysis test for calculating differential expression rates was LRT. The IHW package, as part of DESeq2 suite, is utilized for multiple hypothesis testing and adjusting the respective P values, with alpha significance threshold set at $P<0.05$.

DIABLO Correlation Analysis and Biomarkers minimal signature

DIABLO53 stands for Data Integration Analysis for Biomarker discovery using Latent cOmponents and performs supervised multi-omics data integration, by maximizing the correlation between co-expressed elements in the input datasets. DIABLO algorithm extends sparce Generalized Canonical Correlation Analysis54 and by expanding the Partial Least Squares (PLS) regression, used singular value decomposition for dimensionality reduction and selected co-expressed (correlated) variables that could explain the categorical outcome of interest, in our case the five SOM-derived metabotypes. DIABLO analysis was conducted in R (version 3.6.3) and RStudio (version 1.2.5033) through the package of mixOmics55 (version6.10.9). DIABLO output a set of latent variables (components) based on the dimensionality and the importance of the input datasets. All the datasets in this study carried the same weight, hence the DIABLO dataset matrix initialization design parameter was diagonal. The original input was 289 metabolites, 119 microbial species and 776 genes, all the differentially identified components from the omics datasets. This chosen number of components could extract sufficient information to discriminate all SOM-defined metabotypes. Then, a set of coefficients was attributed to each variable, that indicated the importance of each variable in DIABLO. The goal was to have maximization of the covariance between a linear combination of the variables from each input dataset and each categorical outcome. The algorithm was optimized with a 10-fold validation over 10 training epochs. After tuning these two hyperparameters (number of variables from each dataset, choice of variables that maximize co-variance), DIABLO produced as output a minimal signature of total 113 markers that distinguish the given metabotypes.

RESULTS

Metabolomics based stratification of bariatric surgery population via SOM

To create a multi-omics profile of obesity, a total of 106 individuals from the BARIA[32] bariatric surgery cohort were recruited. The multi-omics analysis included metabolomics on fasting peripheral blood samples, and we employed this dataset for stratification of the

heterogenous group of individuals, independent of traditional clinical indexes, such as Body Mass Index (BMI), hypercholesterolemia, hypertension and treatment for T2D. To enable stratification based on the metabolomics data we built an unsupervised artificial neural network that could group individuals based only on the similarity of their metabolome, a SOM. The SOM[36] evaluated metabolomic similarity by calculating the Euclidean distances between complete metabolomic profiles, and projected BARIA individuals with high inter-group similarity onto a "map" of lower dimensionality compared to that of the initial dataset. The SOM was trained with 106 prototype vectors, where each prototype vector corresponded to a BARIA participant's peripheral plasma metabolite profile, consisting of 986 metabolites (see **Methods**). Iterative training of the SOM resulted in a map of 48 nodes, all projected onto a hexagonal grid (Fig 1A). These 48 nodes considerably reduced the dimensionality and further within-cluster variance was minimized using, which preserved metabolomic distances and identified centroids of core metabolomes[56].

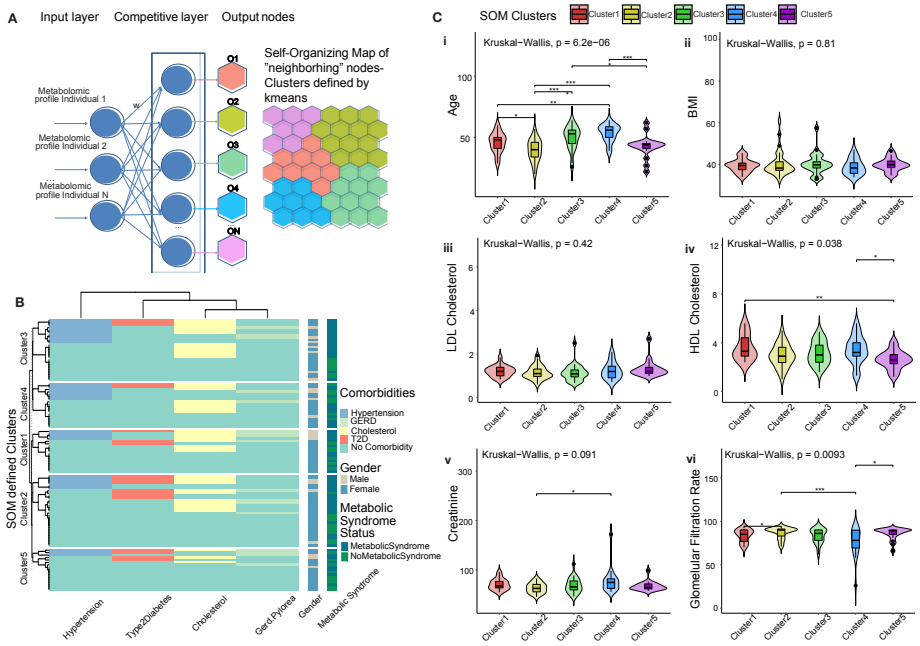


Figure 1. Self-organizing maps reveal five distinct metabotypes within BARIA cohort. (A) Architecture of a competitive artificial neural network. Each individual's complete metabolomic profile is assigned a weight. The weights are in turn assigned to neurons in the competitive layer of the neural network. In the competitive layer, SOM algorithm calculates the similarity metric (here Euclidean distance) between each metabolomic profile and the neurons and then updates the weights. After training, the network assigns the individual's metabolomic profile to the "winner" output node, the node that is essentially more similar to the input metabolomic profile. Once this step is complete, all the nodes are comprising the SOM. Finally, all the nodes of the SOM are subjected to k-means clustering resulting in the partitioned topology, the metabotypes (SOM & k-means defined clusters). **(B)** Clustergram of hierarchical cluster analysis depicting the distribution of medically treated cardiometabolic comorbidities of the individuals in each of the metabotypes (SOM & k-means defined clusters). The treated comorbidities are: hypertension, T2D, GERD and cholesterol. In parallel columns are the gender and metabolic syndrome status of each individual, respectively.

(C) Clinical variables associated with obesity and their statistical significance across the metabotypes (SOM & k-means defined clusters): age **(C. i)**, BMI **(C. ii)**, HDL cholesterol **(C. iii)**, LDL cholesterol **(C. iv)**, creatinine and **(C. v)**, glomerular filtration rate (c. vi); statistical significance among metabotypes is calculated with Kruskal-Wallis test; the symbols indicating significance among metabotypes are '*': P<=0.05, '**': P<=0.01, '***': P<=0.001.

SOM and k-means clustering reveal five distinct metabotypes

Clustering the SOM identified five clusters (metabotypes), each with different features (Fig 1A, Table 1), including unique distributions of comorbidities and medication usage. Polypharmacy is a notable characteristic within this study population, including use of medication for T2D (n = 20), hypertension (n = 30), hypercholesterolemia (n = 42) and gastroesophageal reflux disease (GERD, n = 16). Medication usage was distributed across the five metabotypes and is shown in Fig 1B. Clusters 1 and 3 include most individuals simultaneously treated for hypertension and high cholesterol (four and five individuals respectively), whereas cluster 2 includes individuals co-treated for hypertension, high cholesterol and T2D (four individuals). Nevertheless, the distribution of overlapping treated cardiometabolic comorbidities is quite uniform among clusters and not skewed towards a particular metabotype. In order to assess the effect of missing values and data imputation in SOM clustering, a separate mapping analysis was conducted by using the unimputed metabolomics dataset. The final map clustering did not diverge from the original prediction. Hence, the ability of the SOM to assign similar items on the same node was not affected by the imputation of a minimal set of missing metabolite values. We next assessed the biometric features of each metabotype, by performing differential analysis on the clinical variables available. BMI, body fat, and waist circumference did not significantly differ between clusters, however age, HDL cholesterol and glomerular filtration rate varied between clusters (Fig 1C). Given that all BARIA participants are affected by severe obesity, the stratification based on their SOM metabolomic profile reveals that BMI and treatment of cardiometabolic comorbidities are not the clinical features more accurately describing and differentiating the metabotypes, but age, cholesterol and markers associated with kidney function are important features. We also evaluated if there is any association within the clusters and individuals having the metabolic syndrome and found that there was no such association (Fig 1B). Furthermore, if we grouped the individuals according to having or not-having the metabolic syndrome, we also found no separation according to age or other clinical parameters besides those defining the metabolic syndrome (S1 Fig).

Metabolomic profiles characterized by lipid and amino acid metabolites

Following stratification of the individuals into the five metabotypes we performed differential analysis of the metabolome for the five different metabotypes. Statistical analysis revealed 289 differentially significant metabolites. In comparison we only identified 3 differentially significant metabolites when the cohort was grouped according to presence of

the metabolic syndrome or not (S2 Fig), which shows that driving grouping of the cohort based on the metabolomics data enables more detailed insight into metabolic differences among the individuals. KEGG pathway analysis revealed that the most highly enriched metabolite classes among the 289 metabolites were lipids, amino acids and xenobiotics, followed by cofactors and vitamins, nucleotides, carbohydrates, peptides, energy and partially characterized molecules. Clusters 2 and 3 exhibited the highest relative abundance of differentially significant metabolites, mainly lipids and amino acids (Fig 2A).

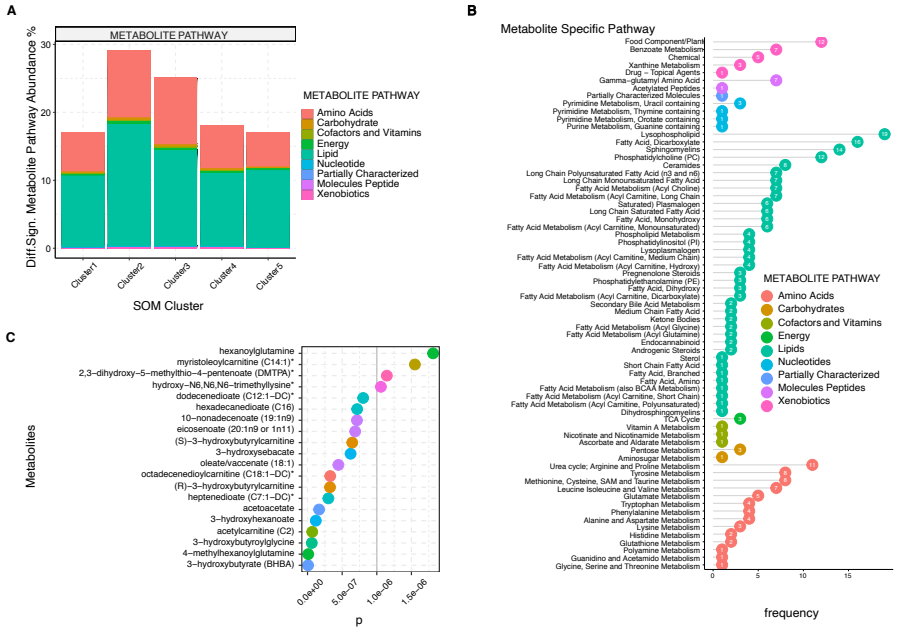


Figure 2. Differentially abundant metabolites and metabolic pathways among the five defined SOM clusters (metabotypes). (A) Relative abundance and distribution of differentially significant metabolites among SOM and k-means defined clusters. Clusters two and three are most abundant in lipids (especially lysophospholipids and sphingomyelins) and amino acids (urea, arginine and proline metabolism). **(B)** Distribution of differentially significant metabolic pathways among SOM and k-means defined clusters, where numbers within each dot indicate how many metabolites of that particular specific pathway were differentially abundant across clusters. **(C)** Top 20 differentially significant metabolites among the SOM and k-means defined clusters, (P<0.05).

Among the enriched KEGG metabolic pathways that had the highest number of differentially significant metabolites were fatty acids (Fig 2B), specifically 19 lysophospholipids, 16 dicarboxylate fatty acids, 14 sphingomyelins and 12 phosphatidylcholines. The amino acid metabolic pathways with the most significant metabolites were arginine and proline metabolism with 11 compounds, tyrosine metabolism with 8 metabolites, methionine, cysteine SAM and taurine metabolism with 8 metabolites, too, while branched-chain amino

acid metabolism for isoleucine and valine had 7 metabolites. The top 20 differentially abundant metabolites are a mixture of lipids, partially characterized molecules, peptides and amino acids, and some of them, despite being the end product of endogenous ketogenesis produced by the liver, also carry the potential of being the result of gut microbial metabolism, such as 3-hydroxybutyrate and acetoacetate (Fig 2C). Our analysis identified lipid metabolites (especially lysophospholipids and sphingomyelins) and amino acid metabolites (urea, arginine and proline metabolism) being significantly altered among the clusters.

Hepatic and adipose tissue transcriptomes enriched for immune, amino acid and lipid metabolism functions

To better understand the relationship between metabolite levels and gene expression, we next sequenced RNA extracted from biopsies taken during bariatic surgery from liver, jejunum, mesenteric adipose tissue and subcutaneous adipose tissue. We identified differentially expressed genes between clusters and conducted gene set enrichment analysis. This analysis revealed 682 hepatic genes differentially expressed across the five metabotypes. In contrast, only four genes were differentially expressed in jejunum, whereas 45 and 49 genes were differentially expressed in mesenteric and subcutaneous adipose tissue, respectively. These liver, mesenteric and subcutaneous adipose tissue gene sets were subjected to enrichment analysis for retrieving their functional profiles (S1-4 Tables). Due to the low number of differentially expressed genes in jejunum, we were unable to obtain a gene set enrichment signature for jejunum tissue. The top represented pathways in the liver included fatty acid elongation /saturation reflecting lipids in the plasma, glycan and sphingolipid biosynthesis, cell function regulation (ErbB signaling pathway, protein export) and immune responses (Fig 3A).

The mesenteric adipose tissue was enriched for amino acid metabolic processes (Fig 3B) reflecting amino acids in the plasma, and subcutaneous adipose tissue was found enriched in many pathways related to pathogens (Fig 3C) and may reflect increased immune activation associated with metabolic disease. To investigate how these pathways are regulated across the five metabotypes, we examined the normalized gene expression levels of differentially expressed genes among the clusters. The metabolic pathways enriched within the hepatic transcriptome exhibited mixed directionality in regulation and were assessed individually, for each metabotype (Fig 3D). Amino acid metabolic pathways in mesenteric adipose tissue exhibited consistent upregulation in clusters 4 and 5 (Fig 3D). Transcriptome analysis from these three tissues showed distinct regulation of lipid, amino acid, immune response and pathogenic pathways amongst the metabotypes.

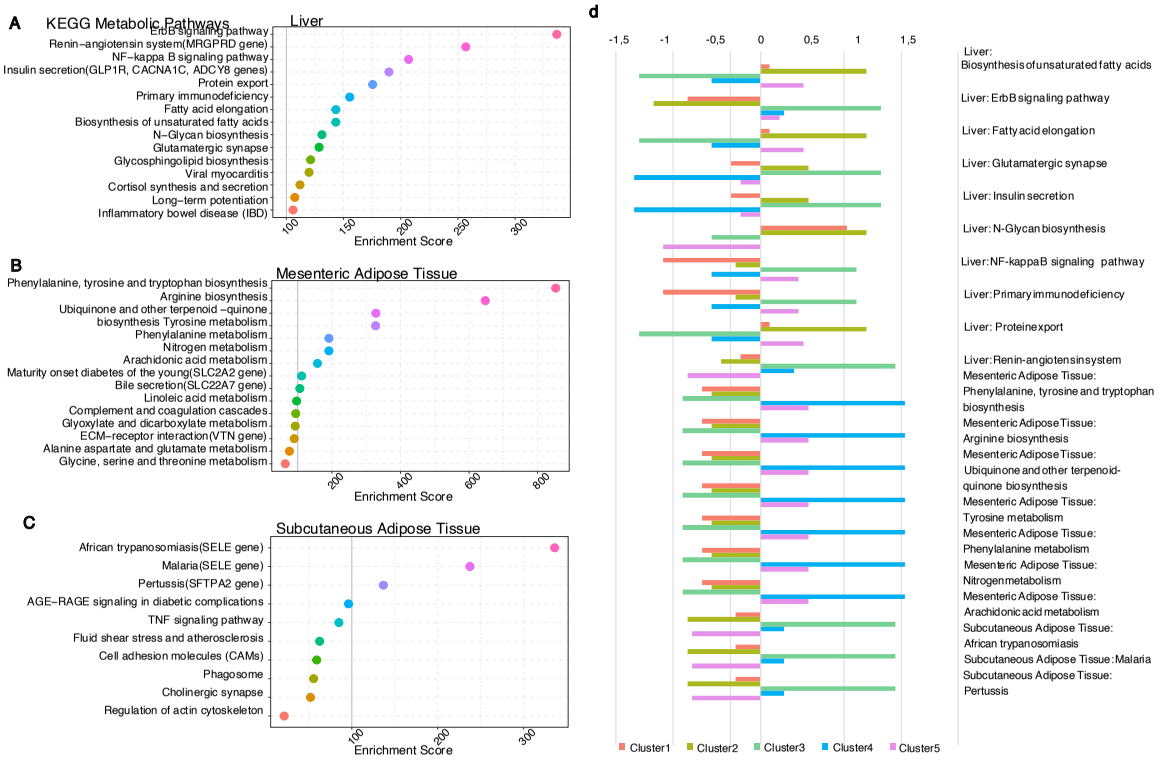


Figure 3. Differentially enriched KEGG metabolic pathways among the five defined SOM clusters (metabotypes). (A) Top 15 differentially enriched KEGG metabolic pathways for hepatic transcriptome among the SOM and k-means defined clusters, ranked based on their scores after differential gene expression analysis (DESeq2, P<0.05) and gene set analysis (GSA with EnrichR). **(B)** Top 15 differentially enriched KEGG metabolic pathways for mesenteric adipose transcriptome among the SOM and k-means defined clusters, ranked based on their scores after differential gene expression analysis (DESeq2, P<0.05) and gene set analysis (GSA) with EnrichR). **(C)** Top 10 differentially enriched KEGG metabolic pathways for subcutaneous adipose tissue transcriptome among the SOM and k-means defined clusters, ranked based on their scores after differential gene expression analysis (DESeq2, P<0.05) and gene set analysis (GSA with EnrichR). **(D)** 20 highest scoring KEGG metabolic pathways according to EnrichR GSA score for liver, mesenteric adipose and subcutaneous adipose tissues. Z score indicates different levels of differentially expressed pathways, for each SOM and k-means defined cluster.

Metabotypes exhibit distinct microbial community composition

Since the gut microbiota is known to be correlated with development of comorbidities linked to obesity[57–59], we also generated a gut microbiota profile for the BARIA individuals from shotgun metagenomic sequencing of fecal DNA. Statistical analysis revealed 119 differentially abundant species among the SOM metabotypes, the top 30 of which are shown in Fig 4A and are dominated by Bacteroidetes and Firmicutes, especially *Lactobacillus*. Out of the 119 differentially abundant species, 70 belonged to Firmicutes phylum, 22 to Bacteroidetes, 11 to Actinobacteria, 11 to Proteobacteria, one to Chloroflexi, one to Cyanobacteria, one to Euryarchaeota, one to Spirochaetes and one to Fusobacteria. Within Firmicutes, *Clostridia* are more highly abundant for cluster 2 and *Weisella* for clusters 2, 4 and 5. Within Bacteroidetes, *Bacteroides* and *Prevotella* species are significantly more abundant in clusters 1, 2 and 3. For Actinobacteria, *Bifidobacterium* are considerably more abundant in 1 and 2, whereas species within Enterobacteriaceae family have higher abundance for clusters 4 and 5 (Fig 4B, S5 Table). In order to assess if there is a difference in alpha and beta diversity among metabotypes, we used a series of metrics (Observed, Chao1, ACE, Shannon, Simpson, Inverse Simpson for alpha diversity and Whittaker index along with dispersion analysis for beta diversity), shown in S3 Fig and S4 Fig. Our metagenomics pipeline displayed that none of the different alpha diversity metrics reach statistical significance. The beta dispersion (with centroids) results coupled with a permutational ANOVA (PERMANOVA) analysis (for 999 permutations) showed F=0.19 and P=0.9431. As seen in S4 Fig, the SOM-defined clusters largely spatially overlap but appear to have different centroids and different dispersions. Nevertheless, the large inter-individual variation cannot account for the negative PERMANOVA results, either. In such cases, there is a need to have a correct specification of the mean-variance relationship by means of multivariate extensions of GLM with methods such as negative binomials, DESeq2[43]. The DESeq analysis revealed that despite the non-statistically significant diversities, there are SOM-defined clusters that are enriched in specific genera, such as *Bacteroides, Prevotella* and *Lactobacillus.*

As comparison, when the patients were grouped after presence or absence of metabolic syndrome, we only identified 54 differentially significant species. Similarly, none of the alpha diversity or beta diversity metrics or ordination were statistically significant. (S5 Fig, S6 Fig). Out of the 54 significant gut microbial species 33 species belonged to Firmicutes phylum, 10 to Bacteroidetes, six to Actinobacteria, four to Proteobacteria and one to Fusobacteria (S7 Fig, S8 Table). Within Firmicutes, there is a trend for *Lactobacillus* species to be two to 8 times significantly less abundant in metabolic syndrome BARIA individuals. In contrast, statistically significant *Streptococcus* species are twice as abundant in metabolic syndrome diagnosed individuals. The majority of the gut microbial species belonging to Bacteroidetes is two to three times depleted in metabolic syndrome diagnosed BARIA individuals,

whereas Actinobacteria levels are elevated in metabolic syndrome. Differentially significant Proteobacteria tend to be depleted in metabolic syndrome.

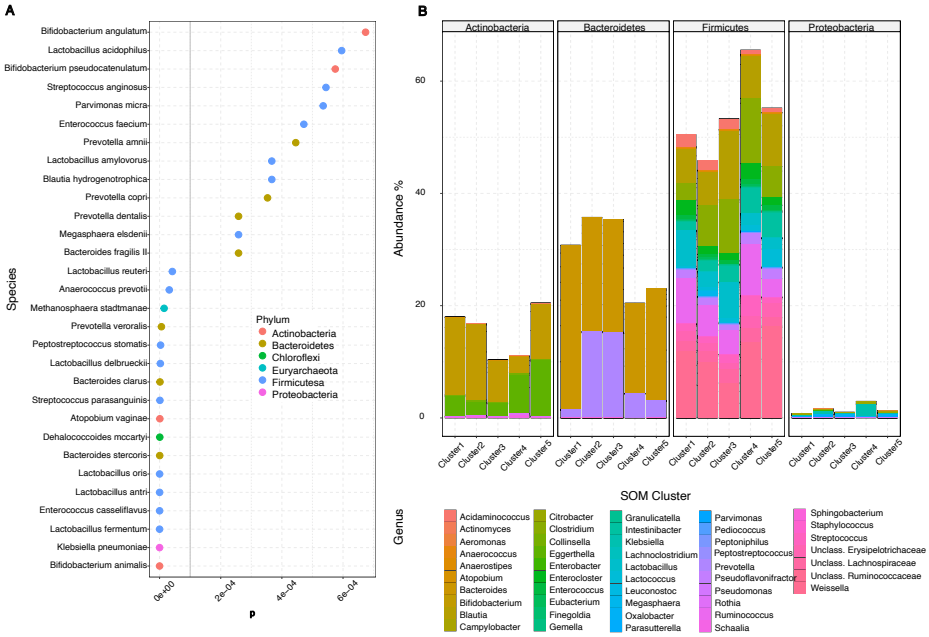


Figure 4. Differentially significant microbial species and phyla among the five defined SOM clusters (metabotypes). (A) Top 30 from 119 differentially significant microbial species among the SOM & k-means defined clusters, after differential analysis with DESeq2 (P<0.05). **(B)** Relative abundance and distribution of differentially significant microbial species for the top 4 most abundant phyla for SOM & k-means defined clusters.

Our analysis showed that metabolic syndrome diagnosis can indeed capture a fraction of the microbial variability within obesity. Even so, our suggested metabotyping approach can identify more gut microbial species across the spectrum of obesity and its related comorbidities.

Individual metabotypes display unique clinical and multi-omics features

Our collective analyses show that the five different metabotypes clearly associate with unique gene expression and microbial community composition patterns and hence represents groups of individuals having distinct differences in their metabolism. To further explore these unique patterns, we next performed a detailed evaluation of the molecular fingerprints of each metabotype using the findings from the multi-omics differential analysis.

17 individuals had metabotype 1 (13 women/four men), and they had the highest fat free mass 60.9 (54.1-93.8) kg and the highest total cholesterol (5.4 ± 1.1 mmol/L). Of

these, four participants were treated for hypertension, three for T2D and four for GERD, whereas almost half the cluster's population (8 participants) was treated for high cholesterol. It is noticeable that three out of four male participants were co-treated for hypertension, GERD/*H. pylori* infection and cholesterol (Fig 1B). Isobutyrylcarnitine was at a higher level in this metabotype (see S6 Table) compared with the other metabotypes, and the same was observed for the tyrosine metabolic pathway intermediate 4-methoxyphenol. When associating the differentially significant fasting metabolites with anthropometric features (S8 Fig, S9 Fig), we observed negative correlations between sphingomyelins, fasting glucose ($r=-0.8, P<0.001$), HbA1c ($r=-0.6, P<0.01$) and age ($r=-0.5, P<0.01$) specifically for this metabotype. In summary, metabotype 1 was characterized by high cholesterol, males using medication, downregulation of immune response pathways in the liver, lower abundance in *Prevotella* and higher abundance in *Bacteroides* (Fig 4B) compared to other metabotypes.

Metabotype 2 was the largest cluster consisting of 29 participants. It was female dominated (25 females/four males) and had the youngest individuals of 40 (20-57) years of average age with a BMI of 38.2 (32.9-60.9) kg/m^2. The highest number of T2D individuals was noted here (n= 8), with the highest mean HbA1c value at 42 ± 12 mmol/mol. The individuals were the most heavily medicated, since it contained 11 individuals with treatment for dyslipidemia, 6 with hypertension along with the individuals affected by T2D, of which four participants were treated for all conditions simultaneously. When considering the metabolome, lysophospholipids, 1-arachidonoyl-GPC*(20:4)*, 1-linoleoyl-GPC(18:3)*, 1-linoleoyl-GPE(18:2)*, 1-oleoyl-GPE(18:1), 1-palmitoyl-GPC(16:0)* and 1-stearoyl-GPC(16:0) were higher in comparison to the majority of the clusters. Similarly, branched-chain amino acid (BCAA) metabolites 1-carboxyethylvaline, 1-carboxyethylisoleucine and valine were all at elevated levels. 3-hydroxyoleoylcarnitine, 3-hydroxydecanoate and 2-hydroxybutyrate-2-hydroxyisobutyrate were positively correlated with both glucose and HbA1c ($r=0.5, P<0.001$) (S8 Fig, S9 Fig). To summarize, metabotype 2 represented the youngest individuals, yet the individuals being most heavily medicated for comorbidities. The individuals have high abundance of BCAAs and hydroxy fatty acids, even though fatty acids biosynthetic pathways were downregulated in mesenteric adipose fat. In the gut microbiome *Prevotella, Bacteroides* and *Lactobacillus* species were found to be highly abundant (Fig 4B).

There were 18 individuals having metabotype 3 and this metabotype had the highest percentage of males among the clusters (11 females/7 males). The individuals exhibited the highest HOMA2-IR at 2.2 (0.5-4.7) and the highest HOMA2-β at 112 (52.7-226.2). It included three individuals with T2D (out of 6 in total treated for T2D), 10 hypertensive (out of 13 in total treated for hypertension) and 12 individuals treated for dyslipidemia (out of 15 in total treated for high cholesterol), whereas three were treated for T2D, hypertension and dyslipidemia at the same time. Even though the anthropometrics differed, the individuals

of metabotype 3 had similar metabolome and microbiome profiles as metabotype 2, but with varying transcriptome patterns. Similar to metabotype 2, lysophospholipids, 1-arachidonoyl-GPC* (20:4)*, 1-linoleoyl-GPC (18:3)* , 1-linoleoyl-GPE (18:2)*, 1-oleoyl-GPE (18:1), 1-palmitoyl-GPC (16:0)* and 1-stearoyl-GPC (16:0) were detected in equally high levels in the individuals of metabotype 3. Noticeably, all sphingomyelins were elevated for individuals in this metabotype (S6 Table). Cluster 3 appeared to be the most insulin resistant and most treated for dyslipidemia, in spite of the highly abundant metabolome in lysophospholipids and sphingomyelins. In essence, hepatic upregulation of immune responses and subcutaneous adipose tissue upregulation of pathogenic-related pathways (Fig 3D), in conjunction with high *Prevotella* and *Lactobacillus* abundance (Fig 4B) completed the cluster's multi-omics profile.

18 individuals had metabotype 4, including two individuals with T2D, 7 with hypertension and 8 treated for high cholesterol. The median age was the highest in this cluster compared to all others at 56 (39-64) years. Cluster 4 had the highest total body fat at 56.5 (40.6-78.9) kg. BARIA individuals stratified within this metabotype exhibited the highest creatinine at 75 (56-172) mmol/L, and lowest glomelular filtration rate at 78 (26-90) kl/1.73m2. The transcriptomics datasets from liver tissues exhibited a very strong negative regulation of cortisol synthesis, glutamatergic synapse, cGMP-PKG signaling pathway and GABAergic synapse. When focusing on the gut microbial species, individuals of metabotype 4 had many changes in the microbial composition, and the abundance of some of the species correlated with plasma glucose, low-density lipoprotein (LDL) cholesterol and cholesterol (S10 Fig). In outline, the individuals of metabotype 4 had potentially impaired kidney function, high body fat, downregulation of synaptic pathways in the liver, upregulation of fatty acid metabolic process in mesenteric adipose tissue, upregulation of pathogenic-related pathways in subcutaneous adipose tissue and increased levels of *Clostridium*, *Streptococcus* and *Klebsiella* in the gut microbial metagenome.

17 individuals had metabotype 5, with dominance of females (14 females/three males) and the median age of the individuals in this cluster was the second youngest, 44 (22-62) years. The participants were relatively treatment naïve, only four were treated for T2D, three for hypertension and three participants for dyslipidemia, with very little overlapping treatments. 1-carboxyethylvaline, 1-carboxyethylisoleucine and valine positively correlated with HbA1c ($r=0.4$, $P<0.01$) (S8 Fig, S9 Fig). In conclusion, metabotype 5 corresponded to a relatively young cluster, with no striking comorbidity treatment, upregulated fatty acid metabolic pathways and immune response pathways in the liver and highly abundant in *Citrobacter.*

Multi-omics integration elucidates discriminatory signature and associations between datatypes

To reveal key interactions between multi-omics data sets, we used DIABLO[60] to identify how the five metabotypes are associated with altered expression in different tissues and an altered gut microbiota. Initially, we provided the differentially abundant metabolites, genes from liver, jejunum, mesenteric and subcutaneous adipose tissue, and gut microbial species for each BARIA individual, along with their respective metabotype membership, as input to the algorithm. DIABLO simultaneously calculates the correlations among all input multi-omics datasets and selects a minimal set of input variables that differentiate the metabotypes. The computational framework used here for integrating various omics datasets successfully identified a highly correlated discriminatory signature for SOM-defined obesity phenotypes that includes multiple *Prevotella* species (*P. veroralis, P. copri, P. multisaccharivorax, P. oulorum, P. denticola, P.* sp. oral taxon 299*, P. bryantii, P. melaninogenica*), *Intestinibacter bartlettii*, *Anaerococcus prevotii*, lipid metabolites (especially phospatidylcholines), liver genes enriched in oxidative phosphorylation, lipid metabolism and cardiomyopathy pathways, subcutaneous adipose tissue *IL6* and *SELE* genes involved in inflammatory and immune system pathways and mesenteric adipose tissue genes enriched in prolactin signaling, T2D and PI3K-Akt signaling pathways (S11 Fig).

Metabotypes are associated with weight loss response to bariatric surgery

In order to define the clinical value of metabotyping, we had to assess the metabotypes' response to bariatric surgery. Hence, we performed a longitudinal biometrics post-operative control of the BARIA obese individuals at three time points: three months, six months and 12 months after surgery, where we monitored the weight, waist circumference and upper leg circumference. It is noteworthy that there are no distinct statistically significant responses in the weight loss or waist circumference reduction immediately after bariatric surgery (3 months after surgery), contrary to what would be expected (Fig 5A, B).

There is a trend that metabotypes 2 and 5 have the highest weight loss one year post-operatively (35kg and 38 kg in average, respectively). Metabotype 2 exhibits the largest waist circumference loss at three months after surgery (12cm) even if this is not deemed statistically significant (Fig 5C, D). However, there is a clear pattern in the reduction of adipose tissue in the upper leg circumference. Metabotypes 1 and 5 are the best responders when it comes to upper leg circumference reduction, with the loss being consistent at all three time points. Upper leg circumference loss is significant ($P<0.05$) when compared to the worst responder cluster, metabotype 3 (Fig 5E, F). This trend is the same for weight loss, regardless of being confirmed by statistical calculations.

Surprisingly, when the BARIA individuals were grouped according to having or not-having metabolic syndrome, there were no notable statistically significant differences in weight and adiposity loss in none of the three time points.

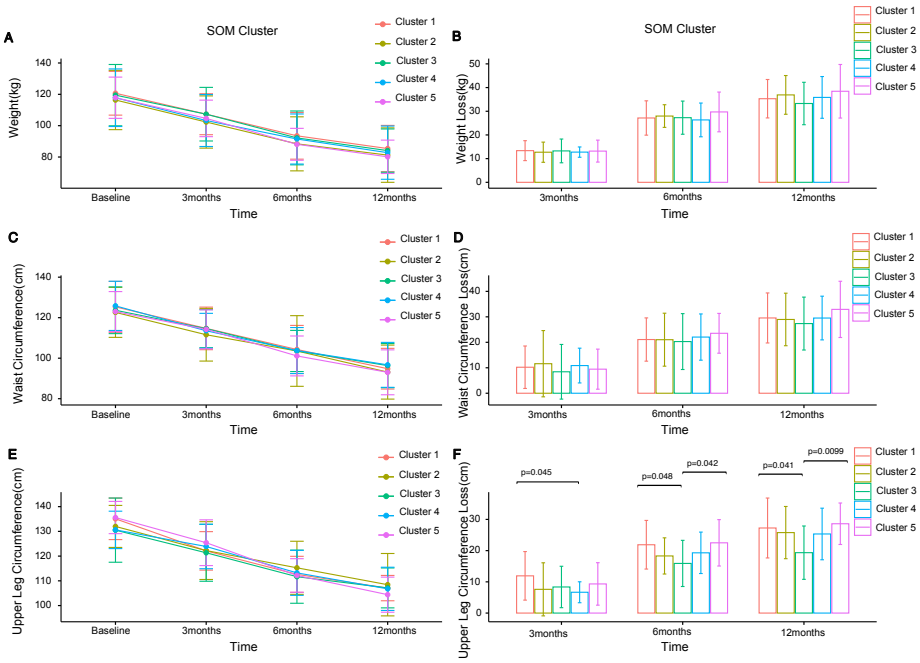


Figure 5. Weight and fat loss progression at distinct time points after bariatric surgery for the five defined SOM clusters (metabotypes). (A) Weight (kg) of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. **(B)** Weight loss(kg) of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. **(C)** Waist circumference (cm) of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. **(D)** Reduction of waist circumference(cm) of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. **(E)** Upper leg circumference (cm) of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. **(F)** Reduction of upper leg circumference(cm)of BARIA individuals at baseline, three months, six months and one year after bariatric surgery for each metabotype. Statistical significance among metabotypes is calculated with t-test and adjusted with FDR; the symbols indicating significance among metabotypes are '*': P<=0.05, '**': P<=0.01, '***': P<=0.001

DISCUSSION

Here we present a novel unsupervised machine learning framework for stratification of individuals in human volunteer cohorts, with a high prevalence and variance of comorbidities. This framework enables a naïve to clinical labels stratification based on fasting metabolome rather than purely clinical parameters that may fail to accurately encompass the multitude of nuances in human population-based studies. The main findings of our study revealed

pronounced changes in lysophospholipids, phosphatidylcholines, dicarboxylate fatty acids, sphingomyelins, and branched-chain amino acid metabolites among the five different metabotypes; KEGG metabolic pathways related to immune functions, fatty acid biosynthesis and elongation, protein signaling and pathogenic pathways were regulated in different ways for each metabotype; the abundance of *Prevotella* and *Lactobacillus* species varied the most between the metabotypes, and metabotypes 4 and 5 had a lower abundance compared to metabotypes 2 and 3. Multi-omics integration enabled reducing the dimensionality and identified a concrete biomarker signature able to differentiate between the five distinct metabotypes. The differences in metabolism among the individuals in the five metabotypes are associated with different responses in terms of weight loss and reduction of waist and upper leg circumference to bariatric surgery.

A considerable advantage of our approach is that SOM and k-means clustering effectively reduced the initial omics dimensionality and resulted in a reusable topological projection of the metabolome. Given the lack of an external multimodal multi-omics dataset for validating our results, establishing a metabolome mapping that can recognize or characterize new unknown inputs can be proven useful. New metabolomes can be projected into the same map, without the requirement of further algorithmic training. That way we can compare metabolic distances among new BARIA inclusions or even the potential post-surgical metabolomes of the initial 106 inclusions. Comparing the post-operative metabolome with the baseline pre-operative state could provide further mechanistic comprehension of the pathophysiological mechanisms of obesity and the responses to the bariatric surgery intervention in the future. Also using the multi-parameter metabolic syndrome as a classifier was here shown not to enable new insight into what drives differences in metabolism within the cohort. Metabotyping identified more gut microbial species among BARIA individuals, whereas the metabolic syndrome classification captured a fraction of the microbial variability. It has been previously attempted in animal studies to model interactions between genes, gut microbiota and the molecular mechanisms underlying obesity[61–63], but their clinical application to humans has been limited[64] so far. Increased microbial variability among metabotypes along with the results from the KEGG pathways enrichment in liver and adipose tissues could be the effect of gut microbial species in the hosts' gene regulation. In the metabotype comparison, the statistically significant anthropometric features of age and glomerular filtration rate along with the differentially significant KEGG pathways could plausibly reflect the process of cellular and biological senescence[65,66]. The detection of senescence in the metabolome by our proposed SOM and k-means methodology, without prior knowledge of biometric characteristics strengthens our claim that the identified metabotypes stand as different representations of human metabolism among the BARIA obese individuals.

A limiting factor that needs to be considered when interpreting our findings is the selection of the eligible individuals for bariatric surgery. The significant variability within human cohorts is often not possible to capture in a finite number of clinical variables. For example, classifiers for obesity-associated comorbidities such as hypertension, T2D, and dyslipidemia may be treated as binary variables (present *vs.* absent)[9], however the overall wellness of an individual with any of these disorders can vary significantly as a function of how well managed each of these conditions are, among many other factors. The BARIA exclusion criteria for surgical interventions have to be strictly met for minimizing the risks and complications of such an evasive procedure. As a result, the BARIA inclusions might not fully represent the obesity spectrum. There is a visible trend in weight loss and leg/waist circumference reduction among the SOM and k-means defined clusters over time. Nonetheless the statistically significant differences among all the identified SOM clusters were not conclusive, probably due to statistical power. Despite the 106 BARIA inclusions and the high quality of the omics dataset, each cluster contains 17-29 individuals, which might account for the values of the statistically significant results.

Conclusions

The principal contribution of this study is the detailed omics dataset for obese individuals, that includes metabolome, microbiome and especially transcriptome from multiple tissues. Our findings suggested that participants' stratification based on metabotyping could enhance our ability to get molecular insights into the causes of diseases from multi-omics integrative analysis. The combination of SOM metabotyping and DIABLO correlation analysis highlights the data-driven nature of this approach. DIABLO analysis enabled the identification of an underlying common yet discriminatory minimal multi-omics signature for the SOM-defined metabotypes, that could lead to predictive markers of the bariatric surgery outcome. In this light, use of biologic parameters such as the plasma metabolome, as a direct readout of the overall status of an entire multiorgan system host and its microbiome, to determine grouping of individuals, offers a unique approach that may more accurately classify individuals into distinct disease physiological states[11,12,67]. Rather than traditional clinical disease classifiers, this grouping method may reduce the confounding effects of such clinical metadata[68,69]. The multi-omics dataset's association framework can be the starting point for selecting candidate compounds for a more thorough examination and provide mechanistic insight into the causality of pathogenicity originating in the tissues, mediated by bacteria and materializing via metabolites and clinical metadata. The multi-omics integrative framework implemented could also be utilized as a hypothesis generating tool for comprehending cardiometabolic disease. Our data suggest that self-organized metabotyping, based only on metabolite distribution, with no other prior knowledge on

the individuals' clinical status in combination with DIABLO integrative analysis, constitute a valuable computational approach studying multifaceted metabolic disorders.

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC. All participants provided written informed consent.

Availability of data and materials

The metagenomics dataset supporting the conclusions of this article is available in the European Nucleotide Archive (ENA) database under the accession number PRJEB47902, (https://www.ebi.ac.uk/ena/browser/view/PRJEB47902?show=reads). The transcriptomics dataset has been deposited in the European Genome-Phenome Archive (EGA) database under the accession number EGAS00001005704. However, access to the transcriptomics data requires additional permission, due to the sensitive nature of the dataset and GDPR. The metabolomics dataset has been deposited in the MetaboLights repository. Access to the metabolomic data can only be granted by the corresponding author upon request, due to the sensitive nature of the dataset and GDPR. All source code can be made available from the corresponding author upon request.

Competing Interests

The authors declare no conflict of interest.

Funding

The BARIA study is funded by the Novo Nordisk Foundation (NNF15OC0016798). The Novo Nordisk Foundation Center for Basic Metabolic Research is supported by an unconditional grant (NNF10CC1016515) from the Novo Nordisk Foundation to University of Copenhagen. The BARIA study is a Scandinavian-Dutch collaboration. Funding from Knut and Alice Wallenberg Foundation is also acknowledged.

Author's contributions

T.W.S., M.N., J.N. and F.B. supervised this work, - D.L. conducted conceptualization, data curation, formal analysis, visualization and original draft preparation, - A.S.M., Ö.A., A.S.V.R., Y.I.Z.A., M.L.D.B, collected medical data and biopsies, - V.T., A.L. performed DNA, RNA and metabolomics isolations and optimizations, - L.M.O. conducted metagenomics data analysis, L.E.O., F.B. and J.N. coordinated project administration, - D.L., A.S.M., K.K.,

L.M.O., Ö.A., A.S.V.R., Y.I.Z.A., M.L.D.B., V.T., L.E.O., A.L., S.H., B.J., V.E.A.G., A.K.G., T.W.S., M.N., F.B. and J.N. conducted hypothesis generation, manuscript review and editing.

Acknowledgements

The computations and RNA Sequencing were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE (SNIC Computational Center of Chalmers University of Technology) partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

REFERENCES

1. McLendon, R. *et al.* Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
2. Manichanh, C., Borruel, N., Casellas, F. & Guarner, F. The gut microbiota in IBD. *Nature Reviews Gastroenterology & Hepatology* **9**, 599–608 (2012).
3. Dao, M. C. & Clément, K. Gut microbiota and obesity: Concepts relevant to clinical care. *European Journal of Internal Medicine* **48**, 18–24 (2018).
4. Vanamala, J. K. P., Knight, R. & Spector, T. D. Can Your Microbiome Tell You What to Eat? *Cell Metabolism* **22**, 960–961 (2015).
5. Wilmanski, T. *et al.* Blood metabolome predicts gut microbiome α-diversity in humans. *Nature Biotechnology* **37**, 1217–1228 (2019).
6. Obesity and overweight. https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight.
7. WHO. *WHO. Obesity: preventing and managing the global epidemic.WHO Technical Report Series number 894*. (2000).
8. Bentham, J. *et al.* Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. *The Lancet* **390**, 2627–2642 (2017).
9. Engin, A. The Definition and Prevalence of Obesity and Metabolic Syndrome. in *Obesity and Lipotoxicity* (eds. Engin, A. B. & Engin, A.) 1–17 (Springer International Publishing, 2017). doi:10.1007/978-3-319-48382-5_1.
10. Haslam, D. W. & James, W. P. T. Obesity. *The Lancet* **366**, 1197–1209 (2005).
11. Ahlqvist, E. *et al.* Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes and Endocrinology* **6**, 361–369 (2018).
12. Wang, T. J. Risk Prediction in Cardiovascular Medicine Assessing the Role of Circulating , Genetic , and Imaging Biomarkers in Cardiovascular Risk Prediction. *Circulation* **123**, 551–565 (2011).
13. Huang, P. L. A comprehensive definition for metabolic syndrome. *Disease Models & Mechanisms* **2**, 231–237 (2009).
14. Bektaş, M., Reiber, B. M. M., Costa, J., George, P. & Donald, L. B. Artificial Intelligence in Bariatric Surgery : Current Status and Future Perspectives. *Obesity Surgery* (2022) doi:10.1007/s11695-022-06146-1.
15. Francisco, S. S., Education, C. & Adult, P. World Health Organization – Defined Metabolic Syndrome Is a Better Predictor of Coronary Calcium Than the Adult Treatment Panel III Criteria in American. *Diabetes Care* **12**, 2977–9 (2004).
16. Fiamoncini, J. *et al.* Plasma metabolome analysis identifies distinct human metabotypes in the postprandial state with different susceptibility to weight loss – mediated metabolic improvements. *The FASEB Journal* **32**, 5447–5458 (2018).
17. Riedl, A. *et al.* Metabotyping and its application in targeted nutrition : an overview. *British Journal of Nutrition* **12**, 1631–1644 (2017).
18. Lacruz, M. E. *et al.* Instability of personal human metabotype is linked to all-cause mortality. *Scientific Reports* **8**, 9810 (2018).
19. Riedl, A. *et al.* Identification of Comprehensive Metabotypes Associated with Cardiometabolic Diseases in the Population-Based KORA Study. *Molecular Nutrition & Food Research* **62**, 180017 (2018).
20. Cirulli, E. T. *et al.* Profound Perturbation of the Metabolome in Obesity Is Associated with Health Risk. *Cell Metabolism* **29**, 488-500.e2 (2019).
21. Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
22. Sommer, F. & Bäckhed, F. The gut microbiota-masters of host development and physiology. *Nature Reviews Microbiology* **11**, 227–238 (2013).
23. Deschasaux, M. *et al.* Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nature Medicine* vol. 24 1526–1531 Preprint at https://doi.org/10.1038/s41591-018-0160-1 (2018).

24. Mardinoglu, A. *et al.* The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Molecular Systems Biology* **11**, 834–834 (2015).

25. Schroeder, B. O. & Bäckhed, F. Signals from the gut microbiota to distant organs in physiology and disease. *Nature Medicine* **22**, 1079–1089 (2016).

26. Bouter, K. E., van Raalte, D. H., Groen, A. K. & Nieuwdorp, M. Role of the Gut Microbiome in the Pathogenesis of Obesity and Obesity-Related Metabolic Dysfunction. *Gastroenterology* **152**, 1671–1678 (2017).

27. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).

28. Savolainen, O. *et al.* Biomarkers of food intake and nutrient status are associated with glucose tolerance status and development of type 2 diabetes in older Swedish women. *American Journal of Clinical Nutrition* **106**, 1302–1310 (2017).

29. Sun, L. *et al.* Gut microbiota and intestinal FXR mediate the clinical benefits of metformin. *Nature Medicine* **24**, 1919–1929 (2018).

30. Schiattarella, G. G. *et al.* Gut microbe-generated metabolite trimethylamine-N-oxide as cardiovascular risk biomarker: A systematic review and dose-response meta-analysis. *European Heart Journal* **38**, 2948–2956 (2017).

31. Nielsen, J. Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine. *Cell Metabolism* **25**, 572–579 (2017).

32. Van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *Journal of Internal Medicine* 0–2 (2020) doi:10.1111/joim.13157.

33. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nature Methods* **13**, 731–740 (2016).

34. Stan Pounds, D. F. HybridMTest: Hybrid Multiple Testing. Preprint at https://doi.org/10.18129/B9.bioc.HybridMTest (2019).

35. Vesanto, J. SOM Toolbox.

36. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43**, 59–69 (1982).

37. Kohonen, T. K. *MATLAB Implementations and Applications of the Self-Organizing Map*. (2014). doi:10.1186/1471-2407-2-7.

38. Kohonen, T. *Self-Organizing Maps*. *Series in Information Sciences* (Springer International Publishing, 1995).

39. Self-Organizing Feature Maps. 119–157 (1989).

40. Liu, Y., Weisberg, R. H. & Mooers, C. N. K. Performance evaluation of the self-organizing map for feature extraction. *Journal of Geophysical Research: Oceans* **111**, 1–14 (2006).

41. Davies D.L., B. D. W. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-1**, 224–22.

42. Philip A Ewels, Alexander Peltzer, Sven Fillinger, Johannes Alneberg, Harshil Patel, Andreas Wilm, Maxime Ulysse Garcia, Paolo Di Tommaso, S. N. nf-core: Community curated bioinformatics pipelines. *bioRxiv* 610741 (2019) doi:doi: 10.1101/610741.

43. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

44. Ignatiadis, N., Klaus, B., Zaugg, J. B. & Huber, W. Data-driven hypothesis weighting increases detection power in genome- scale multiple testing. **13**, (2016).

45. Pantano, L. DEGreport: Report of DEG analysis. Preprint at https://doi.org/10.18129/B9.bioc.DEGreport (2019).

46. Chen, E. Y. *et al.* Enrichr : interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

47. Kanehisa, M. & Goto, S. KEGG : Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27–30 (2000).

48. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies. *Nature Biotechnology* **35**, 1069–1076 (2017).

49. Karlsson, F. H., Nookaew, I. & Nielsen, J. Metagenomic Data Utilization and Analysis (MEDUSA) and Construction of a Global Gut Microbial Gene Catalogue. *PLoS Computational Biology* **10**, (2014).

50. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).

51. Chamberlain, S. A. & Szöcs, E. taxize : taxonomic search and retrieval in R [v2 ; ref status : indexed , http://f1000r.es/24v]. **191**, 1–28 (2014).

52. Mcmurdie, P. J. & Holmes, S. phyloseq : An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217 (2013).

53. Singh, A., Shannon, C. P., Rohart, F., Tebbutt, S. J. & Le, K. Systems biology DIABLO : an integrative approach for identifying key molecular drivers from multi-omics assays. **35**, 3055–3062 (2019).

54. Tenenhaus, A., Guillemot, V., Grill, J. & Frouin, V. Variable selection for generalized canonical correlation analysis. 569–583 (2014) doi:10.1093/biostatistics/kxu001.

55. Singh, A., Cao, K. L. & Rohart, F. mixOmics : An R package for ' omics feature selection and multiple data integration. *PLoS Computational Biology* **13**, 1–19 (2017).

56. Lloyd, S. P. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory* **28**, 129–137 (1982).

57. Wu, H. *et al.* The Gut Microbiota in Prediabetes and Diabetes: A Population-Based Cross-Sectional Study. *Cell Metabolism* **32**, 379-390.e3 (2020).

58. Sze, M. A. & Schloss, P. D. Looking for a signal in the noise: Revisiting obesity and the microbiome. *mBio* **7**, 1–10 (2016).

59. Thingholm, L. B. *et al.* Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell Host and Microbe* **26**, 252-264.e10 (2019).

60. Singh, A., Shannon, C. P., Rohart, F., Tebbutt, S. J. & Le, K. Systems biology DIABLO : an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).

61. Kleinert, M. *et al.* Animal models of obesity and diabetes mellitus. *Nature Reviews Endocrinology* vol. 14 140–162 Preprint at https://doi.org/10.1038/nrendo.2017.161 (2018).

62. Renner, S. *et al.* Porcine models for studying complications and organ crosstalk in diabetes mellitus. *Cell and Tissue Research* vol. 380 341–378 Preprint at https://doi.org/10.1007/s00441-019-03158-9 (2020).

63. Koya, D. & Kanasaki, K. Biology of obesity: Lessons from animal models of obesity. *Journal of Biomedicine and Biotechnology* vol. 2011 Preprint at https://doi.org/10.1155/2011/197636 (2011).

64. Yazdi, F. T., Clee, S. M. & Meyre, D. Obesity genetics in mouse and human: Back and forth, and back again. *PeerJ* **2015**, 1–69 (2015).

65. Childs, B. G., Durik, M., Baker, D. J. & Van Deursen, J. M. Cellular senescence in aging and age-related disease: From mechanisms to therapy. *Nature Medicine* Preprint at https://doi.org/10.1038/nm.4000 (2015).

66. Collado, M., Blasco, M. A. & Serrano, M. Review Cellular Senescence in Cancer and Aging. 223–233 (2007) doi:10.1016/j.cell.2007.07.003.

67. Wang, T. J. *et al.* Metabolite profiles and the risk of developing diabetes. *Nature Medicine* **17**, 448–454 (2011).

68. Holmes, E., Wilson, I. D. & Nicholson, J. K. Forum Metabolic Phenotyping in Health and Disease. *Cell* **134**, 714–717 (2008).

69. Nicholson, J. K. *et al.* Host-gut Metabolic Interactions. *Science* **336**, 1262–1268 (2012).

Supplementary information available at:

Self-organized metabotyping of obese individuals identifies clusters responding differently to bariatric surgery (journals.plos.org/plosone)

6

Systems analysis of metabolic responses to a mixed meal test in an obese cohort reveals links between tissue metabolism and the gut microbiota

Peishun Li, Boyang Ji, Dimitra Lappa, Abraham S Meijnikman, Lisa M. Olsson, Ömrüm Aydin, Sjoerd C. Bruin, Arnold van de Laar, Valentina Tremaroli, Hao Luo, Jun Geng, Kimberly A. Krautkramer, Annika Lundqvist, Hilde Herrema, Albert K. Groen, Victor E.A. Gerdes, Thue W. Schwartz, Fredrik Bäckhed, Max Nieuwdorp, Louise E. Olofsson, Jens Nielsen

Submitted after revisions

ABSTRACT

Individuals with prediabetes and type 2 diabetes mellitus (T2DM) have poor ability to adapt to diet-triggered perturbations. We investigated global metabolic responses to a mixed meal test (MMT) in morbidly obese individuals with different diabetic status by performing plasma metabolomic profiling. Abnormal metabolism of carbohydrates, (branched-chain) amino acids, fatty acids and acylcholines in individuals with (pre)diabetes was observed. Moreover, differences in metabolic responses were associated with altered fecal metagenomics and transcriptomes of liver, jejunum and adipose tissues, which revealed a modified gut microbiome and multi-tissue metabolism in individuals having insulin resistance. Finally, using integrative machine learning models, we built a predictive model based on metabolomics data after 2h MMT, and identified possible new signatures for glycemic control including N–acetylaspartate and phenylalanine-derived metabolites that may be useful for diagnosis, intervention and prevention of T2DM.

INTRODUCTION

Type 2 diabetes mellitus (T2DM), characterized by hyperglycaemia, is one of the fastest increasing diseases worldwide[1]. Before individuals develop T2DM, they almost always have prediabetes (Pre-D). 5-10% of all individuals with Pre-D will annually progress to T2DM, and ~70% will eventually develop T2DM over the course of their lifetime[2,3]. These individuals are characterized by higher than normal blood glucose levels that have not yet reached the threshold for diabetes diagnosis[2,3]. Individuals with Pre-D and T2D not only manifest metabolic disorders at fasting, but also have a reduced ability to adapt to diet-triggered perturbations, e.g., the limited control for postprandial glycemic level[1,4]. Insulin resistance and pancreatic beta-cell dysfunction play important roles in the metabolic imbalance[5,6]. Postprandial blood glucose control is important for diabetes management, and mixed meal test (MMT) have been often used to assess postprandial responses of glucose and insulin[7–9]. Previous studies have revealed postprandial effects on the metabolism of Pre-D and T2DM patients[10,11]. In addition, early studies reported that postprandial glucose responses were predictable based on personal and microbial compositional features using machine learning models[12]. However, few studies have taken a holistic view on how different underlying factors, including metabolisms of the gut microbiota and human host, contribute to abnormally metabolic responses to a MMT in individuals with (pre)diabetes.

Multi-omics profiling and data integrations have been widely applied in Pre-D and T2DM studies[13] Especially, untargeted metabolomics technologies have provided an opportunity to investigate the global metabolic changes in populations with (pre)diabetes[14,15], but the serum metabolome in previous studies has almost been examined at fasting condition. Several studies have shown that disorders of branched-chain amino acids (BCAAs) metabolism contribute to insulin resistance[16]. Here, we applied a combination of metabolomics profiling and MMT to dynamically quantify metabolic processes in response to a MMT, which provides novel insights into metabolic imbalance in obese individuals with Pre-D and T2DM. In addition, transcriptomics studies have previously suggested potential mechanisms involved in the pathogenesis of Pre-D and T2DM[17], but the existing studies have focused on single-tissue transcriptomic profiling, including subcutaneous adipose tissue, visceral adipose tissue and pancreatic islets. Here, we systematically analyzed gene expression profiles of different human tissues (liver, jejunum, mesenteric and subcutaneous adipose tissues), which enables us to identify differences in multi-tissue metabolisms of individuals with different diabetic status. Also, recent studies have shown that the gut microbiota has already significant alterations of metabolic capacity in Pre-D and correlates to T2DM progression[18,19]. Microbial metabolites, including BCAAs and histidine-derived imidazole propionate, have been demonstrated to be associated with insulin resistance[16]. To gain further insights into associations between the gut

microbiota and the metabolic responses of individuals, here shotgun metagenomics was used to determine the microbiota composition and potential microbiome functions.

Based on these obtained multi-omics data, we finally predicted glucose responses to a MMT. Interestingly, the predictive models trained with metabolomics data (especially after 2h MMT) showed best performance. By integrating different data sets, we were able to reveal how metabolic changes in different organs, including the gut microbiome, interplay and based on this identify a number of metabolites and gut microbial signatures that may serve as novel biomarkers of glycemic control.

METHODS

Study population

We studied 106 individuals with morbid obesity in the BARIA cohort scheduled for bariatric surgery[20]. The study was performed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC. All participants provided written informed consent. Firstly, these individuals were classified into T2DM group (n=20) and non-T2DM group (n=86) according to the diagnosis information. The 86 non-T2DM individuals were further classified by their fasting blood glucose and HbA1c levels according to American Diabetes Association (ADA) criteria[5,38]. 27 Individuals were classified into normal glucose tolerance (NGT) group having HbA1c level <39 mmol/mmol and fasting blood glucose level <5.6 mmol/L; 57 individuals were classified into prediabetes (Pre-D) group having HbA1c level 39–47 mmol/mmol or fasting blood glucose level 5.6–6.9 mmol/L; Two individuals were de novo classified into T2DM group having HbA1c level ≥48 mmol/mmol or fasting blood glucose level ≥7 mmol/L. Finally, the 106 individuals in the cohort were classified into three groups, including NGT (n=27), Pre-D (n=57) and T2DM group (n=22), for further analyses.

Measurements of clinical characteristics

Individuals underwent a complete metabolic work-up at the start of their bariatric surgery trajectory. Anthropometric measurements including height, weight and waist circumference were taken. In addition, body fat percentage using bioelectrical impedance and blood pressure were measured. Fasting blood samples were used for the determination of fasting blood glucose, HbA1c, total cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides, ferritin, CRP (C-reactive protein), hemoglobin, leukocyte, creatinine, magnesium, GGT (gamma glutamyl transferase), ALT (alanine aminotransferase), AST (aspartate aminotransferase), vitamin B12, vitamin D and insulin levels.

Mixed meal test

Within three months before surgery, individuals in the cohort underwent a 2-hour mixed meal test (MMT), which was performed to assess insulin resistance and investigate dynamic alterations in circulating metabolites. The MMT consisted of two Nutridrink compact 125ml (Nutricia®), containing 23.3 grams fat, 74.3 grams carbohydrates (of which 38.5 grams sugar) and 24.0 grams protein. The participants received this meal after fasting for a minimum of nine hours. Time point zero refers to the moment at which the participant had fully consumed the meal. Blood samples were drawn via an intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes. Glucose, insulin and triglycerides were measured at these seven time points.

A number of variables related to insulin resistance, pancreatic β cell function and MMT were calculated using previously published methods[69]. The HOMA2 model (the updated HOMA model) was used to estimate insulin resistance (HOMA2-IR index) and pancreatic β cell function (HOMA2-B index) for an individual from fasting plasma glucose and fasting insulin concentrations measured in a MMT[21]. To quantify the postprandial responses of glucose and insulin to a MMT, total area under the curve (tAUC) and incremental AUC (iAUC, subtracting the baseline values) were calculated from their measurements by the trapezoidal method. For calculating the AUC, the k-Nearest Neighbors (KNN) method was performed for imputation of all missing values using R function knnImputation in DMwR with default parameters. Insulin AUC/glucose AUC ratios were calculated to estimate glucose-stimulated insulin secretion during 2h MMT. Besides, the insulinogenic index was calculated by dividing the insulin iAUC during the first 30 minutes by the glucose iAUC during the same period[22].

Metabolomics analysis

106 and 95 EDTA plasma samples were collected from participants at fasting and 2h after MMT, respectively. Samples were shipped to METABOLON (Morisville, NC, USA) for performing analysis using ultra high-performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) untargeted metabolomics, as previously described[23]. The metabolomic abundance obtained, underwent significant curation via metabolites' pre-filtering, imputation for subsets of metabolites' missing values and data normalization, in order to minimize the effect of artifacts in the downstream analysis. The abundances of all metabolites from fasting and post-meal samples were analyzed together in this study. The metabolomics dataset is comprised of 1345 metabolites with 1041 compounds of known identity (named metabolites) and 304 compounds of unknown structural identity (unnamed metabolites). Metabolomics prefiltering and imputation were performed by utilizing a variation of the Perseus platform[24]. Essentially, data has been pre-filtered so as to have a

maximum of 25% missing values for a metabolite across all samples. This was followed by a log transformation of all the measured metabolites' raw intensities across the entire dataset. Then, we calculated the total data mean and standard deviation (by omitting missing values). Taking into account that the metabolite intensities distribution is approximately following normality, we chose a small distribution 2.5 standard deviations away from the original data mean towards the left tail of the original data distribution, with 0.5 standard deviations width. This new shrunken range corresponds to the actual lowest level of detection by the spectrometer. Here by drawing random values from this mini distribution, we filled the missing prefiltered data of choice.

Normalization was conducted as previously described[23]. Out of the 1345 metabolites analyzed by Metabolon, we used 998 metabolites in our downstream analysis after normalization and imputation.

To identify metabolites differential among the three groups with varied diabetic status or between two time points (fasting and 2h after MMT), two main effects (groups and time) and their interaction were assessed by multi-factor ANOVA that was adjusted for covariate age. Student's t-test was used for multiple pairwise comparisons between two groups. P values were corrected for multiple testing using the false discovery rate method. Adjusted $P < 0.05$ was considered as threshold to identify significantly differential metabolites associated with diabetic status or the MMT. Due to the particular interest in (pre)diabetes-related metabolites' responses to a MMT, the metabolites were further classified into three types of response patterns, as shown in Supplementary Fig. 6. Type I metabolites have no significant main effect for time and no interaction of two main effects time and groups, i.e., no response to a MMT. The first plot shows where the time profiles have no change and are parallel for the groups (parallel means no interaction). Type II metabolites have significant main effect for time but no interaction, i.e. parallel response to a MMT. The middle plots show where the time profiles have changes but are still parallel for the groups. Type III metabolites have significant interaction of two main effects, i.e., differential response to a MMT. The last plot shows where the time profiles have different changes for the three groups. The metabolites having differential responses among the three groups were identified by examining the significance of the interaction of two main effects ($P < 0.05$). Moreover, to quantify the MMT-induced metabolic changes of each individual, the ratio of each metabolite abundance at 2h post MMT to fasting abundance was calculated.

Transcriptome analysis

Biopsies from liver (106 samples), jejunum (105 samples), mesenteric adipose fat (104 samples) and subcutaneous adipose fat (105 samples) were collected at the time of the bariatric surgery (Supplementary Table 1), as previously described[20]. RNA was extracted

from biopsies using TriPure Isolation Reagent (Roche) and Lysing Matrix D, 2 mL tubes (MP Biomedicals) in a FastPrep®-24 Instrument (MP Biomedicals) with homogenization for 20 seconds at 4.0 m/sec, with repeated bursts until no tissue was visible; homogenates were kept on ice for 5 minutes between homogenization bursts if multiple cycles were needed. RNA was purified with chloroform (Merck) in phase lock gel tubes (5PRIME) with centrifugations at 4°C, and further purified and concentrated using the RNeasy MinElute kit (QIAGEN, Hilden, Germany). The quality of RNA was analysed on a BioAnalyzer instrument (Agilent), with quantification on Nanodrop (Thermo Fisher Scientific). Due to degradation of the RNA, libraries for RNAseq sequencing were prepared by rRNA depletion; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc.) with 150 bp paired-end reads and 10G data/sample. The average read count per sample from liver and jejunum tissues are 42 ± 15 million. For mesenteric and subcutaneous adipose tissue, the average read count per sample are 43.2 ± 20 million.

Raw RNA-seq reads data were analyzed using nf-core/rnaseq[25], a bioinformatics analysis pipeline for RNA sequencing data. Raw RNA-seq reads data was subjected to quality control using FastQC and multiQC[25]. The alignment of sequencing reads to the reference genome Homo sapiens GRCh38 was performed using STAR[26]. Gene counts were generated using featureCounts and StringTie. The pipeline was built using Nextflow[27].

To identify the differential genes between NGT, Pre-D and T2DM groups, multivariate negative binomial generalized linear models were performed by R package DESeq2. The models were adjusted for covariates age, BMI and gender. Only genes with the sum of counts across all samples ≥10 and existed in at least five samples were considered in the analysis. Raw read counts of genes were normalized using the median of ratios method by DESeq2. P values were corrected for multiple testing using BH method for per pairwise group comparison in each tissue. Adjusted $P<0.05$ was considered as threshold to identify significantly differentially expressed genes. To further explore differences in KEGG functions among the three groups, gene set analysis (GSA) was performed using statistics of all genes (P value and $\log_2$ fold change) and R package PIANO with the reporter algorithm for KEGG pathways. The gene sets with a distinct directional P value<0.05 were chosen in this study, that is only considering gene sets significantly enriched by distinctly up or down-regulated genes.

Microbiome analysis

Fecal samples from 106 participants were collected on the day of surgery and immediately frozen at -80C. Total fecal genomic DNA was extracted from 100 mg feces using a modification of the IHMS DNA extraction protocol[28]. Briefly, fecal samples were extracted in Lysing Matrix E tubes (MP Biomedicals) containing ASL buffer (QIAGEN), and lysis of

cells was obtained, after homogenization by vortexing for 2 minutes, by two cycles of heating at 90°C for 10 minutes followed by three bursts of bead beating at 5.5 m/sec for 60 seconds in a FastPrep®-24 Instrument (MP Biomedicals). After each bead-beating burst, samples were placed on ice for 5 minutes. The supernatants containing fecal DNA were collected after the two cycles by centrifugation at 4°C. Supernatants from the two centrifugations steps were pooled and a 600 µL aliquot from each sample was purified using the QIAamp DNA Mini kit (QIAGEN) in the QIAcube (QIAGEN) instrument using the procedure for human DNA analysis. Samples were eluted in 200 µL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0). Libraries for shotgun metagenomic sequencing were prepared by a PCR-free method; library preparation and sequencing were performed at Novogene (China) on an HiSeq instrument (Illumina Inc.) with 150 bp paired-end reads and 6G data/sample.

MEDUSA is an integrated pipeline for pre-processing of raw shotgun metagenomics sequence data[29], which maps reads to reference databases, combines output from several sequencing runs and manipulates tables of read counts. The input number of total reads from the metagenome analysis were on average 23.4±2.2 million reads per sample and the total aligned reads were 16.6±1.8 million reads per sample (Supplementary Table 6). The sequencing runs had high quality with almost 98% of the reads passing the quality cut-off. Out of the high-quality reads, on average 0.04% aligned to the human genome, although the data had been cleaned for human reads. Out of the high quality non-human reads, 78.4% aligned to the MEDUSA's software gene catalogue that contains more than 11 million genes. Quality filtered reads were mapped to a genome catalogue that contains 1747 species genomes and the gene catalogue using Bowtie2[30]. Raw read counts at different taxonomy levels were normalized by scaling with cumulative sum (i.e. relative abundance). The α-diversity was calculated based on species-levels of each sample using Shannon, Simpson and Invsimpson indices via R package vegan. To visualize and evaluate differences in gut microbiota composition among groups with varied diabetic status, principal coordinates analysis (PCoA) was performed based on species-level Bray–Curtis distances, and PERMANOVA was performed using the R function adonis in vegan. Microbial taxa at each taxonomical level, including class, order, family and genus, were compared by a Kruskal–Wallis test. P values were adjusted by FDR for each taxonomical level separately. To identify the differential species and KOs among the NGT, Pre-D and T2DM groups, multivariate negative binomial generalized linear models were performed by DESeq2 using raw read counts. The models were adjusted for covariates age, BMI and gender. P values were corrected for multiple testing using BH method for per pairwise group comparison. To further explore differences in KEGG functions among the three groups, gene set analysis (GSA) was performed using Piano with the reporter algorithm for KEGG pathways and modules. The differentially enriched KEGG pathways and modules were identified with a distinct directional P value<0.05.

Statistical analysis

All statistical analyses were performed in the R software version 3.5. To identify the differential responses to a MMT, time curves of glucose, insulin and triglycerides concentrations for Pre-D and T2DM group were compared to NGT group using two-way ANOVA with repeated measures. The significant interaction of two main effects time (MMT) and groups (varied diabetic status) was investigated. In addition, Kruskal–Wallis test was used for comparisons among the three groups and Wilcoxon rank-sum test was used for multiple pairwise comparisons between each two groups. To assess the associations between multi-omics and clinical variables, Spearman's rank correlation analysis was performed. P values were adjusted by FDR to control for multiple comparisons error in this study. To assess correlations between distance matrices of multi-omics, Mantel test was performed using the R package ade4 with the permutation number of 9999. The Bray–Curtis dissimilarity matrices were calculated by using the R function vegdist in vegan.

To predict glucose response to a MMT (i.e. tAUC), ridge regression models were trained based on multi-omics data using R package glmnet. The models were adjusted for covariates age, BMI and gender. First, ridge regression model was used to regress the normalized profile of gut microbiota, metabolomics, transcriptomics against glucose tAUC, respectively. Then, ridge regression model was used to regress the combinational normalized profile of multi-omics against glucose tAUC. The optimal lambda was chosen using function cv.glmnet (10-fold cross-validation in package glmnet) based on the minimum Root Mean Square Error (RMSE). To evaluate performance of the ridge model with the optimal lambda, 5-fold cross-validation (106 samples were randomly divided into five equal parts). Four fifths of samples were used to train the predicted model, and the remaining samples were used to test the fitness of it at each time) was performed by considering the measure RMSE.

RESULTS

Study population

In the present study, we recruited 106 individuals with obesity (BMI≥30 kg/m^2) scheduled for bariatric surgery and included in the BARIA cohort with either normal glucose tolerance (NGT, n=27), Pre-D (n = 57) or T2DM (n = 22) based on the American Diabetes Association criteria[1] (Fig. 1a). Baseline characteristics are summarized in Table 1. Individuals with Pre-D and T2DM (47.3 ± 9.5 and 47.1 ± 10.1 years) were older ($P<0.05$) than those with NGT (41.3 ±10.0 years). Fasting glucose, HbA1c and triglycerides levels were, as expected, significantly higher ($P<0.05$) in the T2DM and Pre-D groups compared with the NGT group (Table 1). Similarly, individuals in the T2DM group had elevated insulin resistance

index (HOMA2-IR) and decreased magnesium levels compared with the NGT and Pre-D groups (P<0.05). Additionally, individuals with T2DM used more medication, including insulin, metformin, thiazide, and statin compared with NGT and Pre-D (P<0.05 by Fisher´s Exact test; Supplementary Fig. 1). Furthermore, blood samples for the two-hour MMT and metabolomics profiling were drawn within three months before the bariatric surgery (Fig. 1a and Supplementary Table 1). Also, biopsies from liver, jejunum and adipose tissues, and fecal samples were collected on the day of the surgery.

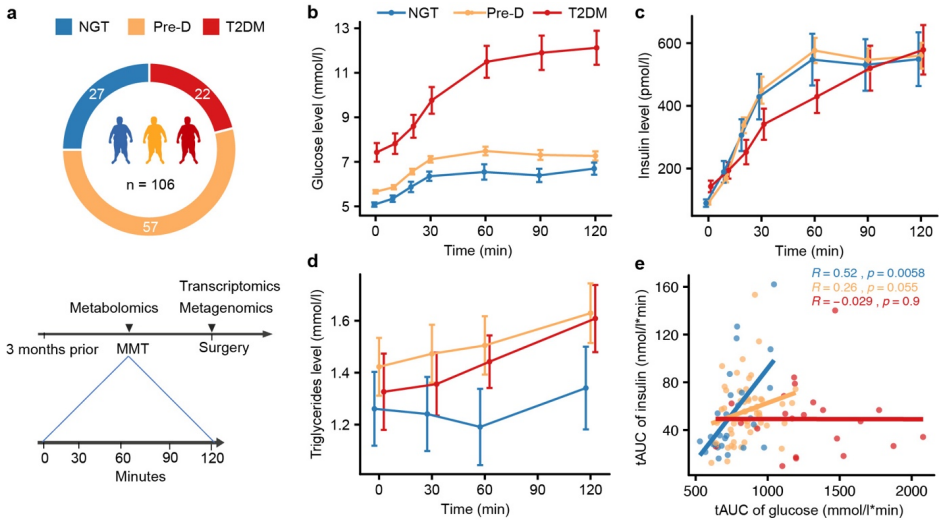


Figure 1: Experimental design and results from a mixed meal test (MMT). a, Schematic illustration of the experimental design. **b**, The time profiles of blood glucose, **c** insulin and **d** triglyceride concentrations during a MMT (Mean ± SEM) in the NGT (n=27), Pre-D (n=57) and T2DM (n=22). **e**, The association between insulin and glucose total AUC in each group. Spearman's rank correlation analysis was performed.

Characteristics	NGT (N = 27)	Pre-D (N = 57)	T2DM (N = 22)
Demographic			
Age (years)	41.3±10.0	47.3±9.5*	47.1±10.1*
Female - no. (%)	24 (88.9)	44 (77.2)	16 (72.7)
Anthropometric			
BMI (kg/m^2)	39.0 (37.5 to 40.9)	39.8 (37.2 to 41.2)	38.7 (35.9 to 42.7)
Weight (kg)	113.0 (108.8 to 121.0)	119.0 (104.8 to 128.0)	118.5 (106.9 to 126.6)
Height (cm)	170.0 (167.0 to 174.5)	172.0 (167.0 to 178.0)	172.5 (165.3 to 179.5)
Waist circumference (cm)	118.0 (111.0 to 124.0)	126.0 (117.8 to 135.0)*	126.0 (117.0 to 130.0)
Total body fat (%)	47.9 (46.8 to 50.1)	47.6 (46.3 to 49.9)	42.8 (40.1 to 48.5)*
Fat free mass (%)	52.3 (50.0 to 52.3)	52.5 (50.1 to 56.7)	57.2 (51.5 to 59.9)*
Total body water (%)	37.9 (35.6 to 39.9)	39.4 (36.5 to 41.6)	42.3 (38.5 to 46.0)*#
Systolic pressure (mmHg)	132.0 (118.0 to 135.0)	136.0 (126.0 to 143.0)	130.0 (122.3 to 138.0)
Diastolic pressure (mmHg)	80.0 (74.0 to 84.0)	82.0 (79.0 to 88.0)	83.0 (77.3 to 86.8)
Clinical lab values			
Fasting glucose (mmol/l)	5.1 (5.0 to 5.4)	5.8 (5.6 to 6)*	7.7 (6.8 to 8.9)*#
HbA1c (mmol/mol)	34.0 (34.0 to 37.0)	39.0 (37.0 to 40.0)*	52.0 (46.0 to 60.0)*#
Fasting insulin (pmol/l)	71 (45 to 116)	80.5 (58.9 to 103.3)	119.0 (64.3 to 212.3)*#
Fasting C-peptide (pmol/l)	0.8±0.3	0.9±0.2	1.0±0.4
Total cholesterol (mmol/l)	4.6±1.0	5.2±1.1*	4.3±1.0#
HDL-cholesterol (mmol/l)	1.1 (1.0 to 1.4)	1.2 (1.0 to 1.4)	1.0 (0.9 to 1.2)
LDL-cholesterol (mmol/l)	3.0±0.9	3.4±1.0	2.7±0.8#
Triglycerides (mmol/l)	1.0 (0.8 to 1.4)	1.4 (1.1 to 1.9)*	1.6 (1.3 to 1.8)*
Ferritin (µg/l)	70.0 (55.0 to 118.0)	128.0 (56.0 to 184.0)	85.5 (27.3 to 162.8)
CRP (mg/l)	5.4 (3.4 to 10.5)	4.9 (2.9 to 7.2)	3.2 (5 to 9.4)
Hemoglobin (mmol/l)	8.7 (8.1 to 9.0)	9.0 (8.6 to 9.4)	8.7 (8.4 to 9.2)
Leukocyte, 10^9/l	6.9 (6.1 to 8.8)	6.9 (5.6 to 8.8)	6.6 (5.6 to 9.4)
Creatinine (µmol/l)	69.0 (61.0 to 78.0)	66.0 (61.0 to 76.0)	65.0 (56.8 to 76.3)
Magnesium (mmol/l)	0.84±0.05	0.82±0.05	0.76±0.08*#
GGT (IU/l)	22.0 (14.0 to 26.5)	25.0 (18.0 to 35.3)	38.0 (23.5 to 52.0)*#
ALT (IU/l)	22.0 (19.5 to 30.5)	27.0 (22.0 to 41.0)	39.5 (27.0 to 56.0)*#
AST (IU/l)	22.0 (20.0 to 25.5)	24.0 (20.5 to 28.0)	28.0 (21.3 to 40.8)
Vitamin B12 (pmol/l)	321.0 (237.5 to 396.0)	282.0 (219.0 to 360.0)	263.0 (220.8 to 348.5)
Vitamin D (nmol/l)	57.0 (31.5 to 71.5)	48.0 (38.0 to 64.0)	50.0 (30.0 to 71.0)
HOMA2-IR	1.4 (0.8 to 2.2)	1.6 (1.1 to 2.1)	2.4 (1.5 to 4.5)*#
HOMA2-B	129.8 (86.7 to 158.5)	97.5 (84.0 to 120.7)	81.0 (60.1 to 126.2)*

Table 1. Baseline characteristics of the obese cohort. Mean±SD. For categorical variables number and percentages are presented. Non-normally distributed variables are presented as median with interquartile range. For comparison between groups, Fisher's Exact test was used for dichotomous variables and Student's t-test or Wilcoxon rank sum test were used as appropriate for continuous variables. For comparison among three groups, Kruskal-Wallis test was used. '*' denotes significant difference in comparison to NGT group (P < 0.05); '#' denotes significant difference between Pre-D and T2DM groups (P < 0.05). BMI: body mass index, CRP: C-reactive protein, GGT: gamma glutamyl transferase, ALT: alanine aminotransferase, AST: aspartate aminotransferase, HbA1c: Hemoblobin A1c, HDL: high-density lipoprotein, LDL: low-density lipoprotein.

Mixed meal test characteristics

To assess the postprandial glucose response and insulin resistance, all individuals underwent a standardized MMT. The time profiles of blood glucose, insulin and triglyceride concentrations in responses to the MMT are shown in Fig. 1b-d and Supplementary Fig. 2. The MMT triggered a temporary increase in plasma glucose and insulin concentrations in NGT, Pre-D and T2DM groups (P<0.01 by ANOVA). However, glucose excursions differed significantly between the three groups (P<0.01 by ANOVA). This translated into significant differences in total area under the curve (tAUC) as well as in incremental AUC (iAUC, subtracting the baseline values) between the three groups (P<0.01 by Kruskal–Wallis test; Supplementary Fig. 3). T2DM individuals had slightly higher insulin levels at fasting, while had decreased insulin AUC/glucose AUC ratios and insulinogenesis index compared with NGT and Pre-D individuals (P<0.05 by Wilcoxon rank-sum test; Supplementary Fig. 3). The plasma triglyceride levels had an increasing trend over time in the Pre-D and T2DM groups, but with large within-group variations in the three groups (Fig. 1d). Overall, the responses of plasma glucose, insulin and triglyceride to the MMT were characterized by a prolonged elevation in the T2DM group. Interestingly, strong positive correlations between glucose AUC and insulin AUC were only observed in the NGT group (Fig. 1e and Supplementary Fig. 4; for tAUC, R=0.52 and P=0.0058; for iAUC, R=0.58 and P=0.0015). Thus, our results confirmed an abnormal plasma glucose and insulin response during a MMT for the Pre-D and T2DM groups.

Metabolic signatures associated with diabetic status and modulated differentially by a MMT

To investigate the global metabolic responses to the MMT in individuals with different diabetic status, we used peripheral plasma samples for untargeted metabolomic profiling from participants at fasting and 2h post MMT and calculated the Euclidean distances between groups with different diabetic status at both time points (Fig. 2a). Interestingly, the three groups clustered separately at the two time points, which indicates that MMT has a large impact on the metabolomic profiles, with 439 metabolites being differentially abundant between the fasting and postprandial conditions (Adjusted P<0.05 by ANOVA; Supplementary Table 2). The main metabolic processes that had physiological responses to the MMT included lipids (n =183), amino acids (n=72) and xenobiotics (n=36) classes (Supplementary Fig. 5). Moreover, the distance between the NGT and T2DM groups at each time point shows an increased trend, compared to the distance between the NGT and Pre-D groups, or between the Pre-D and T2DM groups (Fig. 2a). This suggests a gradually increased difference in metabolomic profiles with progression from NGT to T2DM status. A total of 145 differential metabolites were associated with diabetic status (Adjusted P<0.05 by ANOVA, Supplementary Table 2), mainly composed of metabolites within the classes of lipids (n =

83), carbohydrates (n = 4), amino acids (n = 34), xenobiotics (n=6) and nucleotides (n = 7) (Fig. 2b). We further identified differentially abundant metabolites between any two groups by multiple pairwise comparisons at two time points, respectively (the bar chart in Fig. 2b; Adjusted P < 0.05 by t-test; Supplementary Table 3). No metabolites showed significantly different levels between the NGT and Pre-D groups at either time point. After 2h MMT, 139 metabolites differed between the NGT and T2DM groups, whereas only 55 metabolites differed at fasting. Consequently, metabolomic profiles were associated with diabetic status and showed a larger difference between the NGT and T2DM groups postprandially.

Next, we classified the 145 metabolites associated with diabetes status into three response patterns, type I-III (Fig. 2c; Method; Supplementary Fig. 6; Supplementary Table 3). Of these, 39 metabolites showed a response pattern with no significant difference between the two time points in each group (type I) and 55 metabolites showed a parallel response to the MMT, independent of diabetes status (type II). The remaining 51 metabolites showed differential responses to the MMT among the three groups (type III) and primarily belonged to carbohydrates, amino acids and lipids. As expected, glucose abundance was significantly increased, whereas 1,5-anhydroglucitol abundance was decreased in the T2DM group compared to the NGT and Pre-D groups at both time points (adjusted P < 0.05; Fig. 2c and Supplementary Table 3). Moreover, mannose showed an elevated abundance in the T2DM group (adjusted P < 0.05), which is in agreement with previous results[31,32]. The identified metabolites in the carbohydrate class followed the type III response pattern, which indicates differential responses of carbohydrate metabolism to a MMT among the three groups.

Amino acid-derived metabolites including BCAAs and aromatic amino acids (AAAs) have been reported to be associated with T2DM[16,33,34]. Here, we identified 34 amino acid-derived metabolites associated with diabetes status (adjusted P<0.05 by ANOVA). Most amino acids in the type I response pattern, including arginine, taurine, N–acetyltaurine, C–glycosyltryptophan and hydroxyasparagine, showed reduced abundances in the T2DM group compared with the NGT and Pre-D groups at both time points (adjusted P<0.05; Fig. 2c and Supplementary Table 3). Additionally, amino acids in the type II response pattern, including 1–carboxyethylphenylalanine, 1–carboxyethylisoleucine, 1–carboxyethylleucine, 1–carboxyethylvaline and 2–hydroxybutyrate, were increased in the T2DM group compared to the NGT and Pre-D groups at both time points (adjusted P<0.05). Furthermore, amino acids with the type III pattern, including creatinine, lysine and N–acetylaspartate (NAA), were decreased in the T2DM group compared with the NGT and Pre-D groups at fasting or 2h post MMT (adjusted P<0.05). Particularly, these metabolites responded differentially to the MMT among the three groups, which may reflect the abnormal amino acid metabolism in the T2DM group after diet.

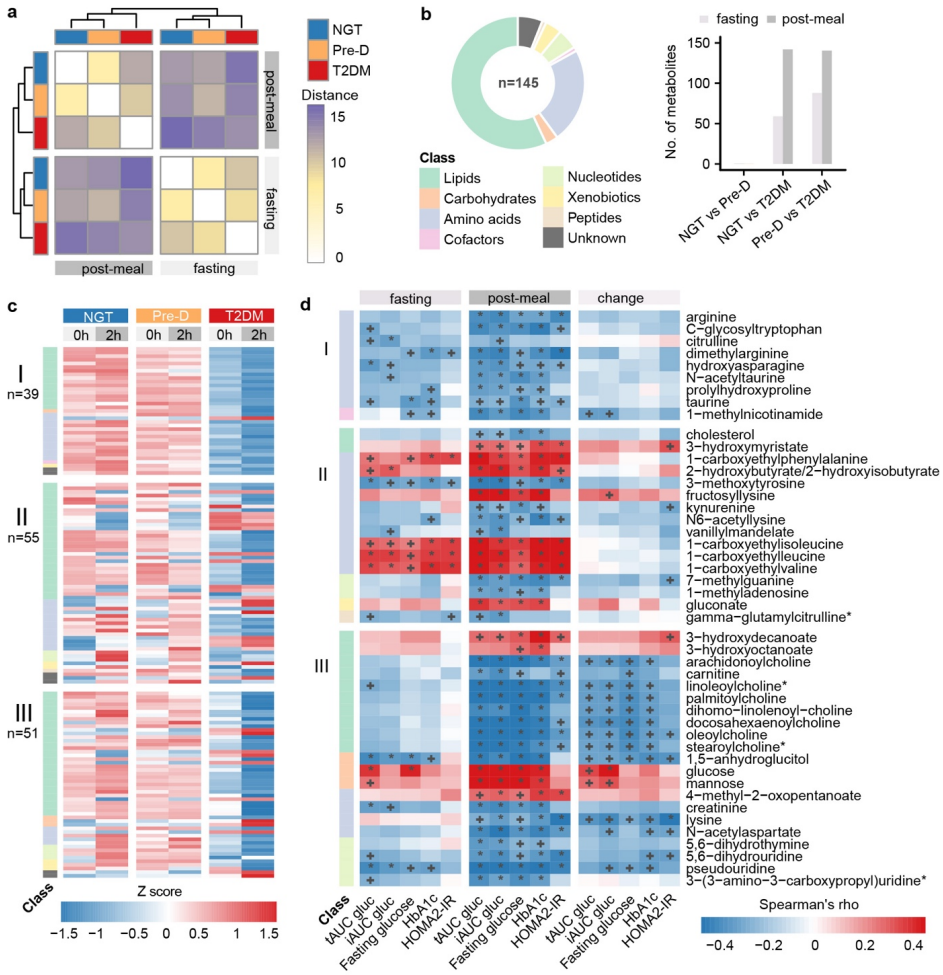


Figure 2: The metabolic changes associated with diabetic status and modulated differentially by the MMT. a, The hierarchical clustering of Euclidean distances between groups with different diabetic status at fasting and 2h post MMT. **b**, The donut chart shows pathways distribution of 145 metabolites differed significantly among the NGT (n=24), Pre-D (n=50) and T2DM (n=21) groups identified by ANOVA (adjusted P < 0.05). The bar chart shows the metabolites differed between the three groups at fasting and 2h post-meal, respectively. **c**, Heatmap showing the mean abundance of the metabolites with three different response patterns in the NGT (n=24), Pre-D (n=50) and T2DM (n=21) groups. **d**, The associations between the metabolomic changes and T2DM-related clinic variables. Only metabolites involved in the metabolic processes, including carbohydrates, amino acids, cofactors, nucleotides, xenobiotics, peptides, acylcholines, fatty acids, carnitine and sterol metabolism are shown. Spearman's rank correlation analysis was performed. '+' denotes adjusted P < 0.05; '*' denotes adjusted P < 0.01. "biochemical name*" indicates a compound that has not been confirmed based on a standard but is confident in its identity.

Many T2DM patients also have dyslipidemia, and our analysis showed that abundances of metabolites in the fatty acid subclass were increased in the T2DM group at fasting or 2h post MMT (adjusted P<0.05, Supplementary Table 3). Six fatty acids, including 3–hydroxydecanoate, 3–hydroxyoctanoate, linoleate (18:2n6), linolenate (18:3n3 or 3n6), had

differential responses to a MMT among the three groups (i.e., type III pattern). However, most metabolites in the lipid class had a decreased trend in the T2DM group compared to the NGT and Pre-D groups at both time points (Fig. 2c and Supplementary Table 3), including sphingomyelin, carnitine, sterol, hexosylceramides, lactosylceramides and acylcholine subclasses. Interestingly, all seven acylcholines, including arachidonoylcholine, linoleoyl-choline*, palmitoylcholine, responded differentially to a MMT among the three groups.

Associations of metabolomic changes with clinical variables

To assess the links between phenotypic characteristics and the postprandial changes of the T2DM-related circulating metabolites, we performed correlation analyses between the clinical variables and these metabolites at fasting or 2h post MMT (Fig. 2d). At fasting, most metabolites with type I and II response patterns had significant correlations with clinical variables glucose tAUC, HbA1c and HOMA2-IR (adjusted P<0.05, Fig. 2d). The carboxyethyl derivatives of BCAAs and phenylalanine were positively correlated with glucose tAUC and HOMA2-IR at both time points (adjusted P<0.05, R=0.29~0.55). Interestingly, correlations between most metabolites and these clinical variables showed an increased trend after 2h MMT. For example, NAA only showed significant correlations with glucose AUC (both iAUC and tAUC) and HOMA2-IR after 2h MMT (adjusted P<0.05, for tAUC, R= -0.42; for iAUC, R= -0.51; for HOMA2-IR, R= -0.36). Furthermore, correlations between the clinical variables and the MMT-induced metabolite changes (i.e., ratio of metabolite abundance at 2h post MMT to fasting) were investigated (Fig. 2d). The postprandial changes of most metabolites with type III pattern had significant correlations with the clinical variables (adjusted P<0.05). Especially, the postprandial changes of metabolites in acylcholine subclass were negatively correlated with glucose AUC (adjusted P<0.05, R= -0.4~-0.33). In addition, the postprandial changes of NAA and lysine correlated negatively with HOMA2-IR (adjusted P<0.05, for NAA, R= -0.32; for lysine, R= -0.49), which suggests that the postprandial regulation of these metabolites might be associated with insulin resistance.

Transcriptional changes associated with diabetes status

To identify differences in metabolic functions of individuals with variable diabetic status, gene expression profiles in four human tissues, including liver, jejunum, mesenteric and subcutaneous adipose tissue, were quantified using RNA-sequencing. As shown in Fig. 3a, the first and second principal component analysis (PCA) components clearly separated samples from different tissues, which accounted for 34% and 24% of the variability, respectively. Differential gene expression analysis by multiple pairwise comparisons between the NGT, Pre-D and T2DM groups resulted in identification of 194, 30, 235 and 11 significantly differentially expressed genes in liver, jejunum, mesenteric and subcutaneous adipose tissue,

respectively (adjusted P<0.05; Supplementary Table 4). These differentially expressed genes show tissue-specific (Fig. 3b). Furthermore, gene set analysis (GSA) identified enrichments of KEGG pathways in the four different tissues (P<0.05; Fig. 3c). Our results showed differences in metabolic pathway related to valine, leucine and isoleucine degradation in liver, jejunum and subcutaneous adipose tissue among the three groups. Differential genes involved in these pathways are summarized in Supplementary Table 5 (P<0.01). In liver, insulin secretion, cAMP signaling pathway and cGMP–PKG signaling pathway were enriched with down-regulated genes, while steroid biosynthesis, terpenoid backbone biosynthesis and propanoate metabolism were enriched with up-regulated genes in the T2DM group compared with the NGT and Pre-D groups (P<0.05; Fig. 3c). The ryanodine receptor gene RYR2 related to insulin secretion and genes related to cGMP-PKG signaling pathway including IRS2, MYLK3, ADRA2C, NPPA, were down-regulated in the T2DM group (P<0.01; Supplementary Table 5). In mesenteric adipose tissue, MAPK, TNF and NF-kappa B signaling pathway and cellular senescence were enriched with up-regulated genes in the T2DM group (P<0.05; Fig. 3c). In subcutaneous adipose tissue, the gene NEU4 (encoding neuraminidase 4) related to sphingolipid metabolism was up-regulated in the Pre-D group compared with the NGT group (P<0.01 and $|\log_2$ (fold change)$|>1$). The gene FASN (encoding fatty acid synthase) involved in fatty acid metabolism was down-regulated in the T2DM group compared with the NGT group (P<0.01 and $|\log_2$ (fold change)$|>0.6$). In jejunum, genes DGAT2, APOA4, MTTP, AGPAT2 related to fat digestion and absorption were up-regulated in the T2DM group compared with the Pre-D group (P<0.01; Supplementary Table 5).

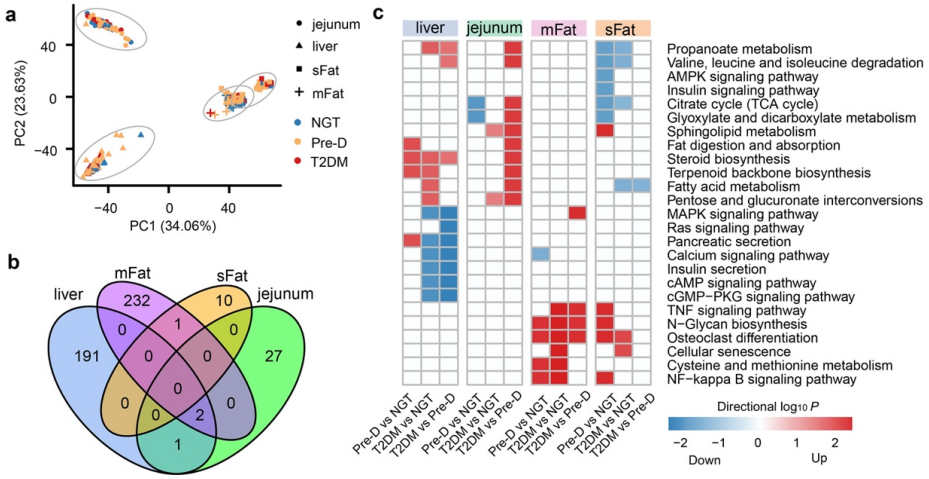


Figure 3: Transcriptional profiles of different human tissues from the NGT, Pre-D and T2DM individuals. a, Principle component analysis (PCA) of transcriptomic profiles in liver (n=106), jejunum (n=105), mesenteric(n=104) and subcutaneous adipose tissues (n=105). Nodes with circle, triangle, rectangle and crisscross represent samples from jejunum, liver, subcutaneous and mesenteric adipose tissue, respectively. **b**, The Venn diagram depicting the

distribution of significantly differentially expressed genes among the three groups in the four tissues (adjusted P < 0.05). **c**, The enriched KEGG pathways comparing the NGT, Pre-D and T2DM groups in the four tissues (P < 0.05). The red color indicates up-regulated gene sets; the blue color indicates down-regulated gene sets.

Microbiota alterations associated with diabetes status

To determine the role of gut microbiota in the metabolic response to a MMT, the gut metagenome of the 106 individuals was quantified using shotgun DNA sequencing (Supplementary Table 6). Principal coordinate analysis (PCoA) shows that the second principal coordinate separates NGT and T2DM groups, which accounts for 11% of the variability (Fig. 4a). PERMANOVA analysis also shows that the diabetic status is associated with dissimilarities in gut microbiota composition (R^2= 0.027, P <0.05). To identify differences in the bacterial composition among the three groups, microbial taxa at each taxonomical level, including class, order, family and genus, were compared by the Kruskal–Wallis test (Supplementary Table 7). Class Epsilonproteobacteria and from within this class, order Campylobacterales, family Campylobacteraceae and genus Campylobacter were more abundant in the Pre-D group than in the NGT and T2DM groups (adjusted P < 0.1 by Kruskal–Wallis test; Fig. 4b and Supplementary Fig. 7). In addition, family Peptostreptococcaceae with the genus unclassified Peptostreptococcaceae showed significantly decreased abundance in the T2DM group compared with the NGT and Pre-D groups (adjusted P < 0.1 by Kruskal–Wallis test; Fig. 4b and Supplementary Fig. 7), which is in accordance with a recent study[35].

Furthermore, 56 differential species were identified by multiple pairwise comparisons between NGT, Pre-D and T2DM groups (adjusted P <0.01; Supplementary Table 8), mainly belonging to the phylum Firmicutes. A total of 24 species exhibited differential abundance in two or three pairwise comparisons between the three groups (Fig. 4c). The abundances of nine species of genus Streptococcus, Lactobacillus sanfranciscensis and Lactobacillus ruminis were increased, whereas the abundances of seven species of genus Clostridium (Clostridium butyricum, Clostridium novyi, etc.), Turicibacter sanguinis, Anaerococcus lactolyticus and Paenibacillus polymyxa were decreased in the T2DM group (adjusted P <0.01). Moreover, Butyrivibrio crossotus and Anaerococcus vaginalis were enriched in the Pre-D and T2DM groups compared to the NGT group (adjusted P <0.01). Especially, C. novyi had a significantly negative correlation with glucose iAUC (R=-0.49, P <1.0e-06; Supplementary Fig. 8a). Additionally, we investigated correlations between the differential species and the metabolites associated with diabetic status at fasting or 2h post MMT (Supplementary Table 9). C. novyi was positively correlated with NAA after 2h MMT (R= 0.46; P <1.0e-05; Supplementary Fig. 8b). Interestingly, the correspondence between the gut microbiota composition and postprandial metabolomic profiles had an increased trend compared to the fasting condition (Supplementary Fig. 9). The carboxyethyl derivatives of

BCAAs and phenylalanine were negatively correlated with several Clostridium species at both time points (P<0.01; Supplementary Fig. 9).

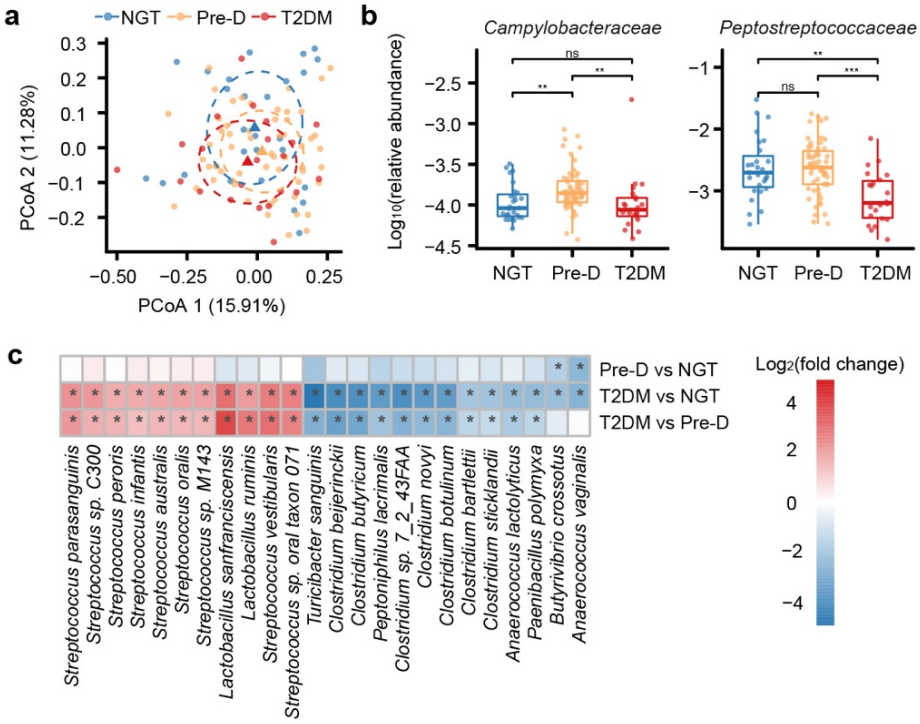


Figure 4: Alterations in gut microbiota associated with diabetes status. a, PCoA of microbiota community at species level based on Bray-Curtis distance (n=106). The centroid for each group is represented as a triangle and the ellipse covers the samples belonging to the group with 95% confidence. **b**, Log_{10} relative abundances of families Campylobacteraceae and Peptostreptococcaceae in the NGT (n=27), Pre-D (n=57) and T2DM (n=22), compared by Wilcoxon rank-sum test for multiple pairwise comparisons. 'ns' denotes no significance; '*' denotes adjusted P < 0.05; '**' denotes adjusted P < 0.01; '***' denotes adjusted P < 0.001; '****' denotes adjusted P < 0.0001. **c**, Heatmap showing log_2 fold changes of 24 significantly differentially species between the NGT (n=27), Pre-D (n=57) and T2DM (n=22). Only species exhibiting differential abundance in two or three pairwise comparisons are shown. '+' denotes adjusted P < 0.05; '*' denotes adjusted P < 0.01.

By investigating the functional capacity of the gut microbiome, we identified 60 significantly differential KEGG orthologs (KOs) among the NGT, Pre-D and T2DM groups (adjusted P <0.05; Supplementary Table 10). Moreover, we observed alterations in potential of phenylalanine and phenylacetate metabolism in the microbiome of individuals with Pre-D and T2DM by gene set analysis (P<0.05; Table 2). The microbial genes including hcaC, hcaF, tynA, feaB, paaA and paaE involved in phenylalanine metabolism, were more abundant in the T2DM group compared to the NGT or Pre-D group (P<0.01 and $|\log_2$ (fold change)|>3; Table 2). Especially, 1-carboxyethylphenylalanine correlated positively with genes feaB, hcaC and paaE (P<0.05 and R>0.2; Supplementary Fig. 10).

	Differential genes (P < 0.01)		
	Pre-D vs NGT	**T2DM vs NGT**	**T2DM vs Pre-D**
KEGG pathway			
Phenylalanine metabolism ↑	–	tynA, feaB, paaA, paaC, paaD, paaE, paaJ	hcaC, hcaF, paaJ
KEGG module			
Phenylacetate degradation ↑	paaE	paaA, paaC, paaD, paaE, paaJ	–

Table 2. The enriched KEGG pathways and modules in gut microbiome between the NGT (n=27), Pre-D (n=57) and T2DM (n=22) groups identified by gene set analysis. '↑' denotes significantly enriched pathway or module comparing two groups (P < 0.05). '–' denotes no differential genes in the pathway or module.

Prediction of postprandial glucose response based on omics data

To systematically investigate potential contributing factors for metabolic responses to a MMT, we first quantified the associations between multi-omics data by the Mantel test using the Bray–Curtis distance (Fig. 5a; Supplementary Table 11). Significant correlations between the gut microbiome and metabolomics at fasting and 2h post MMT were observed (Mantel r = 0.10~0.16, P < 0.05), which suggests that the gut microbiota is linked to metabolism of individuals in this study. Moreover, transcriptomics data correlated significantly in the different human tissues (Mantel r = 0.21~0.28, P < 0.01). Interestingly, metabolomic changes (i.e., ratio of metabolite abundance at 2h post MMT to fasting) were correlated with the metabolomic profile at fasting (Mantel r = 0.084, P < 0.05), as well as transcriptional profiles in jejunum and mesenteric adipose tissue (Mantel r = 0.094 and 0.085, P < 0.05). This demonstrates that metabolic responses were associated with the metabolic profiles at baseline.

To further investigate possible driving factors for postprandial glucose regulation, we predicted glucose tAUC based on multi-omics data using ridge regression models with 5-fold cross-validation. The models trained with metabolomics data (especially after 2h MMT) performed best with minimum root mean square error (RMSE) (Fig. 5b). The performance was improved when the model was trained using taxonomic (species) profiles compared to using functional (KOs) profiles and transcriptomics data. The correlation coefficients between the predictive and actual glucose tAUC were 0.92, 0.91 and 0.9 when using metabolomic profiles at fasting and 2h post MMT, and metabolomic changes as the training sets, respectively (Fig. 5c and Supplementary Fig. 11). Using species and KOs profiles of the gut microbiota as training sets, correlation coefficients between the predictive and actual glucose tAUC were 0.69 and 0.72, respectively (Fig. 5c). At fasting, glutamine, creatinine, pseudouridine, arginine, alanine, mannose, phenylalanine and lysine were identified to be the most important metabolites for prediction of glucose tAUC (Supplementary Fig. 12a). After 2h MMT, mannose, allantoin, phenylalanine, 1–carboxyethylphenylalanine and NAA were predicted to be the most important metabolites (Fig. 5d). Moreover, the postprandial

changes of metabolites N-acetylalanine, carnitine, lysine, histidine, N-acetylserine and mannose were predicted to be important for glucose control (Supplementary Fig. 12b). The regression coefficients for species and KOs correlated with glucose tAUC are shown in Supplementary Table 12. We found that several Clostridium species, such as Clostridium sp. D5, Clostridium sp. SS2 and Clostridium bartlettii were identified to be correlated with glucose tAUC. Consequently, our results revealed that glycemic response to a MMT was associated with the interaction of the gut microbiota and metabolism of individuals in this study.

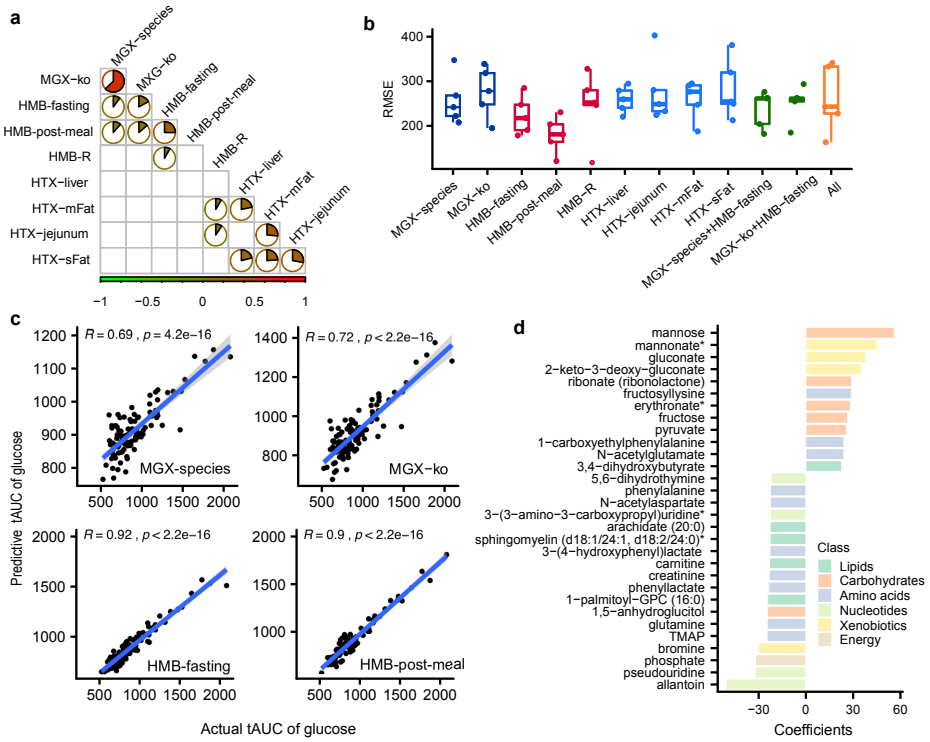


Figure 5: Predicting glucose response to a MMT by ridge regression models. a, The correlations between Bray-Curtis distance matrices of multi-omics. Mantel test was performed. The size of pie in the circle indicates the absolute value of correlation coefficient. Red and green colors represent positive and negative correlation coefficients, respectively. **b**, The performances of the ridge regression models evaluated by 5-fold cross-validation based on multi-omics and root mean square error (RMSE). **c**, The significant correlations between the actual glucose tAUC and the predicted glucose tAUC by ridge regression models using microbiota and metabolomics profiles, respectively. Spearman's rank correlation analysis was performed. **d**, The regression coefficients of the top 30 metabolites for predicting glucose tAUC based on post-meal metabolomics data. MGX-species, microbiota composition at species level; MGX-ko, microbiota KO function profile; HMB-fasting, metabolomic profile at fasting; HMB-post-meal, metabolomic profile after 2h MMT; HMB-R, the ratios of metabolite abundance at 2h post MMT to fasting, which means the postprandial metabolic changes; HTX-liver, HTX-jejunum, HTX-mFat, HTX-sFat indicate human transcriptional profiles from liver, jejunum, mesenteric and subcutaneous adipose tissues, respectively; MGX-species+HMB-fasting, the combination of microbiota composition and metabolomic profile at fasting; MGX-ko+HMB-fasting, the combination of microbiota KOs profile and metabolomic profile at fasting; All, the integration of all multi-omics data.

DISCUSSION

Here, we revealed that the global metabolic responses to a MMT were different in individuals with varied glucose tolerance status. From plasma metabolomic profiling we found more differential metabolites between the NGT and T2DM groups after the meal intake compared to fasting condition, thus enabling us to discover abnormal metabolism related to (pre) diabetes that did not appear at fasting condition. Furthermore, we identified three different types of response patterns in the 145 metabolites that were associated with diabetic status. Following the MMT, 39 metabolites were unaltered; these were mainly amino acid-derived metabolites including arginine and taurine, that had reduced abundances in the T2DM group.

Another 55 metabolites showed a parallel response to the MMT in the NGT, Pre-D and T2DM groups, including the carboxyethyl derivatives of BCAA and phenylalanine that increased with elevating glucose level during the MMT. Also, these carboxyethyl derivatives were positively correlated with HOMA2-IR. Consistently, several studies in both rodents and humans have observed alterations in BCAA and amino acid metabolites in relation to insulin resistance[16,33,36]. Increase in circulating BCAA in cardiometabolic disease are considered to result from decreased catabolism in adipose tissue and from inactivation of the branched-chain ketoacid dehydrogenase (BCKDH) complex in the liver[16]. In line, we observed alterations in BCAA metabolism (i.e., valine, leucine and isoleucine degradation) in liver and subcutaneous adipose tissue by gene set analysis. Previous studies have suggested that the gut microbiome of individuals with insulin resistance also has an increased capacity to produce amino acids and specifically BCAA[13,23]. By investigating the functional capacity of the gut microbiome, we also observed that amino acid metabolism (i.e., phenylalanine and phenylacetate metabolism) was enriched in the microbiome of individuals with insulin resistance. Interestingly, 1-carboxyethylphenylalanine correlated positively with microbial genes feaB, hcaC and paaE involved in phenylalanine metabolism. Microbial products of aromatic amino acid metabolism, in particular phenylacetic acid, has previously been linked to insulin resistance and thrombosis risk[37,38]. Recently it was reported that phenylalanine-derived metabolites increased after autologous fecal microbiota transplantation (FMT) in individuals with liver steatosis[39]. Through integrative analysis, the carboxyethyl derivatives of BCAA and phenylalanine correlated negatively with several Clostridium species, indicating that a reduction of this bacterial species might influence changes in the circulating metabolites. Another study reported that quantification of 1–carboxyethyvaline peptides of beta-hemoglobin can be useful for assessing glycemic status[40]. Thus, these carboxyethyl derivatives of amino acids could be potential biomarkers for (pre)diabetes. Although our results are associative in nature, we further strengthen the hypothesis that the

gut microbiome is capable of inducing alterations in circulating plasma metabolites. The strong correlations between circulating metabolites and insulin resistance, raise the question whether these metabolites serve as a biomarker or are causal agents in insulin resistance.

A total of 51 out of the 145 metabolites showed differential responses to the MMT in individuals with variable diabetic status. These metabolites might be involved in metabolic processes related to impaired adaptive response in (pre)diabetes, which commonly were not discovered at fasting. We observed different responses of carbohydrate metabolism (glucose, 1,5-anhydroglucitol, mannose) to the MMT, which is consistent with an earlier study[41]. Also, mannose has been identified as a biomarker of insulin resistance previously[31], which is in line with the fact that mannose is the most important metabolite for predicting glucose response in this study. Besides, several amino acids including N–acetylaspartate (NAA) responded differentially to the MMT. Interestingly, NAA involved in neuronal metabolism, was negatively correlated with HOMA2-IR, which is in agreement with a previous report[42]. NAA has been suggested to induce oxidative stress and nitric oxide (NO) production that has been reported to be associated with insulin resistance[42]. Thus, NO synthesis may be down-regulated due to the decreased NAA abundance in individuals with insulin resistance. Meanwhile, the decreased arginine abundance may explain the reduced NO synthesis from arginine in the T2DM group[43]. In addition, reduced NO/cGMP signaling has been demonstrated to contribute to insulin resistance[44]. Consistently, the cGMP/PKG signaling pathway was down-regulated in the T2DM group in the liver, along with the decreased abundances of NAA and arginine. Through integrative analysis, C. novyi showed significantly positive correlation with NAA and negative correlation with glucose response. Thus, our results suggest potential interplay between C. novyi, NAA and insulin resistance via the NO/cGMP signaling pathway.

Additionally, fatty acids including 3–hydroxydecanoate, 3–hydroxyoctanoate, linoleate (18:2n6) and 3–hydroxysebacate, had a higher abundance in the T2DM group and showed differential responses to the MMT, which have previously been reported to be associated with insulin resistance[45]. Furthermore, the responses of acylcholines to the MMT were different in individuals with insulin resistance and negatively correlated with glucose response. Previous studies have reported that acylcholines can act as agonist of muscarinic acetylcholine receptors (mAChRs) and play an important role in stimulating insulin secretion and maintaining glucose homeostasis[46]. Overall, the abnormal metabolism of carbohydrates, amino acids, fatty acids and acylcholines after a MMT in individuals with T2DM were revealed by metabolomic analysis in our study.

Interestingly, our results indicated that metabolic responses were associated with the metabolic status at baseline through integrative analysis. Therefore, differences in metabolic responses can be traced back to differences in other omics sets, such as liver, adipose tissue

and jejunum transcriptomics data. By mapping genes onto KEGG pathways, we observed alterations in several pathways involved in crucial metabolic and inflammatory pathways in mesenteric adipose tissue, such as MAPK, TNF, NF-kappa B signaling and cellular senescence. The MAPK and NF-kappa signaling pathways have been suggested to be activated by TNF-alpha in adipose tissue, which is associated with insulin resistance[46]. Moreover, data from several human clinical studies has shown a clear correlation between insulin resistance and cellular senescence[47]. Furthermore, a previous study has suggested that RyR2 channels regulate insulin secretion and glucose homeostasis[48]. Our results also showed that the gene RYR2 involved in insulin secretion was down-regulated in the T2DM group in liver, which is in accordance with the decreased HOMA2-B.

Several limitations of the current study must be acknowledged. Individuals with T2DM report a considerably higher number of medications than the NGT and Pre-D groups in this cohort, including glucose lowering agents, which may confound the T2DM-related changes in the gut microbiome and serum metabolomics (Supplementary Table 13). In addition, the discrepancy in sampling time points may influence the results of our integrative analysis, but in this study it is no more than three months and has limited effect. To reduce the surgical risk, individuals have lost weight before the operation, which might introduce relevant biases in particular pre-operative weight-loss. However, in contrast to most bariatric surgery trajectories, these individuals did not adhere to a specific diet. Moreover, overfitting may happen due to the limited sample number and high number of features when we predicted glucose response using multi-omics data. Also, these identified contributing factors for postprandial glucose response need to be further validated in a new cohort. Furthermore, the heterogeneity of prediabetes or T2DM was insufficiently investigated here, which has been demonstrated in previous studies[49].

In conclusion, our study systematically characterized the metabolic response to a MMT in individuals with different glucose tolerance, which provides new insights into the metabolic imbalance of (pre)diabetes. We first identified the abnormal metabolic processes related to (pre)diabetes after meal intake, including carbohydrates, amino acids, fatty acids and acylcholines. Further, we revealed that differences in metabolic responses could be traced back to other omics sets including fecal metagenomics and transcriptomics data of liver, adipose tissue and jejunum. Using machine learning models, we identified possible new biomarkers for glycemic control including NAA and phenylalanine derived metabolites. However, future studies should test whether these potential biomarkers can be used for the early identification of individuals that are at risk of developing T2DM. Also, further studies are needed to validate the biological causality of the identified metabolic imbalance of (pre) diabetes.

Acknowledgments

The BARIA study is funded by the Novo Nordisk Foundation (NNF15OC0016798). The Novo Nordisk Foundation Center for Basic Metabolic Research is supported by an unconditional grant (NNF10CC1016515) from the Novo Nordisk Foundation to University of Copenhagen. The BARIA study is a Scandinavian-Dutch collaboration. The computations and RNA Sequencing were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE (SNIC Computational Center of Chalmers University of Technology) partially funded by the Swedish Research Council through grant agreement no. 2018-05973. Funding from Knut and Alice Wallenberg Foundation is also acknowledged.

Contributions

T.W.S., M.N., J.N. and F.B. supervised this work. P.L. and B.J. performed data analyses and visualization. D.L. processed metabolomics and transcriptomics data. L.M.O. processed metagenomics data. A.S.M., O.A., S.C.B. and A.V.D.L. collected medical data and biopsies. V.T. and A.L. performed DNA, RNA and metabolomics isolations and optimizations. H.L., J.G. and K.K. helped with ridge regression models, gut microbiota and metabolomics analysis. L.E.O., F.B. and J.N. coordinated project administration. P.L. wrote the first draft. P.L., B.J., A.S.M., L.E.O., H.H., A.K.G., V.E.A.G., T.W.S., M.N., F.B. and J.N. conducted hypothesis generation, manuscript review and editing.

Availability of data and code

The related raw sequence files were submitted to public databases: the gut metagenomic data has been deposited in the European Nucleotide Archive (ENA) database under the accession number PRJEB47902; the transcriptomic data has been deposited in the European Genome-Phenome Archive (EGA) database with the access number EGAS00001005704, but the access to the transcriptomics data has been restricted to those who will get permission due to the GDPR. The metabolomic data has been deposited in the MetaboLights repository (access number xxx) but can be accessed by contacting the leading authors with signing required documents. In addition, scripts used for the processing and analysis of data and other materials that support the findings of this study can be provided upon request.

Competing interests

The authors declare no conflict of interests.

REFERENCES

1. American Diabetes Association. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care* **43**, S14–S31 (2020).
2. Tabák, A. G., Herder, C., Rathmann, W., Brunner, E. J. & Kivimäki, M. Prediabetes: A high-risk state for diabetes development. *The Lancet* **379**, 2279–2290 (2012).
3. Tabak, A. G. *et al.* Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet* **373**, 2215–2221 (2009).
4. Corpeleijn, E., Saris, W. H. M. & Blaak, E. E. Metabolic flexibility in the development of insulin resistance and type 2 diabetes: effects of lifestyle. *Obes Rev* **10**, 178–193 (2009).
5. Defronzo, R. A. Banting Lecture. From the triumvirate to the ominous octet: a new paradigm for the treatment of type 2 diabetes mellitus. *Diabetes* **58**, 773–95 (2009).
6. DeFronzo, R. A., Tobin, J. D. & Andres, R. Glucose clamp technique: A method for quantifying insulin secretion and resistance. *American Journal of Physiology Endocrinology Metabolism and Gastrointestinal Physiology* **6**, (1979).
7. Besser, R. E. J., Shields, B. M., Casas, R., Hattersley, A. T. & Ludvigsson, J. Lessons from the mixed-meal tolerance test: use of 90-minute and fasting C-peptide in pediatric diabetes. *Diabetes Care* **36**, 195–201 (2013).
8. Laiteerapong, N. *et al.* The Legacy Effect in Type 2 Diabetes: Impact of Early Glycemic Control on Future Complications (The Diabetes & Aging Study). *Diabetes Care* **42**, 416–426 (2019).
9. Ahmed, M., Gannon, M. C. & Nuttall, F. Q. Postprandial plasma glucose, insulin, glucagon and triglyceride responses to a standard diet in normal subjects. *Diabetologia* **12**, 61–67 (1976).
10. Pellis, L. *et al.* Plasma metabolomics and proteomics profiling after a postprandial challenge reveal subtle diet effects on human metabolic status. *Metabolomics* **8**, 347–359 (2012).
11. Shankar, S. S. *et al.* Standardized Mixed-Meal Tolerance and Arginine Stimulation Tests Provide Reproducible and Complementary Measures of β-Cell Function: Results From the Foundation for the National Institutes of Health Biomarkers Consortium Investigative Series. *Diabetes Care* **39**, 1602–1613 (2016).
12. Zeevi, D. *et al.* Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079–1095 (2015).
13. Pedersen, H. K. *et al.* Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* (2016) doi:10.1038/nature18646.
14. Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).
15. Zhou, W. *et al.* Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* **569**, 663–671 (2019).
16. White, P. J. & Newgard, C. B. Branched-chain amino acids in disease. *Science* **363**, 582–583 (2019).
17. Fadista, J. *et al.* Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc Natl Acad Sci U S A* **111**, 13924–13929 (2014).
18. Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
19. Thingholm, L. B. *et al.* Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell Host Microbe* (2019) doi:10.1016/j.chom.2019.07.004.
20. van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *J Intern Med* joim.13157 (2020) doi:10.1111/joim.13157.
21. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, T. R. Homeostasis model assessment insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412–9 (1985).
22. Brodovicz, K. G. *et al.* Postprandial metabolic responses to mixed versus liquid meal tests in healthy men and men with type 2 diabetes. *Diabetes Res Clin Pract* **94**, 449–455 (2011).
23. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).

24. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods 2016 13:9* **13**, 731–740 (2016).

25. Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276–278 (2020).

26. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15 (2013).

27. di Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316–319 (2017).

28. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* (2017) doi:10.1038/nbt.3960.

29. Karlsson, F. H., Nookaew, I. & Nielsen, J. Metagenomic Data Utilization and Analysis (MEDUSA) and Construction of a Global Gut Microbial Gene Catalogue. *PLoS Comput Biol* **10**, (2014).

30. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).

31. Lee, S. *et al.* Integrated Network Analysis Reveals an Association between Plasma Mannose Levels and Insulin Resistance. *Cell Metab* **24**, 172–184 (2016).

32. Yu, D. *et al.* Plasma metabolomic profiles in association with type 2 diabetes risk and prevalence in Chinese adults. *Metabolomics* **12**, 1–11 (2016).

33. Lynch, C. J. & Adams, S. H. Branched-chain amino acids in metabolic signalling and insulin resistance. *Nature Reviews Endocrinology* vol. 10 Preprint at https://doi.org/10.1038/nrendo.2014.171 (2014).

34. Jang, C. *et al.* A branched-chain amino acid metabolite drives vascular fatty acid transport and causes insulin resistance. *Nat Med* **22**, 421–426 (2016).

35. Doumatey, A. P. *et al.* Gut Microbiome Profiles Are Associated With Type 2 Diabetes in Urban Africans. *Front Cell Infect Microbiol* **10**, 63 (2020).

36. Newgard, C. B. *et al.* A Branched-Chain Amino Acid-Related Metabolic Signature that Differentiates Obese and Lean Humans and Contributes to Insulin Resistance. *Cell Metab* **9**, 311–326 (2009).

37. Hoyles, L. *et al.* Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat Med* **24**, 1070–1080 (2018).

38. Nemet, I. *et al.* A Cardiovascular Disease-Linked Gut Microbial Metabolite Acts via Adrenergic Receptors. *Cell* (2020) doi:10.1016/j.cell.2020.02.016.

39. Witjes, J. J. *et al.* Donor Fecal Microbiota Transplantation Alters Gut Microbiota and Metabolites in Obese Individuals With Steatohepatitis. *Hepatol Commun* hep4.1601 (2020) doi:10.1002/hep4.1601.

40. Jagadeeshaprasad, M. G. *et al.* Targeted quantification of N-1-(carboxymethyl) valine and N-1-(carboxyethyl) valine peptides of β-hemoglobin for better diagnostics in diabetes. *Clin Proteomics* **13**, 1–11 (2016).

41. Wopereis, S. *et al.* Multi-parameter comparison of a standardized mixed meal tolerance test in healthy and type 2 diabetic subjects: the PhenFlex challenge. *Genes Nutr* **12**, (2017).

42. Daniele, G. *et al.* Plasma N-Acetylaspartate Is Related to Age, Obesity, and Glucose Metabolism: Effects of Antidiabetic Treatment and Bariatric Surgery. *Front Endocrinol (Lausanne)* **11**, (2020).

43. Tessari, P. *et al.* Nitric oxide synthesis is reduced in subjects with type 2 diabetes and nephropathy. *Diabetes* **59**, 2152–2159 (2010).

44. Rizzo, N. O. *et al.* Reduced NO-cGMP signaling contributes to vascular inflammation and insulin resistance induced by high-fat feeding. *Arterioscler Thromb Vasc Biol* **30**, 758–765 (2010).

45. Roden, M. *et al.* Mechanism of free fatty acid-induced insulin resistance in humans. *Journal of Clinical Investigation* **97**, 2859 (1996).

46. Gautam, D. *et al.* A critical role for beta cell M3 muscarinic acetylcholine receptors in regulating insulin release and blood glucose homeostasis in vivo. *Cell Metab* **3**, 449–461 (2006).

47. Palmer, A. K., Gustafson, B., Kirkland, J. L. & Smith, U. Cellular senescence: at the nexus between ageing and diabetes. *Diabetologia* (2019) doi:10.1007/s00125-019-4934-x.

48. Santulli, G. *et al.* Calcium release channel RyR2 regulates insulin release and glucose homeostasis. *J Clin Invest* **125**, 1968–1978 (2015).

49. Olofsson, L. E. & Bäckhed, F. The Metabolic Role and Therapeutic Potential of the Microbiome. *Endocr Rev* **43**, 907–926 (2022).

Supplementary data available at

https://www.biorxiv.org/content/10.1101/2022.04.26.489057v1

7

A systems biology approach to study non-alcoholic fatty liver (NAFL) in women with obesity

Abraham S. Meijnikman#, Dimitra Lappa#, Hilde Herrema, Omrum Aydin, Kimberly A. Krautkramer, Valentina Tremaroli, Louise E. Olofsson, Annika Lundqvist, Sjoerd Bruin, Yair Acherman, Joanne Verheij, Siv Hjorth, Victor E. A. Gerdes, Thue W. Schwartz, Albert K. Groen, Fredrik Bäckhed, Jens Nielsen, Max Nieuwdorp

# equal contribution

iScience, 2022 Aug 5;25(8):104828.

SUMMARY

Non-Alcoholic Fatty Liver Disease (NAFLD) is now the most frequent global chronic liver disease. Individuals with NAFLD exhibited an increased risk of all-cause mortality driven by extrahepatic cancers and liver and cardiovascular disease. Once the disease is established, women have a higher risk of disease progression and worse outcome. It is therefore critical to deepen the current knowledge on the pathophysiology of NAFLD in women. Here, we used a systems biology approach to investigate the contribution of different organs to this disease. We analyzed transcriptomics profiles of liver and adipose tissues, fecal metagenomes and plasma metabolomes of 55 women with and without NAFLD. We observed differences in metabolites, expression of human genes and gut microbial features between the groups and revealed that there is substantial crosstalk between these different omics sets. Multi-omics analysis of individuals with NAFLD may provide novel strategies to study the pathophysiology of NAFLD in humans.

INTRODUCTION

As a consequence of the pandemic spread of obesity, NAFLD is now recognized as the most prevalent chronic liver disease worldwide[1]. In the general population, one in four individuals is affected by NAFLD; this prevalence increases to over 80% in individuals with obesity[1]. NAFLD comprises a spectrum of clinical and histopathological abnormalities. These include simple steatosis and steatosis with mild inflammation (non-alcoholic fatty liver, NAFL) as well as steatosis with ballooning and inflammation (non-alcoholic steatohepatitis, NASH). Accumulation of fat in hepatocytes has long been considered a relatively benign condition. However, an estimated 30% of people with NAFL will develop NASH, a progressive form of liver disease that can lead to liver fibrosis, cirrhosis and hepatocellular carcinoma[2]. Advanced forms of NASH often require liver transplantation and are the main cause of liver-related deaths in NAFLD[1]. A recent report of a large nationwide cohort study investigating overall and cause specific mortality in long-term follow up of individuals with NAFLD however, showed that individuals with NAFL also exhibited an increased risk of all-cause mortality driven by extrahepatic cancers and liver and cardiovascular disease (CVD)[3]. Of concern, especially women with NAFLD are more susceptible to develop excess CVD events compared to age matched men[4]. In fact, NAFLD has a cardiovascular aging effect of approximately 18 years in women. Moreover, in general, women have a lower risk of developing NAFLD, but once the disease is established, women have a higher risk of disease progression[5]. The rapidly growing prevalence of NAFLD and lack of effective treatment options to tackle this potentially debilitating disease, will further increase obesity-related burden on public health and economies. In order to develop appropriate, sex-specific non-invasive diagnostic methods and treatment options, it is critical to deeply investigate the complex pathophysiology of NAFLD.

The underlying mechanisms that govern hepatic lipid accumulation and the predisposition to inflammation and fibrosis are complex and multifactorial, which is recapitulated in the multi-hit hypothesis that implicates that a myriad of factors are acting in a parallel and synergistic manner[6]. These factors include: insulin resistance, adipocyte dysfunction, genetic variants, bile acid metabolism, the gut microbiome, and lipotoxicity[7,8]. The complexity of the contributing factors can mask different structural associations between metabolic activities in different tissues, prohibiting in-depth insight into molecular mechanisms underlying disease development. By applying a systems biology approach using multi-omics data, it is possible to deep phenotype individuals with or without metabolic diseases and, through data integration, identify the crosstalk between different relevant biological layers.

We here used a global approach to investigate factors that may contribute to NAFL development in women. Our systems biology approach allowed for integration of transcriptomics, metagenomics and plasma metabolomics datasets from obese women with and without NAFL. Analyses of these integrated omics sets revealed a robust NAFL-signature and highlight the additive value of a multi-omics approach to study NAFL pathophysiology.

Method details

Ethical approvals and patients' clinical information

The recruitment of participants was conducted from the BARIA study[11] with a total of 55 individuals included. The baseline characteristics of these participants are described in ***Table 1***. The study was performed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC (Trialregister: BARIA study NL8983). All participants provided written informed consent.

Material collection

Individuals underwent a complete metabolic work-up at the start of their bariatric surgery trajectory. Anthropometric measurements including height, weight and waist and hip circumference were taken. In addition, body fat percentage using bioelectrical impedance and blood pressure were measured. Fasting blood samples were used for the determination of hemoglobin, HbA1c, glucose, lipid profile, alanine aminotransferase, aspartate aminotransferase, insulin, and creatinine levels. Within three months before surgery, a 2-hour mixed meal tolerance test (MMT) was performed to assess insulin resistance and investigate dynamic alterations in circulating metabolites. The MMT consisted of two Nutridrink compact 125ml (Nutricia®), containing in total 23.3 grams fat, 74.3 grams carbohydrates (of which 38.5 grams sugar) and 24.0 grams protein. The participants received this meal after fasting for a minimum of nine hours. Time point zero refers to the moment at which the participant had fully consumed the meal. Blood samples were drawn *via* an intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes. All samples were stored at -80°C until further processing.

Liver biopsies and histology

Liver histological sections were stained with Haematoxylin-Eosin and Sirius red and then reviewed by members of the Dutch Liver Pathology Panel after training sessions for NAFLD according to the Steatosis, Activity and Fibrosis (SAF) score[9]. Difficult borderline cases were discussed during panel meetings for consensus. NAFLD was categorized into NAFL when steatosis was present in >5% of hepatocytes alone or with mild inflammation but without ballooning, or NASH when steatosis was present in >5% of hepatocytes and if ballooning

and inflammation were both present in the biopsy. In the present study, no individuals were diagnosed with NASH based on histology.

Metabolome processing

EDTA plasma samples under fasting and two-hours after a MMT postprandial conditions were collected from 55 BARIA participants. All EDTA plasma samples were shipped to METABOLON (Morisville, NC, USA) for performing analysis using ultra high-performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) untargeted metabolomics. The metabolomic counts obtained, underwent significant curation via metabolites' pre-filtering, imputation for subsets of metabolites' missing values and data normalization, in order to minimize the effect of artifacts in the downstream analysis. Metabolomics prefiltering and imputation were performed by utilizing a variation of the Perseus platform[10]. Essentially, data has been pre-filtered so as to have a maximum of 25% missing values for a metabolite across all samples. This was followed by a log transformation of all the measured metabolites' raw intensities across the entire dataset. Then, we calculated the total data mean and standard deviation (by omitting missing values). Taking into account that the metabolite intensities distribution is approximately following normality, we chose a small distribution 2.5 standard deviations away from the original data mean towards the left tail of the original data distribution, with 0.5 standard deviations width. This new shrunken range corresponds to the actual lowest level of detection by the spectrometer. Here by drawing random values from this mini distribution, we fill the missing prefiltered data of choice. Originally METABOLON measured 1345 metabolites, but after applying previously mentioned methodology of imputation and normalization we included 988 metabolites for fasting metabolome and 1018 metabolites for postprandial metabolome[11].

Transcriptome processing

Biopsies from liver (55 samples), mesenteric adipose tissue (54 samples) and subcutaneous adipose tissue (55 samples) were collected at the time of the bariatric surgery. RNA was extracted from biopsies using TriPure Isolation Reagent (Roche, Basel, Switzerland) and Lysing Matrix D, 2 mL tubes (MP Biomedical, Irvine, CA, USA) in a FastPrep®-24 Instrument (MP Biomedical, Irvine, CA, USAs) with homogenization for 20 seconds at 4.0 m/sec, with repeated bursts until no tissue was visible; homogenates were kept on ice for 5 minutes between homogenization bursts if multiple cycles were needed. RNA was purified with chloroform (Merck, Darmstadt, Germany) in phase lock gel tubes (5PRIME) with centrifugations at 4°C, and further purified and concentrated using the RNeasy MinElute kit (Qiagen, Venlo, The Netherlands). The quality of RNA was analysed on a BioAnalyzer instrument (Agilent), with quantification on Nanodrop (Thermo Fisher Scientific, Waltham,

MA, USA). Due to degradation of the RNA, libraries for RNAseq sequencing were prepared by rRNA depletion; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end reads and 10G data/sample. The average read count per sample from liver was 42 ± 15 million. For mesenteric and subcutaneous adipose tissue, the average read count per sample were 43.2 ± 20 million. The extracted fastq files were analyzed with nf-core/rnaseq (Ewels et al. 2020), a bioinformatics analysis pipeline used for RNA sequencing data. The workflow processed raw data from FastQ inputs (FastQC, TrimGalore!), aligned the reads (STAR) with *Homo sapiens* GRCh38 as reference genome, generates gene counts (featureCounts, StringTie) and performed extensive quality-control on the results (RSeqQC, dupRadar, Preseq, edgeR, multiQC). The pipeline was built using Nextflow.

Microbiome processing

Fecal samples from 55 participants were collected on the day of surgery and immediately frozen at -80C. Total fecal genomic DNA was extracted from 100 mg of feces using a modification of the IHMS DNA extraction protocol Q[13]. Briefly, fecal samples were extracted in Lysing Matrix E tubes (MP Biomedicals) containing ASL buffer (Qiagen, Venlo, The Netherlands), and lysis of cells was obtained, after homogenization by vortexing for 2 minutes, by two cycles of heating at 90°C for 10 minutes followed by three bursts of bead beating at 5.5 m/sec for 60 seconds in a FastPrep®-24 Instrument (MP Biomedicals). After each bead-beating burst, samples were placed on ice for 5 minutes. The supernatants containing fecal DNA were collected after the two cycles by centrifugation at 4°C. Supernatants from the two centrifugations steps were pooled and a 600 µL aliquot from each sample was purified using the QIAamp DNA Mini kit (QIAGEN, Venlo, The Netherlands) in the QIAcube (QIAGEN Venlo, The Netherlands) instrument using the procedure for human DNA analysis. Samples were eluted in 200 µL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0). Libraries for shotgun metagenomic sequencing were prepared by a PCR-free method; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc. San Diego, CA, USA) with 150 bp paired-end reads and 6G data/sample.

MEDUSA pipeline was used for pre-processing of raw shotgun metagenomics sequence data. MEDUSA is an integrated pipeline for analysis of short metagenomic reads, which maps reads to reference databases, combines output from several sequencing runs and manipulates tables of read counts. The input number of total reads from the metagenome analysis were on average 23.4±2.2 million reads per sample and the total aligned reads 16.6±1.8 million reads per sample. The sequencing runs had high quality with almost 98% of the reads passing the quality cut-off. Out of the high-quality reads, on average 0.04% aligned

to the human genome, although the data had been cleaned for human reads. Out of the high quality non-human reads, 78.4% aligned to the MEDUSA's software gene catalogue. Quality filtered reads were mapped to a genome catalogue and gene catalogue using Bowtie2[14].

Quantification and statistical analysis

Differential analysis of the plasma metabolome was conducted with two methods: ANOVA and Kruskal Wallis, with the use of HybridMTest package, that performs hybrid multiple testing using Empirical Bayes Probability (EBP). The cut-off significance level of $P<0.1$ was used for identifying differentially significant metabolites with an adjusted EBP value.

Differential gene expression analysis for individuals with and without NAFL was performed for liver, subcutaneous adipose and mesenteric adipose tissues, respectively, in R with DESeq2 package[15]; log normalization is based on gene counts geometric distribution. The statistical analysis method for calculating differential expression rates is the Wald test. After False discovery rate (FDR) correction with multiple hypothesis testing with IHW package[16], we analyzed genes with $P<0.1$ by DEGreport's degPatterns function, to identify subgroups of co-expressed genes between individuals with and without NAFL. For these differentially significant co-expressed genes we performed gene enrichment with Enrichr platform[17] using KEGG metabolic pathways[18].

Statistical analysis on the metagenomics data was performed on rarefied count, (20 M reads per sample). The taxon ids were input to taxize package[19], to get full taxonomic information and ranking for the species. This dataset was input to DESeq2 and phyloseq packages[20] for conducting downstream differential statistical analysis. Similar to the BARIA transcriptomics counts, log normalization based on gene counts geometric distribution has been conducted with it. Statistical analysis methods for calculating differential expression rates was Wald Test. The IHW package, as part of DESeq2 suite, was utilized for multiple hypothesis testing and adjusting the respective P values, with alpha significance threshold set at $P<0.1$.

Multi-omics integrative analysis has been conducted with DIABLO. DIABLO extends sparce Generalized Canonical Correlation Analysis (sGCCA), uses singular value decomposition for dimensionality reduction and selects co-expressed (correlated) variables that can explain the categorical outcome of interest21, in our case non-NAFL or NAFL. DIABLO output a set of latent variables (components) based on the dimensionality of the input datasets. The chosen number of components could extract sufficient information to discriminate all phenotype groups. Then, a set of coefficients was attributed to each variable, that indicated the importance of each variable in DIABLO. The goal was to have maximization of the covariance between a linear combination of the variables from each input dataset and each categorical outcome. After tuning these two hyperparameters, DIABLO output

a list of selected variables from each omic dataset, associated to each component, that could distinguish the given phenotypes.

Key resources material

Reagent or Resource	Source	Identifier
Biological samples		
Human fecal metagenomics data	BARIA cohort (PI prof M. Nieuwdorp)	[11]
Human liver RNA sequencing data	BARIA cohort (PI prof M. Nieuwdorp)	[11]
Human subcutaneous adipose tissue sequencing data	BARIA cohort (PI prof M. Nieuwdorp)	[11]
Human visceral adipose tissue sequencing data	BARIA cohort (PI prof M. Nieuwdorp)	[11]
Human plasma metabolomics data	BARIA cohort (PI prof M. Nieuwdorp)	[11]
Deposited data		
Liver and adipose tissue transcriptomics	European Nucleotide Archive	ENA PRJEB47902
Fecal metagenomics	European Genome-Phenome Archive	EGAS00001005704
Software and algorithms		
MEDUSA pipeline	n/a	[22]
Bowtie2	n/a	[14]
DESeq2	n/a	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
phyloseq	n/a	https://bioconductor.org/packages/release/bioc/html/phyloseq.html
DIABLO	n/a	http://mixomics.org/mixdiablo/
Other		
HiSeq instrument	Illumina	N/A
DNA extraction kit	QIAamp DNA Mini kit	N/A

RESULTS

To take a comprehensive approach to investigate factors that may contribute to NAFL development, we included individuals from our bariatric surgery cohort (the BARIA study)[11], but excluded type 2 diabetes mellitus (T2DM) patients to avoid confounding effects of long-term hyperglycemia or medication use. Since there are strong sex differences in hepatocellular and systemic processes in the pathophysiology and progression of NAFLD[23,24], we focused on women. The study cohort comprised of 55 women for whom a multi-omics dataset was available, including fasting and two-hour post mixed meal test (MMT) plasma metabolome, liver and adipose tissue (subcutaneous and mesenteric) transcriptome, along with gut microbial metagenome. In addition, we analyzed the glucose and insulin response

during the MMT before and one year after bariatric surgery to investigate differences in glucose metabolism between women with and without NAFL.

In total, 23 individuals (BMI 39.4 ± 3.0 kg/m^2, age 45 ± 11 years) fulfilled the criteria for NAFL (biopsy-proven) whereas 32 individuals (BMI 40.2 ± 4.7 kg/m^2, age 41 ± 10 years) had no NAFL (***Table 1***). NAFL ranged from grade 1 to grade 2 steatosis; none of our individuals had hepatocyte ballooning, a prerequisite for NASH diagnosis according to the SAF criteria[9]. As expected, the ALT levels were increased in the NAFL group, whereas comorbidities such as insulin resistance (as assessed by MMT) and medication did not differ between groups, indicating a homogenous study population (***Table 1***, ***Figure S1 and S2***).

Characteristics	Non-NAFL = 32	NAFL = 23
Demographic		
Age *(years)*	41 ± 10	45 ± 11
Anthropometric		
BMI *(kg/m2)*	40.2 ± 4.7	39.4 ± 3.0
Type 2 diabetes mellitus *(n)*	0	0
Clinical lab values (normal range)		
ALP *(30-135 U/L)*	85 ± 21	84 ± 19
g-GT *(10 - 40 IU/l)*	26 (18-26)	28 (18-41)
ALT *(0 - 50 IU/l)*	25 (18-27)	36 (22-42)*
AST *(0 - 35 IU/l)*	22 ± 4	26 ± 6
FPG *(<5.6 mmol/l)*	5.4 ± 0.5	5.6 ± 0.6
HbA1c *(<5.6%)*	5.4 ± 0.3	5.6 ± 0.2
HbA1c *(mmol/mol)*	35 ± 3	37 ± 2
Total cholesterol *(1.5 - 6.5 mmol/l)*	4.9 ± 1.1	4.9 ± 1.1
Triglycerides *(<1.7 mmol/l)*	1.4 (0.9-1.5)	1.7 (1.1-1.9)
HDL cholesterol *(≥1.0 mmol/l)*	1.3 ± 0.4	1.2 ± 0.3
LDL cholesterol *(< 3.0 mmol/l)*	3.1 ± 1.1	3.2 ± 0.8
Histological parameters (number)		
Steatosis grade score *(0,1,2,3)*	32,0,0,0	0,22,1,0
Lobular inflammation score *(0,1,2)*	14,17,1	0,21,2
Hepatocyte ballooning score *(0,1,2)*	32,0,0	23,0,0

Table 1. Baseline characteristics of the 55 women included. Data is expressed as mean ± standard deviation or as median (interquartile range) depending on normality of the data. For histological scores, the number of individuals with a certain score is shown according to the Steatosis Activity and Fibrosis score (SAF). NAFL: Non-Alcoholic Fatty Liver, BMI: body mass index, ALP: alkaline phosphatase, g-GT: gamma glutamyl transferase, ALT: alanine aminotransferase, AST: aspartate aminotransferase, FPG: fasting plasma glucose, HbA1c: Hemoglobin A1c, HDL: high-density lipoprotein, LDL: low-density lipoprotein. *indicate significant ($p<0.05$) difference. Significance was calculated by either independent T test or Mann-Whitney U test depending on normality.

The gut microbial communities of individuals with and without NAFL significantly differ

To characterize the gut microbiome in individuals with and without NAFL we performed whole-genome shotgun sequencing of the fecal DNA and used MEDUSA to obtain taxonomic information[22]. In order to assess if there is a difference in microbial alpha-diversity between individuals with and without NAFL, we used a series of different metrics (Observed, Chao1, ACE, Shannon, Simpson, Inverse Simpson). According to the alpha diversity metrics, the microbial diversity was similar in the two groups, which is in contrast to previous reports[25,26], (***Figure S3*****).** These previous observations analyzed individuals with a more progressive form of NAFLD (*i.e.,* NASH). In agreement with previous studies[27–29], we observed that the microbiome was dominated by Firmicutes in individuals with NAFL, while Bacteroidetes was the most dominant phylum in individuals without NAFL (***Figure 1a***). We next assessed the bacterial species composition per individual (***Figure S4*****).** Even though we observed large inter-individual variation in the gut microbiome composition, PERMANOVA and beta dispersion analysis revealed that the two groups largely spatially overlap but have different centroids and different dispersions (***Figure S5***). In total 57 bacterial species differed significantly between individuals with and without NAFL (***Figure 1a***). Three bacterial species were at least twice as abundant in individuals with NAFL (***Table S1***). One of these species belonged to the phylum Actinobacteria (*Collinsela stercoris*) whereas two belonged to Firmicutes (*Lactobacillus buchneri, Lactobacillus iners*). In individuals without NAFL, 11 bacterial species were at least twice as abundant compared to individuals with NAFL. Of these 11 bacterial species, six belonged to the phylum Bacteroidetes, (*Prevotella oulorum, Prevotella sp. oral taxon 317, Prevotella sp. Oral taxon 472, Prevotella multisaccharivorax, Prevotella dentalis* and *Prevotella bryantii*); two belonged to Firmicutes (*Lactobacillus delbrueckii, Enterococcus casseliflavus*) and three belonged to Proteobacteria (*Citrobacter rodentium, Yersinia enterocolitica* and *Haemophilus pittmaniae*). In summary, even though alpha diversity did not differ significantly between the two groups, 57 bacterial species differed significantly and mainly belonged to the Bacteroidetes and Firmicutes phylum.

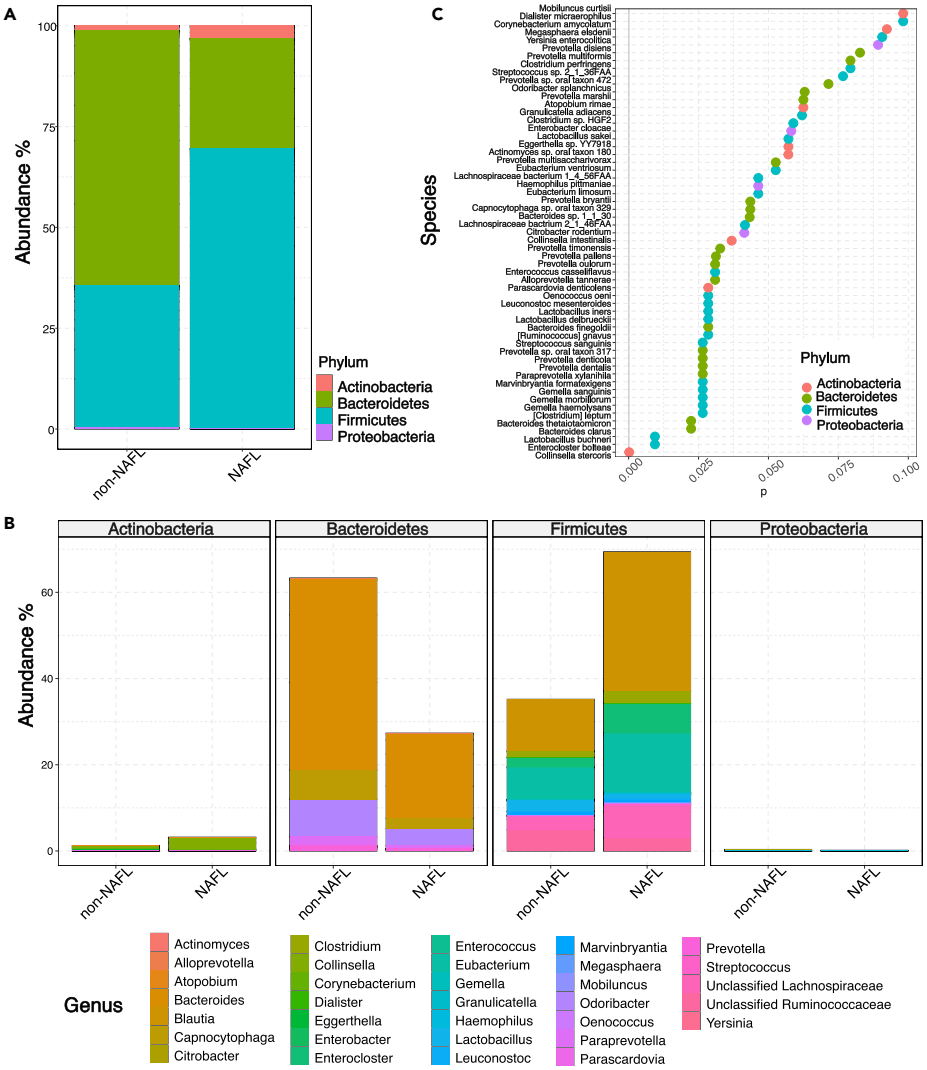


Figure 1. Microbial species and phyla between individuals with and without NAFL. (a) Difference in total abundance of bacterial species indicated at the Phylum level between individuals with and without NAFL. **(b)** Relative abundance and distribution within of differentially significant microbial species between individuals with and without NAFL. **(c)** 57 differentially significant microbial species between individuals with and without NAFL, after differential microbial species analysis with DESeq2 (adjusted $P < 0.1$) Likelihood Ratio Test for significance.

The NAFL-associated metabolome is characterized by increased lipid and amino acids in postprandial conditions

Since microbiome-associated factors such as microbial metabolites are more and more recognized as disease-modifying factors, including in NAFL development[8], we performed plasma metabolomics analyses on fasting and post MTT samples to reveal metabolite-based phenotypes of NAFL. Out of 988 metabolites, phosphathidylcholine 1-palmitoyl-2-arachidonoyl-GPC (16:0/20:4n6) was the only significantly altered metabolite in fasted individuals and was lower in individuals with NAFL (***Figure 2, Table S2***). Since humans rarely reside in a fasting state for a long period of time, the liver is continuously exposed to nutrients and (microbial) metabolites from the intestine. Thus, postprandial plasma samples, might provide a more representative view on circulating metabolites in individuals with NAFL. Indeed, seven metabolites differed significantly in the postprandial state. Five metabolites were more abundant in NAFL and two were more abundant in non-NAFL individuals (***Figure 2, Table S3***). Two sphingomyelin metabolites were decreased in individuals with NAFL whereas diacylglycerol, a signaling lipid previously linked to hepatic insulin resistance and NAFLD[30], was more abundant in individuals with NAFL. In agreement with previous studies demonstrating that circulating amino acids are increased in individuals with NAFL[31,32], the branched-chain amino acids (BCAA) derivatives 1–carboxyethylisoleucine and 1-carboxyethylvaline were increased in individuals with NAFL. Alterations in amino acids in cardiometabolic disease have been tightly linked to insulin resistance. Since both insulin and glucose levels did not differ during the MMT (***Figure S2)***, our data suggests that these alterations may be independent of altered glucose metabolism.

Distinct transcriptional profiles in liver, subcutaneous and mesenteric adipose tissue

Since several studies have demonstrated that (microbial) metabolites exert metabolic actions on distal tissues and organs[33], we profiled hepatic, mesenteric and subcutaneous adipose tissue transcriptomes to improve our understanding of the interrelation between alterations in the plasma metabolome and gene expression. By using DESeq2[15], we identified differently expressed genes between individuals with and without NAFL. Analyses of the hepatic transcriptome identified 52 genes that were differently expressed between individuals with and without NAFL. Of these genes, 13 were upregulated and 39 were downregulated in individuals with NAFL compared to individuals without NAFL (***Table S4*)**. KEGG pathway enrichment analysis using EnrichR[17] identified that pathways involved in several cancers were enriched in individuals with NAFL, which may indicate increased cell proliferation. Furthermore, the hypoxia-inducible factor 1 (HIF-1) signaling pathway, which has previously been linked to NAFLD pathogenesis[34], was enriched in the liver of individuals with NAFL. The only significant pathway that was enriched in individuals without NAFL was the

pathway involved in arginine and proline metabolism (***Table 2***). Since adipose tissue and the liver communicate with each other[35],we next investigated the transcriptome of two different adipose tissue depots. In subcutaneous adipose tissue, 19 genes were significantly different between the groups of which 15 were higher in individuals with NAFL and four were higher in individuals without NAFL (***Table S5***). The mesenteric adipose tissue transcriptome revealed that 56 genes differed significantly between individuals with and without NAFL. Of these, 34 genes were upregulated, and 22 genes were downregulated in individuals with NAFL compared to individuals without NAFL ***(Table S6)***.

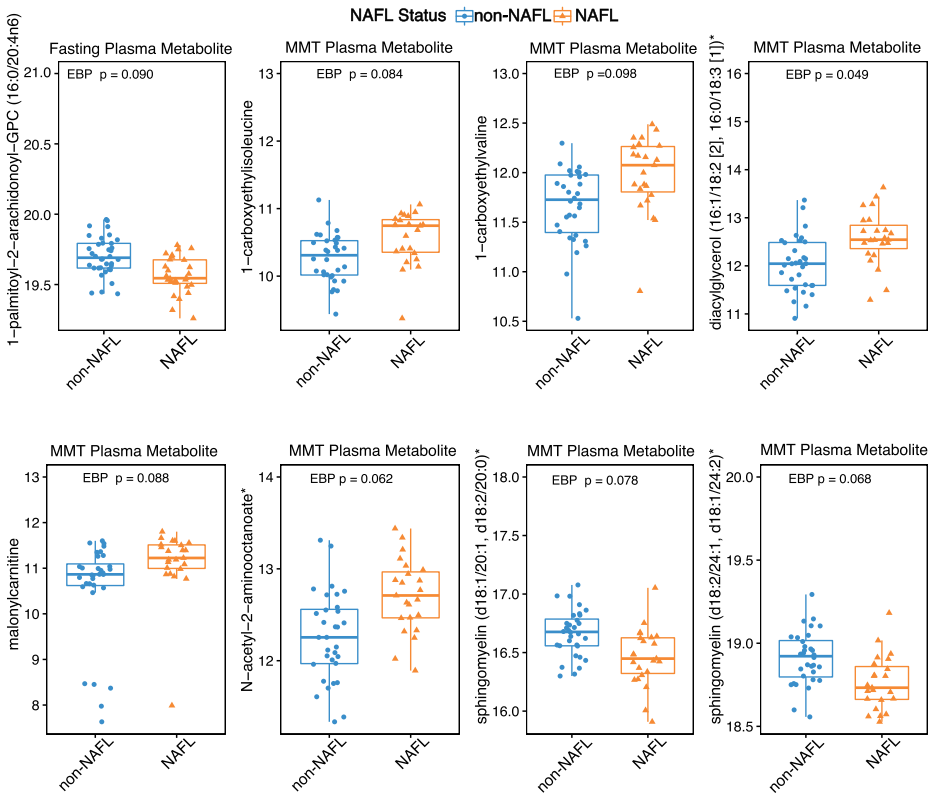


Figure 2. Log scale abundance of differentially significant metabolites between individuals with and without NAFL in fasting and postprandial plasma metabolomics. Differential metabolite analysis was conducted with the HybridMtest package and p-adjusted based on Estimated Bayesian Probability (P<0.1).

According to KEGG pathway analysis, interleukin (IL)-17, advanced glycation end products (AGE) and tumor necrosis factor (TNF)-signaling pathways were enriched in individuals with NAFL in subcutaneous adipose tissue, whereas response to oxidative stress was not enriched underscoring the well-established link between adipose tissue

inflammation and NAFL[36] (***Table 2)***. In mesenteric adipose tissue, carbohydrate, galactose, sucrose and protein metabolism pathways were enriched in mesenteric adipose tissue from individuals with NAFL, while pathways involved in infectious disease were not enriched (***Table 2***). Furthermore, pathways associated with fat digestion and absorption were enriched in individuals with NAFL. This further strengthens the link between alterations in diacylglycerol and adipose tissue dysfunction. Transcriptome analyses from all three tissues showed distinct differences in gene expression and pathways relevant for the development of NAFL such as HIF-1 signaling, inflammation and fat digestion and absorption.

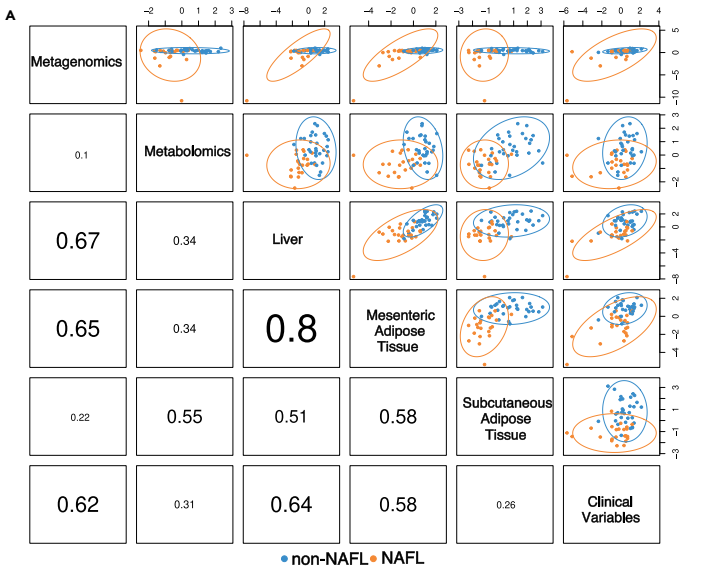
Tissue	Regulation	Pathway	P-value
Liver	Upregulated in NAFL	HIF-1 signalling pathway	0.0019
		Bladder cancer	0.026
		Endometrial cancer	0.037
		Central carbon metabolism in cancer	0.041
		Non-small cell lung cancer	0.042
		Arginine and proline metabolism	0.089
	Downregulated in NAFL	Pyrimidine metabolism	0.103
		Cortisol synthesis and secretion	0.116
		Bile secretion	0.128
		Drug metabolism	0.186
Mesenteric adipose tissue	Upregulated in NAFL	Galactose metabolism	2.119E-7
		Carbohydrate digestion and absorption	5.6684E-5
		Protein digestion and absorption	4.767E-4
		Starch and sucrose metabolism	0.002
		Fat digestion and absorption	0.002
	Downregulated in NAFL	Prion diseases	0.036
		Legionellosis	0.056
		Complement and coagulation cascades	0.080
		Systemic lupus erythematosus	0.1308
		Herpes simplex virus 1 infection	0.407
Subcutaneous adipose tissue	Upregulated in NAFL	IL-17 signalling pathway	4.253E-5
		AGE-RAGE signalling pathway in diabetic complications	5.283E-5
		TNF signalling pathway	7.019E-5
		Prion diseases	3.079E-4
		African trypanosomiasis	3.444E-4
	Downregulated in NAFL	Regulation of response to oxidative stress	0.002
		Regulation of response to stress	0.002
		Positive regulation of G2/M transition of mitotic cell cycle	0.003
		Positive regulation of cell cycle G2/M phase transition	0.003
		Positive regulation of peptidyl-threonine phosphorylation	0.005

Table 2: KEGG metabolic pathways up or downregulated in individuals with and without NAFL

Multi-omics integration creates a signature for NAFL

The individual omics datasets thus far revealed differences in the fecal metagenome, the plasma metabolome and tissue transcriptome between the groups. However, discriminative, analyses of these individual omics sets do not provide insight in the interrelation between the different biological layers. We therefore constructed a multivariate model to identify crosstalk events between these different tissues and metagenomic, metabolomic and clinical datasets by fitting a sparse Partial Least Squares discriminant analysis with Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO)[37]. DIABLO simultaneously calculates the correlations between all input omics datasets and selects a minimal set of input variables that differentiate between individuals with and without NAFL. This approach revealed correlations between the different tissues. The correlation between liver and mesenteric adipose tissue transcriptomics particularly stands out (r=0.8), followed by liver transcriptomics and the fecal metagenome (r=0.67*; **Figure 3a***).

In addition, the full correlation matrix revealed the interrelation between metabolites, bacterial species and genes that can be used to generate biological hypotheses that can be used to further unravel the pathophysiology or develop next-generation therapeutic strategies for NAFL (***Figure 3b***). For example, N-acetyl-2-aminooctanoate, *Lactobacillus sakei,* hepatic *TRIP6* (Thyroid Hormone Receptor Interactor 6), *ERBB2* (Erb-B2 Receptor Tyrosine Kinase 2) and *MIR34AHG* (*MIR34A* Host Gene affiliated with the lncRNA class), were all upregulated in NAFL and correlated positively with each other (r≥0.6). Suggesting that this metabolite could be of bacterial origin, or that the circulating levels are influenced by the gut microbiome. Moreover, the correlation between this metabolite and *TRIP6* and *ERBB2,* two genes that were recently identified to play a role in the pathophysiology of NAFLD[38,39] suggests that upregulation of these genes can be induced via circulating metabolites. In addition, N-acetyl-2-aminooctanoate was positively correlated with *AADACL* (Arylacetamide Deacetylase Like 3) in mesenteric adipose tissue, which is a gene involved in lipolysis of adipose tissue and thus contributes to hepatic triglyceride accumulation[40]. To which extent these genes are regulated by bacterial strains or metabolites needs to be further investigated. 1-carboxyethylvaline was positively correlated with *ACAN* in subcutaneous adipose tissue. Furthermore, diacylglycerol was positively correlated with *WFDC1* and *ACAN* in subcutaneous adipose tissue. *WFDC1* and *ACAN* in the subcutaneous adipose tissue were highly enriched in gene sets involved in mitochondrial translation/elongation, suggesting a strong association between BCAAs and potential regulative signaling from adipose tissue.



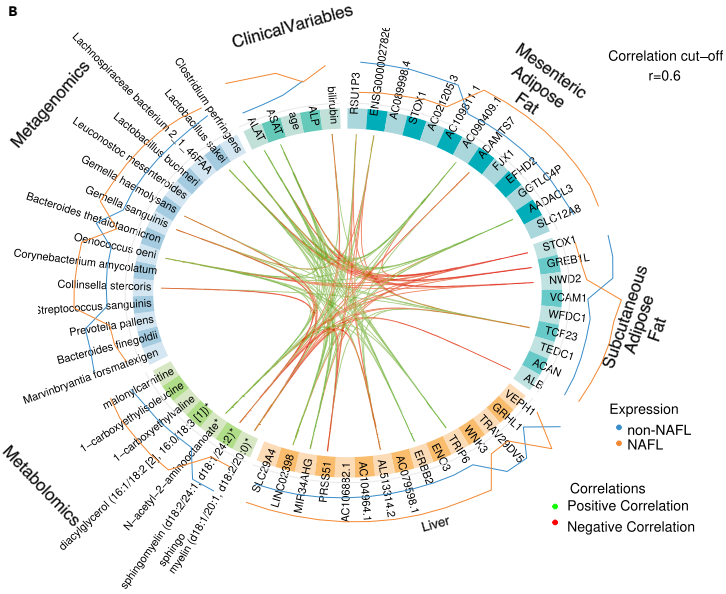


Figure 3. DIABLO analysis and correlations among multi-omics datasets for individuals with and without NAFL. (a) Total correlation matrix for all the different omic datasets after Sparce Principal Least Squares Regression with mixOmix DIABLO. Highest correlation is observed for genes from liver and mesenteric adipose tissue. **(b)** Circular correlation plot by Data Integration Analysis for Biomarker discovery using a Latent cOmponents (mixOmics DIABLO), for top contributing components to from each omics dataset (metabolites, genes, bacterial species). Correlation cut-off is r=0.6. Signature involves Prevotella species, branched-chain amino acid metabolites, sphingolipid metabolites, diacyglycerols, liver genes highly involved in cancer pathways, renin- angiotensin system, mesenteric adipose tissue genes involved in carbohydrate metabolism and subcutaneous adipose tissue genes involved in mitochondrial translation/elongation.

Finally, to quantify the robustness of the individual omics signatures obtained by the integrative analysis, the power of every chosen omics subset by DIABLO to predict NAFL was assessed **(Figure 4).** A series of Generalized Linear Models (GLMs) aimed to investigate whether the minimal discriminatory signal of the omics could outperform the clinical variables capacity to correctly predict NAFL. As anticipated, the performance of the signature found in the liver transcriptome was very high with an area under the curve (AUC) of 0.98, followed by the visceral adipose tissue and subcutaneous adipose tissue. The post prandial metabolites and gut microbial species signatures appear to be more accurate prognostic markers of NAFL when compared to the chosen clinical variables.

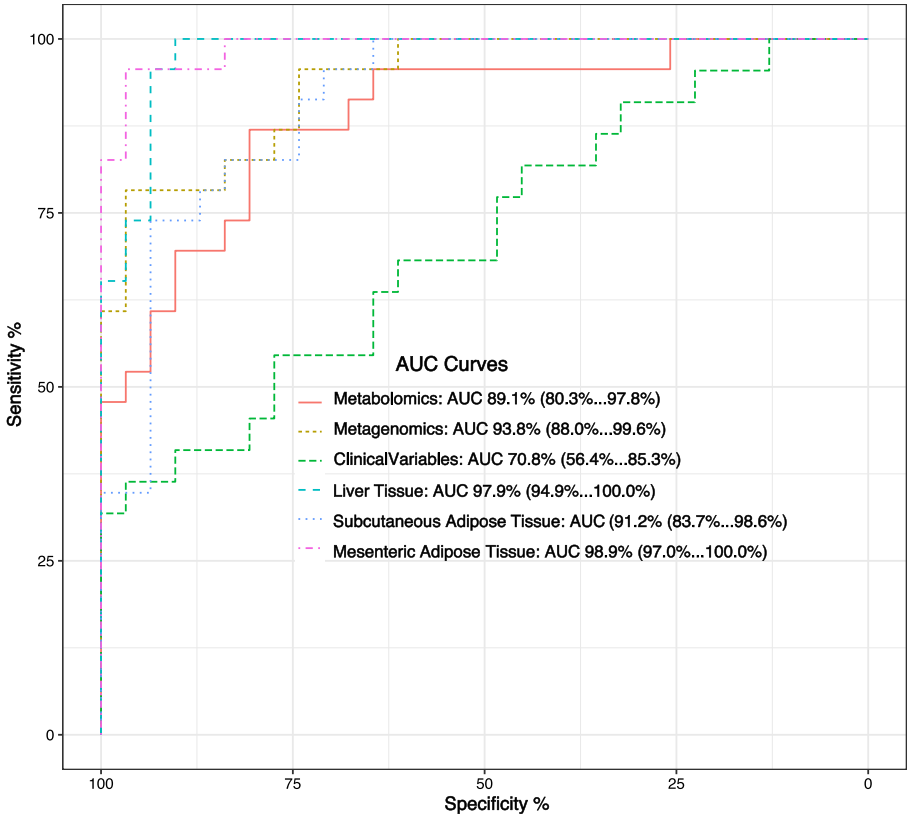


Figure 4. AUC predictive capacity for each omic dataset from DIABLO analysis. All the transcriptomics datasets and the chosen genes can very accurately predict NAFL. Both DIABLO chosen Metabolome and Metagenome datasets outperform the Clinical variables in NAFL predictive capacity, with AUC=89.1% and 93.8% respectively versus AUC=70.8%.

To differentiate if these signatures are driven by hepatic steatosis (the prerequisite for NAFLD diagnoses) or lobular inflammation, we performed the same analyses but then between women with (n=41) and without lobular inflammation (n=14) and observed no distinct differences between al the omics sets suggesting that these signatures are driven mainly by the steatosis component (data now shown). In summary, the computational framework used here for integrating various omics datasets successfully identified a highly correlated discriminatory signature for NAFL that included BCAA metabolites, diacylglycerol, liver genes involved in HIF-1 signaling, mesenteric adipose tissue genes involved in fat metabolism and subcutaneous adipose tissue genes that are part of mitochondrial translation/elongation.

Women with NAFL have a different response upon MMT after massive weight loss

To further substantiate that the alterations have clinical relevance, we analyzed the mixed meal data one year after bariatric surgery. Interestingly, at baseline we did not observe significant differences between individuals with and without NAFLD in glucose and insulin response during the MMT, but one year after bariatric surgery and massive weight loss, a clear difference was observed between the two groups in insulin but not in glucose during the MMT, which was significant (**Figure 5**). These results further suggests that whole body metabolism is indeed different in this early phase of the disease. Given the fact that weight loss and decrease in liver transaminases were not different, these data suggest that these differences are due to the inherent differences in whole body metabolism.

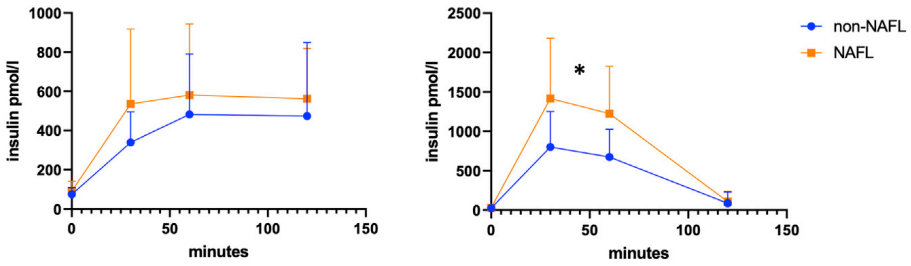


Figure 5. Insulin excursions during the mixed meal test in women with and without NAFL before (A) and one year after bariatric surgery (B).

DISCUSSION

Here we used a systems biology approach to identify factors that may contribute to NAFL development by analyzing six omics data sets of 55 women -who only differed in the presence of hepatic steatosis- including fecal metagenomics, plasma metabolomics and liver, subcutaneous and mesenteric adipose tissue transcriptomics. NAFLD is a multifactorial disease, which is underscored in the present study by showing that in each individual omics data set, differences between women with and without NAFL could be observed. Suggesting that whole body metabolism is already altered in this early stage of the disease. The alterations in gut microbial composition are in line with previous work conducted by other independent groups showing that in subjects with NAFL, the gut microbiome is dominated by members of Firmicutes[27,29]. However, our findings are in contrast with a recent report where liver steatosis was anti correlated with Firmicutes[25]. Nevertheless, it is plausible that there is not one unique microbiome signature for NAFLD, bearing in mind that the human microbiome is shaped by multiple factors such as age, sex and disease state[41]. On species level, we observed a decrease in *Prevotella* species in individuals with NAFL. Interestingly, most of the *Prevotella* species were of oral origin, which is in contrast to previous findings[42]. However, the mechanism and clinical significance underlying the increased transfer of oral bacteria to the gut remain to be elucidated. Subtle changes in the plasma metabolome were observed, especially in the post MMT samples, emphasizing that early changes in metabolism are more pronounced post meal than in fasting conditions. Alterations in BCAA composition in individuals with cardiometabolic disease is often explained to be caused by impaired amino acid metabolism linked to insulin resistance in the liver or muscle[43]. Since insulin and glucose levels did not differ during the MMT, this indicates that these changes are independent of insulin resistance and opens up the prospect that these changes have been derived from another origin, potentially the gut microbiome[33]. Diacylglycerol, which is associated with NAFLD[30], was increased in post MMT plasma of individuals with NAFL. Diacylglycerol is synthesized intracellularly from specific lipid precursors such phosphatidylcholines, possibly including the metabolite that was increased in fasting conditions in individuals with NAFL.

To further investigate the relation between alterations in microbial composition and metabolites in host metabolism, we analyzed the transcriptome of liver and two adipose tissue depots obtained during surgery. Pathways previously suggested to play a pivotal role in the development of NAFLD such as the HIF-1 signaling pathway in the liver, fat and glucose metabolism and inflammation in adipose tissue were increased in individual with NAFL. Nevertheless, the exact mechanisms that contribute to these pathways, especially in early disease, remain largely unknown. Therefore, we constructed a multivariable model to objectively quantify the crosstalk between these different omics datasets. We observed

strong correlations between omics data sets, especially between mesenteric adipose tissue and liver transcriptomic data (r=0.8) and between liver and subcutaneous adipose tissue (r=0.51). These observations are in line with the current concept that adipocyte dysfunction plays a pivotal role in the pathophysiology of NAFLD[36,44]. Adipose tissue expansion of both the subcutaneous and the visceral compartment leads to hypoxia-induced hypersecretion of adipocytokines such as Tumor Necrosis Factor (TNF) and interleukin (IL) 6 by the adipocytes as well as by the inflammatory immune cells that accumulate in adipose tissue of individuals with obesity[44,45]. When reaching the liver through the portal vein, these mediators, together with increased levels of lipid metabolites such as diacylglycerols observed during metabolic dysregulation, can contribute to the development and progression of NAFLD[36]. Interestingly, KEGG pathway enrichment of the differential significant genes of both mesenteric and subcutaneous adipose tissue revealed that pathways involved in fat and glucose metabolism and TNF signaling were upregulated in NAFL, respectively, underscoring the potential role of the adipose tissue in the development of NAFLD.

Recently, it was shown that there is a considerable link between the liver, the gut microbiome and gut microbial metabolites[25]. In this study, post prandial metabolomes and fecal metagenomes in general did not correlate with each other. A more in-depth view, however, revealed associations among metabolites belonging to amino acid metabolism, bacterial species and liver genes. For example, N-acetyl-2-aminooctanoate, *Lactobacillus sakei and* hepatic *TRIP6*, *ERBB2* and *MIR34A* were all upregulated in NAFL and correlated positively with each other (r≥0.6). Of interest, *TRIP6,* is an upstream activator of the transcriptional co-activators YAP (or YAP1) and TAZ and are involved in the pathogenesis of NAFLD[38,39]. Also, genes involved in the hippo-signaling pathway were associated with this metabolite and bacterial strains. Hippo-signaling and downstream effectors are involved in a multitude of cell and non-cell autonomous functions including metabolism, cell proliferation and survival[38]. Interestingly, in the total correlation matrix, non-coding RNAs (*LINC02398* and *MIR34AHG*) and clone (*AC106882.1; AC109811.1*) of liver and mesenteric adipose fat were included. To what extent these non-coding and clones are associated with transcriptional regulation and are involved in the pathogenesis of NAFL remain to be investigated. Although our results are of associative nature, DIABLO full matrix correlation highlights the interrelation between metabolites, bacterial species and genes and can be used to generate hypothesis to further study the pathophysiology of NAFL in humans.

In conclusion, our study provides a comprehensive multi-omics analysis of women with NAFL, providing a different strategy to study the pathophysiology of NAFL in women. Even though it is increasingly recognized that NAFL, also referred as “simple steatosis”, is more than just the passive accumulation of excessive fat, we further emphasize this by showing differences in metabolites, genes and gut microbial species between individuals with and

without NAFL. This is important work considering the fact that women with NAFLD have a higher change of CVD events, mortality and disease progressions, even in the absence of severe hepatic inflammation and scarring[4]. To what extent these findings are related to the severe outcome in women remain to be investigated. Lasty, by building a multivariate model, we revealed that there is substantial crosstalk between these different omics sets. Our model suggests that in early stages of the disease, adipocyte dysfunction is the predominant factor in disease development followed by gut microbial composition and plasma metabolites.

In conclusion, our study provides a comprehensive multi-omics analysis of women with NAFL, providing a different strategy to study the pathophysiology of NAFL in women. Even though it is increasingly recognized that NAFL, also referred as "simple steatosis", is more than just the passive accumulation of excessive fat, we further emphasize this by showing differences in metabolites, genes and gut microbial species between individuals with and without NAFL. This is important work considering the fact that women with NAFLD have a higher change of CVD events, mortality and disease progressions, even in the absence of severe hepatic inflammation and scarring[4]. To what extent these findings are related to the severe outcome in women remain to be investigated. Lasty, by building a multivariate model, we revealed that there is substantial crosstalk between these different omics sets. Our model suggests that in early stages of the disease, adipocyte dysfunction is the predominant factor in disease development followed by gut microbial composition and plasma metabolites.

Limitations of the study

We note that the analyses of human omics data sets in our study has some limitations. Here, we used tissue and plasma samples obtained from women who underwent bariatric surgery, which may introduce relevant biases in particular pre-operative weight-loss with a subsequent decrease in liver volume. However, individuals who had lost more than 3% of weight in the month prior to surgery or more than 5% six months before surgery were excluded. We therefore ensure that the samples were obtained in a relatively stable period. The relatively low number of individuals (n=55) in this study could potentially introduce bias to this particular modelling approach, especially since we did not have a validation cohort available to confirm these signatures. Therefore, external validation of these metabolites is warranted or should be further evaluated. Nevertheless, it is considerably challenging to come across similar multi-omics datasets in an external cohort, postprandial metabolome in particular, that include the same metabolites. Another limitation is that with the current study design we were not able to investigate to what extent the robust NAFL signature in each omics set contribute to the increased risk of developing CVD or adverse clinical outcome.

Acknowledgements

The BARIA study is funded by the Novo Nordisk Foundation (NNF15OC0016798). The Novo Nordisk Foundation Center for Basic Metabolic Research is supported by an unconditional grant (NNF10CC1016515) from the Novo Nordisk Foundation to University of Copenhagen. The BARIA study is a Scandinavian-Dutch collaboration. Funding from Knut and Alice Wallenberg Foundation is also acknowledged. The computations and RNA Sequencing were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at C3SE (SNIC Computational Center of Chalmers University of Technology) partially funded by the Swedish Research Council through grant agreement no. 2018-05973. We Thank Torsten Scheithauer for making the figures.

Contributions

T.W.S., M.N., J.N. and F.B. supervised this work, - A.S.M & D.L. conducted conceptual design, data curation, data analysis, visualization and main manuscript preparation, - A.S.M., O.A., S.B., M.L.D.B, collected medical data and biopsies, - J.V. and D.L.P.P reviewed all liver biopsies, - V.T., A.L. performed DNA, RNA and metabolomics isolations and optimizations, - L.M.O. conducted metagenomics data analysis, L.E.O., F.B. and J.N. coordinated project administration, - D.L., A.S.M., H.H., K.K., L.M.O., Ö.A., Y.I.Z.A., M.L.D.B., V.T., L.E.O., A.L., S.H., V.E.A.G., A.K.G., T.W.S., M.N., F.B. and J.N. conducted hypothesis generation, manuscript review and editing.

Competing Interests

The authors declare no conflict of interest.

REFERENCES

1. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
2. Farrell, G. C. & Larter, C. Z. Nonalcoholic fatty liver disease: From steatosis to cirrhosis. *Hepatology* **43**, S99–S112 (2006).
3. Simon, T. G., Roelstraete, B., Khalili, H., Hagström, H. & Ludvigsson, J. F. Mortality in biopsy-confirmed nonalcoholic fatty liver disease: Results from a nationwide cohort. *Gut* (2020) doi:10.1136/gutjnl-2020-322786.
4. Allen, A. M. *et al.* Women with nonalcoholic fatty liver disease lose protection against cardiovascular disease: A longitudinal cohort study. *American Journal of Gastroenterology* (2019) doi:10.14309/ajg.0000000000000401.
5. Balakrishnan, M. *et al.* Women Have a Lower Risk of Nonalcoholic Fatty Liver Disease but a Higher Risk of Progression vs Men: A Systematic Review and Meta-analysis. *Clinical Gastroenterology and Hepatology* Preprint at https://doi.org/10.1016/j.cgh.2020.04.067 (2021).
6. Buzzetti, E., Pinzani, M. & Tsochatzis, E. A. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism* (2016) doi:10.1016/j.metabol.2015.12.012.
7. Marjot, T., Moolla, A., Cobbold, J. F., Hodson, L. & Tomlinson, J. W. Nonalcoholic Fatty Liver Disease in Adults: Current Concepts in Etiology, Outcomes, and Management. *Endocr Rev* (2020) doi:10.1210/endrev/bnz009.
8. Aron-Wisnewsky, J. *et al.* Gut microbiota and human NAFLD: disentangling microbial signatures from metabolic disorders. *Nat Rev Gastroenterol Hepatol* **17**, 279–297 (2020).
9. Bedossa, P. *et al.* Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. *Hepatology* **60**, 565–575 (2014).
10. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).
11. van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *J Intern Med* joim.13157 (2020) doi:10.1111/joim.13157.
12. Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276–278 (2020).
13. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* **35**, 1069–1076 (2017).
14. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
15. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
16. Ignatiadis, N., Klaus, B., Zaugg, J. B. & Huber, W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat Methods* **13**, 577–580 (2016).
17. Chen, E. Y. *et al.* Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 1–14 (2013).
18. Kanehisa, M. & Goto, S. KEGG : Kyoto Encyclopedia of Genes and Genomes. **28**, 27–30 (2000).
19. Chamberlain, S. A. & Szöcs, E. taxize : taxonomic search and retrieval in R [v2 ; ref status : indexed , http://f1000r.es/24v]. **191**, 1–28 (2014).
20. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).
21. Tenenhaus, A. *et al.* Variable selection for generalized canonical correlation analysis. *Biostatistics* **15**, 569–583 (2014).
22. Karlsson, F. H., Nookaew, I. & Nielsen, J. Metagenomic Data Utilization and Analysis (MEDUSA) and Construction of a Global Gut Microbial Gene Catalogue. *PLoS Comput Biol* **10**, (2014).
23. Lonardo, A. *et al.* Sex Differences in Nonalcoholic Fatty Liver Disease: State of the Art and Identification of Research Gaps. *Hepatology* Preprint at https://doi.org/10.1002/hep.30626 (2019).

24. Vandel, J. *et al.* Hepatic molecular signatures highlight the sexual dimorphism of Non-Alcoholic SteatoHepatitis (NASH). *Hepatology* (2020) doi:10.1002/hep.31312.

25. Hoyles, L. *et al.* Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat Med* **24**, 1070–1080 (2018).

26. Leung, C., Rivera, L., Furness, J. B. & Angus, P. W. The role of the gut microbiota in NAFLD. *Nat Rev Gastroenterol Hepatol* **13**, 412–425 (2016).

27. Boursier, J. *et al.* The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. *Hepatology* (2016) doi:10.1002/hep.28356.

28. Caussy, C. *et al.* A gut microbiome signature for cirrhosis due to nonalcoholic fatty liver disease. *Nat Commun* **10**, 1406 (2019).

29. Loomba, R. *et al.* Gut Microbiome-Based Metagenomic Signature for Non-invasive Detection of Advanced Fibrosis in Human Nonalcoholic Fatty Liver Disease. *Cell Metab* **25**, 1054-1062.e5 (2017).

30. Samuel, V. T. & Shulman, G. I. Nonalcoholic Fatty Liver Disease as a Nexus of Metabolic and Hepatic Diseases. *Cell Metab* **27**, 22–41 (2018).

31. Kalhan, S. C. *et al.* Plasma metabolomic profile in nonalcoholic fatty liver disease. *Metabolism* (2011) doi:10.1016/j.metabol.2010.03.006.

32. Gaggini, M. *et al.* Altered amino acid concentrations in NAFLD: Impact of obesity and insulin resistance. *Hepatology* (2018) doi:10.1002/hep.29465.

33. Krautkramer, K. A., Fan, J. & Bäckhed, F. Gut microbial metabolites as multi-kingdom intermediates. *Nature Reviews Microbiology 2020 19:2* **19**, 77–94 (2020).

34. Semenza, G. L. Life with Oxygen. *Science (1979)* **318**, 62–64 (2007).

35. Azzu, V., Vacca, M., Virtue, S., Allison, M. & Vidal-Puig, A. Adipose Tissue-Liver Cross Talk in the Control of Whole-Body Metabolism: Implications in Nonalcoholic Fatty Liver Disease. *Gastroenterology* **158**, 1899–1912 (2020).

36. Arab, J. P., Arrese, M. & Trauner, M. Recent Insights into the Pathogenesis of Nonalcoholic Fatty Liver Disease. *Annu Rev Pathol* **13**, 321–350 (2018).

37. Singh, A. *et al.* DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).

38. Machado, M. V. *et al.* Accumulation of duct cells with activated YAP parallels fibrosis progression in non-alcoholic fatty liver disease. *J Hepatol* **63**, 962–970 (2015).

39. Wang, X. *et al.* Hepatocyte TAZ/WWTR1 Promotes Inflammation and Fibrosis in Nonalcoholic Steatohepatitis. *Cell Metab* **24**, 848–862 (2016).

40. Quiroga, A. D. & Lehner, R. Pharmacological intervention of liver triacylglycerol lipolysis: The good, the bad and the ugly. *Biochem Pharmacol* **155**, 233–241 (2018).

41. Meijnikman, A. S., Gerdes, V. E., Nieuwdorp, M. & Herrema, H. Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. *Endocr Rev* **39**, 133–153 (2018).

42. Atarashi, K. *et al.* Ectopic colonization of oral bacteria in the intestine drives T H 1 cell induction and inflammation. *Science* **358**, 359–365 (2017).

43. White, P. J. & Newgard, C. B. Branched-chain amino acids in disease. *Science* **363**, 582–583 (2019).

44. du Plessis, J. *et al.* Association of Adipose Tissue Inflammation With Histologic Severity of Nonalcoholic Fatty Liver Disease. *Gastroenterology* **149**, 635e14-648.e14 (2015).

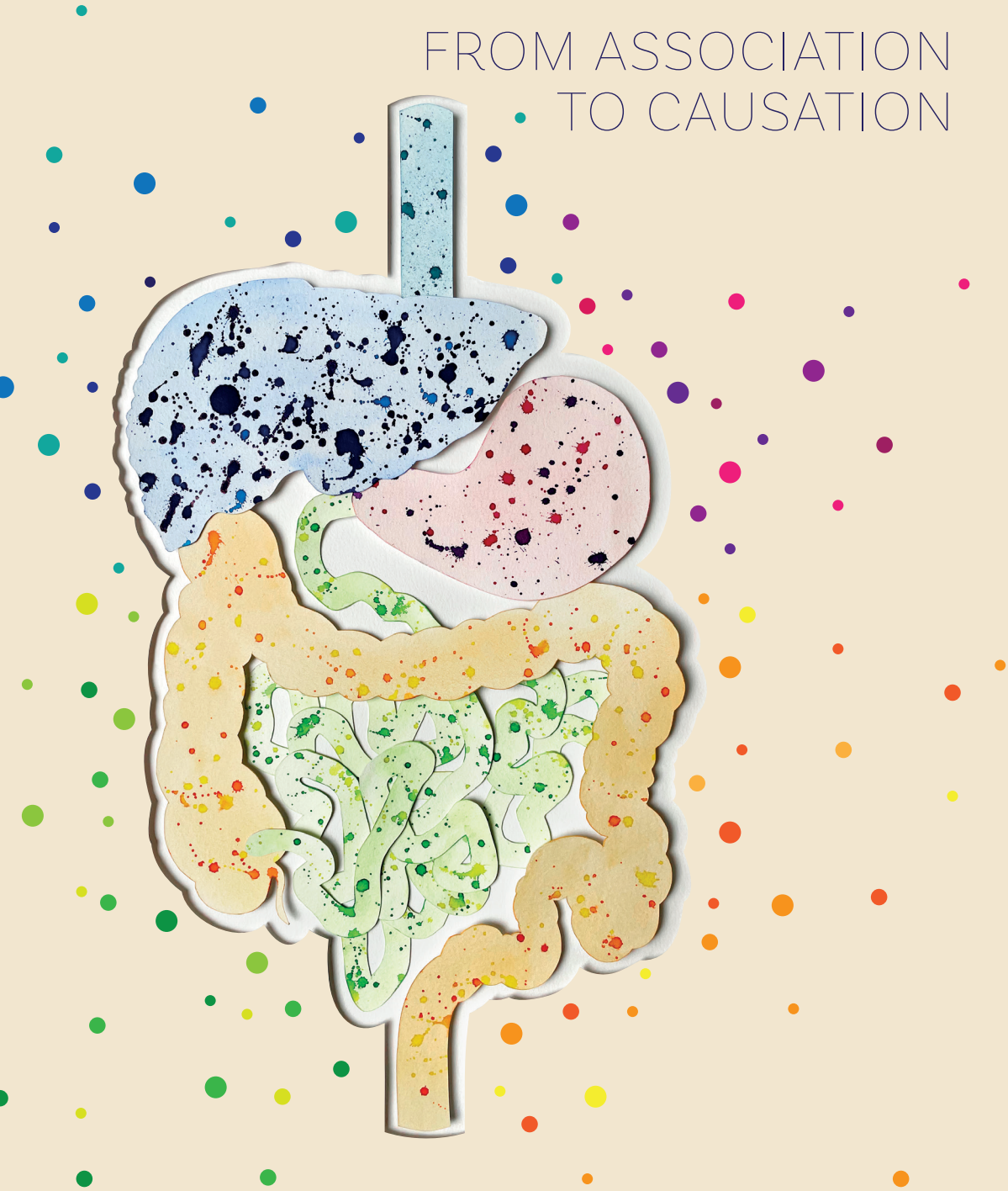
45. Fernandez-Real, J. M. *et al.* Circulating interleukin 6 levels, blood pressure, and insulin sensitivity in apparently healthy men and women. *J Clin Endocrinol Metab* **86**, 1154–1159 (2001).

Supplementary information available at:

A systems biology approach to study non-alcoholic fatty liver (NAFL) in women with obesity: iScience (cell.com)

PART 2

FROM ASSOCIATION TO CAUSATION



8

Microbiome-derived ethanol in non-alcoholic fatty liver disease

Abraham S. Meijnikman, Mark Davids, Hilde Herrema, Omrum Aydin, Valentina Tremaroli, Melany Rios-Morales, Han Levels, Sjoerd Bruin, Maurits de Brauw, Joanne Verheij, Marleen Kemper, Adriaan G. Holleboom, Maarten E.M. Tushuizen, Thue W. Schwartz, Jens Nielsen, Dees Brandjes, Eveline Dirinck, Jonas Weyler, An Verrijken, Christophe De Block, Luisa Vonghia, Sven Francque, Ulrich Beuers, Victor E. A. Gerdes, Fredrik Bäckhed, Albert K. Groen, Max Nieuwdorp

Nature Medicine, 2022 Oct;28(10):2100-2106

ABSTRACT

To test the hypothesis that the gut microbiota of individuals with non-alcoholic fatty liver disease (NAFLD) produce enough ethanol to be a driving force in the development and progression of this complex disease, we performed one prospective clinical study and one intervention study. Ethanol was measured in fasting and 120-minutes post mixed meal test (MMT) in 146 individuals. In a subset of 37 individuals' and in an external validation cohort, ethanol was measured in portal vein blood. In an intervention study, 10 individuals with NAFLD and 10 overweight but otherwise healthy controls were infused with the selective alcohol dehydrogenase inhibitor (ADH) before an MMT. Median portal vein ethanol concentrations increased with disease progression: 2.1mM; NAFL 8.0mM; NASH 21.0mM and were 187 (IQR:17-516) times higher compared to fasted peripheral blood. Inhibition of ADH induced a 15-fold (IQR:1.6-20) increase in peripheral blood ethanol concentration in individuals with NAFLD, though this effect was abolished after antibiotic treatment. Specifically, *Lactobacillaceae* correlated with post prandial peripheral ethanol concentrations (spearman rho:0.42; $p<10^{-5}$) in the prospective study. Our data shows that first pass effect obscures the levels of endogenous ethanol production and suggest that microbial ethanol could be considered in the pathogenesis of this highly prevalent liver disease.

INTRODUCTION

The gut microbiome has the capacity to produce clinically relevant amounts of ethanol, which might contribute to development of non-alcoholic fatty liver disease (NAFLD)[1,2]. NAFLD is the most frequent global chronic liver disease affecting one in four individuals in the general population[3] and comprises a spectrum of clinical and histopathological abnormalities[4]. It is well-known that individuals with NAFLD and alcoholic fatty liver disease (ALD) share histologic features including hepatic steatosis and the large number and size of Mallory bodies, suggesting common pathophysiology[5–7].

High levels of circulating microbial ethanol have been thought to result from functional impairment of hepatic insulin-dependent alcohol dehydrogenase (ADH)[8]. Overt microbial ethanol production in NAFLD *per se* has, however, also been postulated[1,2,9–11]. If produced chronically in relevant amounts, microbial produced ethanol could have the ability to alter lipid and glucose metabolism and induce steatosis and inflammation in the liver[12]. Yet, reported peripheral blood concentrations of microbial-derived ethanol are generally very low, which questions its clinical relevance in NAFLD pathogenesis. The liver has a massive capacity to metabolize ethanol via the alcohol dehydrogenase and cytochrome P4502E1 pathways (~3.5mg/kg/day)[13]. This likely results in a significant first-pass effect, rendering low circulating levels of ethanol and explaining the lack of systemic alcohol misuse symptoms[14]. However, there are exceptions, such as patients with auto brewery syndrome or end stage liver disease, where microbial ethanol production exceeds the liver's capacity to clear ethanol from the portal circulation[2]. Besides these exceptions, insight in the production of gut microbial ethanol is stagnated by the fact that portal vein blood, which is enriched in microbial metabolites and has not been subjected to this first-pass effect, is difficult to obtain in humans. The putative role of microbial ethanol in human disease has therefore never been demonstrated unequivocally. We here overcome methodological challenges and report on gut microbial ethanol production in individuals with and without NAFLD.

METHODS

Study design

We designed and performed one prospective clinical study and one intervention study, which were in accordance with the Declaration of Helsinki and were approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC (Trialregister BARIA study NL8983, ETHANASH study NL7693). All participants provided written informed consent. In addition, we used samples from an external observational study as validation cohort,

which was collected in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Antwerp University Hospital (Belgian registration number B300201524515) and requiring written informed consent of the patient.

In the prospective study, we enrolled 146 individuals (Table 1) from our bariatric surgery cohort[15]. All individuals underwent an extensive metabolic work-up prior to their bariatric surgery procedure and did not lose more than three to five percent of body weight three or six months prior to surgery. Within two months before surgery, a two-hour mixed meal test was performed to calculate glucose excursions, deduce insulin resistance and measure post prandial circulating microbial metabolites. Excessive ethanol consumption (>14 units/week) was an exclusion criterion and was assessed using an ethanol timeline followback assessment during the screening and mixed meal test. Overnight fasted participants ingested two Nutridrink compact drinks at t=0 (125ml each, Nutricia®). Caloric intake totaled to 23.3 grams of fat, 74.3 grams of carbohydrates (of which 38.5 grams sugar) and 24.0 grams of protein. Ethanol was measured fasting and at t=120 minutes. In a subset of 37 individuals, samples from portal vein blood were drawn on the day of surgery. Fecal samples were obtained on the day of surgery and processed as previously described[16]. All samples were stored at -80°C until further processing.

During surgery, small intestinal biopsies were obtained from the jejunum. Wedge liver biopsies were obtained from the diaphragmatic surface of segment three or five of the liver. Biopsies were snap-frozen in liquid nitrogen and stored at -80°C until further processing and paraffin-embedded for histology. Paraffin embedded liver sections were stained by Haematoxylin-Eosin and Sirius Red. Members of the Dutch Liver Pathology Panel, trained for NAFLD scoring according to the Steatosis, Activity and Fibrosis (SAF) score[17] reviewed the sections. Borderline cases were discussed during pathology consensus meetings. NAFLD was categorized into NAFL when steatosis was solely present (in >5% of hepatocytes) or when in concert with mild inflammation without ballooning. NAFLD was categorized into NASH when steatosis was present (in >5% of hepatocytes) in concert with inflammation and ballooning. The activity part of the SAF score (SAF-A) that incorporates the scores for hepatocellular ballooning and lobular inflammation was also assessed. In addition, the Nonalcoholic Fatty Liver Disease Activity Score (NAS), which is the sum of the scores for steatosis [range, 0 to 3], ballooning [range, 0 to 2], and lobular inflammation [range, 0 to 3], with higher scores indicating greater disease activity) was assessed[18]. The external validation cohort consisted of 51 individuals of whom both portal vein blood and liver biopsies obtained during bariatric were available. Liver biopsies were subjected to the same stains and were scored according to the same criteria[17].

In the intervention study, we enrolled ten individuals with biopsy-proven non-cirrhotic NASH according to the SAF score[19] from our outpatient visit clinic. Ten overweight but

otherwise healthy individuals were recruited via advertisement and were screened for cardiometabolic diseases including hypertension, type 2 diabetes and hepatic steatosis using an ultrasonography of the liver prior to inclusion. Ethanol consumption of more than two units was an exclusion criterium that was assessed similarly as in the prospective trial. At baseline, all participants underwent a four-hour mixed meal test consisting of 200 g of carbohydrates. This meal differed from the one given in the prospective trial where a lower amount of carbohydrates was given ensuring that the meal could be repeated during the follow-up (i.e., after bariatric surgery) without inducing severe dumping syndrome. The meal in the intervention study was a better reflection of the oral challenges routinely encountered daily. Blood was drawn at t=0, 30, 60, 90, 120, 180 and 240 minutes. Within two weeks after the baseline mixed meal test, individuals underwent a second mixed meal test, which was preceded by an infusion with the selective alcohol dehydrogenase inhibitor 4-methylpyrazole. The first-pass effect of the liver for ethanol is thereby blocked. The total amount of 4-methylpyrazole (15mg/kg) was infused within one hour immediately followed by a mixed meal test as described above. Nine out of ten individuals with NASH were then given a one-week oral antibiotics course (metronidazole, 500 mg two times daily; clindamycin, 300 mg three times daily and ciprofloxacin, 500 mg once daily) to deplete the gut microbiome. After the antibiotics course, these individuals underwent a third mixed meal test preceded by an infusion with 4-methylpyrazole as described for the second mixed meal test. Fecal samples were obtained at every site visit. All mixed meal test were conducted under permanent medical supervision.

Ethanol measurement

Plasma was deproteinized using perchloric acid (5% v/v) and after centrifugation (5 sec at 12000g) neutralized with NaOH (5M). Subsequently ethanol was assayed using an enzyme-based kit from DiaSys (Holzheim, Germany). The presence of ethanol was validated by measuring ethanol on HPLC. Separation of ethanol was carried out using a Reprogel H column (250 mm × 4.6 mm, 9 µm. Screening Devices, Amersfoort, Netherlands) using Jasco pump (PU4285, Jasco Benelux, De Meern, Netherlands) in an isocratic setting with demineralized water as mobile phase at a flowrate of 0.6 ml/min. After separation ethanol was detected using a refractive index detector at a detection-cell temperature of 30° C sample (RI 2031, Jasco Benelux, De Meern, Netherlands). During measurements all samples were kept at 4 °C in closed vials in the Jasco autosampler (AS4285, Jasco Benelux, De Meern, Netherlands). Finally, the absolute ethanol concentrations were calculated using Chrom-Nav chromatography software (Version 2.0, Jasco, de Meern, Netherlands).

Fecal total genomic DNA isolation and microbiome analyses

Total fecal genomic DNA was extracted from 100 mg of feces using a modification of the IHMS DNA extraction protocol Q[20]. Briefly, fecal samples were extracted in Lysing Matrix E tubes (MP Biomedicals) containing ASL buffer (Qiagen, Venlo, The Netherlands). Lysis was obtained after homogenization by vortexing for 2 minutes, by two cycles of heating at 90°C for 10 minutes followed by three bursts of bead beating at 5.5 m/sec for 60 seconds in a FastPrep®-24 Instrument (MP Biomedicals). After each bead-beating burst, samples were laced on ice for 5 minutes. The supernatants containing fecal DNA were collected after the two cycles by centrifugation at 4°C. Supernatants from the two centrifugations steps were pooled and a 600 µL aliquot from each sample was purified using the QIAamp DNA Mini kit QIAGEN, Venlo, The Netherlands) in the QIAcube (QIAGEN Venlo, The Netherlands) instrument using the procedure for human DNA analysis. Samples were eluted in 200 µL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0). Libraries for shotgun metagenomic sequencing were prepared by a PCR-free method; library preparation and sequencing were performed at Novogene (UK, Cambridge) on an HiSeq instrument (Illumina Inc. San Diego, CA, USA) with 150 bp paired-end reads and 6G data/sample.

The MEDUSA[21] pipeline was used for pre-processing of raw shotgun metagenomics sequence data. MEDUSA is an integrated pipeline for analysis of short metagenomic reads, which maps reads to reference databases, combines output from several sequencing runs and manipulates tables of read counts. The input number of total reads from the metagenome analysis were on average 23.4±2.2 million reads per sample and the total aligned reads 16.6±1.8 million reads per sample. The sequencing runs had high quality with almost 98% of the reads passing the quality cut-off. Out of the high-quality reads, on average 0.04% aligned to the human genome, although the data had been cleaned for human reads. Out of the high quality non-human reads, 78.4% aligned to the MEDUSA's software gene catalogue. Quality filtered reads were mapped to a genome catalogue and gene catalogue using Bowtie2[22]. The taxon ids were input to taxize package[23], so as to get full taxonomic information and ranking for the species. Fungal composition was determined using the human mycobiome scan[24]. KEGG othologs K00001, K00121, K04072, K11440, K13951, K13953, K13954, and K00132, K04072, K04073 were taken to assess the metagenomic potential for alcohol dehydrogenase (EC:1.1.1.1) and acetaldehyde dehydrogenase (EC:1.2.1.10).

Small intestinal total genomic DNA isolation and microbiome analyses

Small intestinal biopsies were expected to be rich in host DNA and therefore a bacterial 16S targeted method was applied. Small intestinal biopsies were lysed using repeated bead beating in STAR buffer (Roche Diagnostics) following Proteinase K treatment. Total genomic DNA was isolated from the lysates using a Maxwell device (RSC Blood kit,

Qiagen). Bacterial 16S rRNA was amplified using the V3-V4 341F forward primer and the 805R reverse primer. The PCR was performed in a total volume of 30 µl containing 1× High Fidelity buffer (Thermo Fisher Scientific, Waltham, MA, USA); 1 µl deoxynucleoside triphosphate (dNTP) mix (10 mM; Promega, Leiden, The Netherlands); 1 U of Phusion green high-fidelity DNA polymerase (Thermo Fisher Scientific, Waltham, MA, USA); 500 nM the forward 8-nucleotide (nt) sample-specific barcode primer containing the Illumina adapter, pad, and link (341F [5′-CCTACGGGNGGCWGCAG-3′]); 500 nM the reverse 8-nt sample-specific barcode primer containing the Illumina adapter, pad, and link (805R [5′-GACTACHVGGGTATCTAATCC-3′]); 100 ng/µl of template DNA; and nuclease-free water. The amplification program was as follows: an initial denaturation step at 98 °C for 30 s; 30 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 20 s, and elongation at 72 °C for 90 s; and an extension step at 72 °C for 10 min. PCR products (~540 bp) were checked on gel (1% wt/vol agarose, containing ethidium bromide, AppliChem GmbH, Darmstadt, Germany). Bacterial PCR products were purified using AMPure XP beads (Beckman Coulter, Brea, CA, USA). Amplicon DNA concentration was measured using Qubit (Thermo Fisher Scientific, Waltham, MA, USA) and DNA quality was addressed on a Agilent Bioanalyzer. The purified products were equimolarly pooled and libraries were sequenced on an Illumina MiSeq platform (paired-end run, 251 cycles, GATCBiotech, Constance, Germany) using V3 chemistry. Forward and reverse reads were truncated to 240 and 210 bases, respectively, and merged[25]. Merged reads that did not pass the Illumina chastity filter, had an expected error rate higher than 2, or were shorter than 380 bases were filtered. Amplicon sequencing variants (ASVs) were inferred for each sample individually with a minimum abundance of 4 reads with UNOISE3[26]. Unfiltered reads were then mapped against the collective ASV set to determine the abundances. Bacterial taxonomy was assigned using the RDP classifier[27] and SILVA 16S ribosomal RNA gene database V132[28]. Dataset were rarified down to 500 reads before further analysis.

RNA extraction and transcriptomics analyses

RNA was extracted from liver biopsies using TriPure Isolation Reagent (Roche, Basel, Switzerland) and Lysing Matrix D, 2 mL tubes (MP Biomedical, Irvine, CA, USAs) in a FastPrep®-24 Instrument (MP Biomedical, Irvine, CA, USAs) with homogenization for 20 seconds at 4.0 m/sec, with repeated bursts until no tissue was visible. Homogenates were kept on ice for 5 minutes between homogenization bursts if multiple cycles were needed. RNA was purified with chloroform (Merck, Darmstadt, Germany) in phase lock gel tubes (5PRIME) with centrifugations at 4 °C, and further purified and concentrated using the RNeasy MinElute kit (Qiagen, Venlo, The Netherlands). The quality of RNA was analysed on a BioAnalyzer instrument (Agilent), with quantification on Nanodrop (Thermo Fisher Scientific, Waltham,

MA, USA). Due to degradation of the RNA, libraries for RNAseq sequencing were prepared by rRNA depletion; library preparation and sequencing were performed at Novogene (Nanjing, China) on an HiSeq instrument (Illumina Inc., San Diego, CA, USA) with 150 bp paired-end reads and 10G data/sample. The average read count per sample was 42 ± 15 million. Obtained reads were quality trimmed using Trimmomatic[29] with sliding_window_threshold_q_score of 15. Quality trimmed reads of at least 36 nt length were pseudo aligned against *Homo sapiens* GRCh38 transcriptome release 97 using kallisto[30] (0.46.1). One sample was excluded from the analysis as it did not conform to the expression profile of the liver.

Statistical analysis

All statistical analyses were performed in R version 4.0.5[31]. Differences in patient characteristics were tested using Kruskal-Wallis Rank Sum Test (two sided) or Pearson's Chi-squared Test where appropriate. Differences in ethanol concentrations were assessed using Wilcoxon test and correlations were tested using Spearman's rank correlation coefficient. The post prandial ethanol increase was tested using a linear mixed model. Differential liver gene expression analysis for individuals with and without NAFL, NAFL and NASH was performed with DESeq2[32] package (1.30.1); log normalization was based on gene counts geometric distribution. The statistical analysis method for calculating differential expression rates is the Wald test. After False discovery rate (FDR) correction for multiple hypothesis testing. For these differentially significant co-expressed genes we performed gene enrichment with Enrichr platform[41] using GO terms (2021), KEGG human pathways[42] (2021) and WikiPathways Human (2019). Statistical analysis for the fecal microbiome was performed on rarefied count, (20 M reads per sample). Associations with microbiome alpha diversity metrics were determine using Wilcoxon test while beta diversity associations were tested using PERMANOVA as implemented in the vegan[33] package. Effects of microbiome-altering medication on ethanol levels were tested using linear models. To test for differential abundant taxa DESeq2[32] packages were used while spearman correlations were calculated for ethanol concentrations. Similar to the BARIA transcriptomics counts, log normalization based on gene counts geometric distribution has been conducted with it. Statistical analysis method for calculating differential expression rates was Wald Test. Obtained p-values were corrected per comparison using the Benjamini-Hochberg procedure[34].

Data availability

Fecal metagenomics and liver transcriptomics is deposited in the European Nucleotide Archive (ENA PRJEB47902) and European Genome-Phenome Archive (EGAS00001005704) respectively. All relevant clinical data is uploaded in the source data file. A comprehensive data analysis report can be found at https://amcmc.github.io/BARIA_ETHANASH/

RESULTS

Study overview and baseline characteristics

We designed and performed one prospective clinical study and one intervention study. In addition, we used samples from an external observational study as validation cohort. In the design of the studies, we did not account for differences in sex due to the nature of the studies performed. In the prospective study, we enrolled 146 individuals mean age 45, BMI 39±4 kg/m^2, female/male 108/38 (**Table 1**) from our bariatric surgery cohort[15]. In a subset of the prospective study, 37 individuals, samples from portal blood were drawn on the day of surgery **(Supplementary Table S1)**. The external validation cohort consisted of 51 individuals of whom both portal vein blood and liver biopsies obtained during bariatric surgery were available **(Supplementary Table S2)**. In the intervention study, we enrolled ten individuals with biopsy-proven non-cirrhotic NASH and ten overweight but otherwise healthy individuals **(Table 2)**. In both cohorts as well as in the intervention study, NAFLD was assessed according to the Steatosis, Activity and Fibrosis (SAF) score[17].

Portal and peripheral vein ethanol concentrations

Gut microbial-produced ethanol was detected in portal vein plasma in the subset of the 37 individuals from the prospective cohort in three biopsy-proven NAFLD classifications (14 no steatosis; 18 NAFL; 5 NASH. Portal vein ethanol levels were higher in individuals with NAFLD as compared to individuals without steatosis (medians: no steatosis 2.1mM; NAFL 8.0 mM; NASH 21.0mM (**Figure 1A).** Assuming a minimal portal flow of 0.3L/min, the median hepatic load of ethanol in individuals with NAFL and NASH would amount to 92 and 241 mg per min, respectively. To validate these findings, we repeated portal vein ethanol measurements in 51 individuals (22 no steatosis; 12 NAFL; 17 NASH) in the Antwerp bariatric surgery cohort. Indeed, portal vein microbial ethanol concentrations in the NASH group were in a comparably high range though ethanol concentrations were not increased in the NAFL group (no steatosis 1.7 mM; NAFL 1.7 mM; NASH 11.6mM, **Figure 1B**). The differences in results between cohorts are mainly driven by the association between ethanol and steatosis grades (**Extended Figure 1).**

We hypothesized that ethanol levels in the fasted state could be an underestimate and levels would spike a few hours postprandially. In a larger cohort (146 individuals: 58 no steatosis; 73 NAFL; 15 NASH, **Table 1**), we linked peripheral fasting and post prandial (*i.e.,* after a mixed meal test) ethanol concentrations to NAFLD classifications, histological scores, hepatic transcriptomics and gut microbiome characteristics.

Patient Characteristics Prospective Cohort	No Steatosis (N=58)	NAFL (N=73)	NASH (N=15)	P-value
Demographic				
Age (years)	43.5 (38.2-50)	50 (42-55)	48 (45-54)	**0,0153**
Female	51/58	48/73	9/15	**0,0069**
Anthropometric				
BMI (kg/m2)	39 (36.5-40.7)	38.7 (36.1-40.7)	39.1 (37.6-40.3)	0,9200
Type 2 diabetes mellitus (n)	7/58	24/73	5/15	**0,0165**
Clinical lab values (normal range)				
ALP (30-135 U/L)	82 (66-97)	82 (67.5-100.2)	70 (66.5-79)	0,3110
g-GT (10 - 40 IU/l)	22 (18-30)	29.5 (22.8-41.2)	20 (17.5-36.5)	**0,0108**
ALT (0 - 50 IU/l)	25.5 (19-36)	30 (22-42)	37 (26-42.5)	**0,0158**
AST (0 - 35 IU/l)	23 (20-25)	23.5 (20.2-28)	27 (26-32.5)	**0,0330**
FPG (<5.6 mmol/l)	5.6 (5.2-6)	6 (5.5-7.2)	6.1 (6-6.9)	**0,0002**
HbA1c (%)	5.5 (5.3-5.7)	5.8 (5.5-6.4)	6.0 (5.8-6.3)	**<0.0001**
HbA1c (mmol/mol)	37 (34-39)	40 (37-46)	42 (40-45)	
Insulin Fasted (pmol/L)	72 (48.5-101.5)	99.5 (69.6-156)	120.2 (90-151.7)	**0,0005**
Insulin Post prandial (pmol/L)	366 (248-581.6)	538 (334.6-769.5)	737.9 (670.7-927.5)	**0,0007**
Triglycerides (<1.7 mmol/l)	1.2 (0.9-1.5)	1.4 (1.1-1.9)	1.6 (1.2-2.2)	**0,0148**
HDL cholesterol (>1.0 mmol/l)	1.2 (1-1.4)	1.1 (1-1.4)	1.2 (0.9-1.4)	0,3271
LDL cholesterol (< 3.0 mmol/l)	3 (2.4-3.7)	3 (2.4-3.7)	3.4 (3-3.8)	0,2557
Total cholesterol (1.5 - 6.5 mmol/l)	4.7 (4.3-5.6)	4.8 (3.8-5.6)	5.1 (4.4-5.6)	0,4935
Fat free mass (percentage)	58.5 (53.9-62.1)	62.2 (56.8-76.5)	62.7 (58-70)	**0,0115**
Total body water (L)	42.9 (39.2-47.5)	45.1 (41.3-56.4)	46.4 (43.5-52)	**0,0140**
Extracellular water (L)	19.1 (17.2-20.8)	20.3 (18.5-26)	20.6 (19.2-23.6)	**0,0060**
Intracellular water (L)	23.9 (22-26.4)	26.1 (23.5-32.1)	25.8 (24-28.4)	**0,0056**
Histological parameters (number)				
Steatosis grade score (0/1/2/3)	58/0/0/0	0/60/11/2	0/6/7/2	**<0.0001**
Lobular inflammation score (0/1/2)	25/30/3	12/55/6	0/8/7	**<0.0001**
Hepatocyte ballooning score (0/1/2)	57/1/0	72/0/0	0/9/6	**<0.0001**
Fibrosis state (0/1/2/3/4)	7/42/9/0/0	3/55/15/0/0	0/9/6/0/0	**0,0394**
SAF-Activity (0/1/2/3/4)	25/29/4/0/0	12/55/6/0/0	0/0/4/9/2	**<0.0001**
NAFLD Activity Score	1 (0-1)	2 (2-2)	4 (4-5)	**<0.0001**

Table 1. Values are denoted as median (IQR1, IQR3). Continuous variables were tested using the Kruskal-Wallis test. Categorical variables were tested using the chi-squared test. ALP, alkaline phosphatase; g-GT, gamma-glutamyl transferase; ALT, alanine transaminase; AST, aspartate aminotransferase; FPG, fasting plasma glucose. Continues variables were tested using Kruskal-Wallis test, Categorical variables were tested using chisq test.

Peripheral ethanol concentrations were lowest in fasting conditions (means: 0.051 mM,0.064 mM and 0.098 mM and increased 120 minutes after intake of a mixed meal in 101 out of 109 individuals (**Figure 1C**). Although fasting ethanol concentrations differed between the no steatosis, NAFL and NASH groups, the post prandial increase in plasma ethanol was more profound in the NAFL (+0.039 mM; $p<0.01$) and NASH (+ 0.073 mM; $p<0.001$) groups as compared to no steatosis (+0.020; $p<0.0001$). Ethanol in both fasting and postprandial conditions increased in a dose dependent matter with an increase in hepatic

steatosis score, hepatic ballooning score, fibrosis score but not with lobular inflammation (**Extended Figure 2**). Fasting ethanol concentrations significantly correlated with fasting insulin levels (R = 0.23, p = 0.0064). In addition, post prandial insulin concentrations correlated significantly with portal vein ethanol (R = 0.43, p = 0.016) and with ethanol in the fating state (R = 0.21, p = 0.013) (**Extended Figure 3**).

Patient Characteristics Intervention Study	**No Steatosis (N=10)**	**NASH (N=10)**	**P-value**
Demographic			
Age (years)	39 (37.2-47)	51 (48.8-51.8)	0,1385
Female	2/10	2/10	1
Anthropometric			
BMI (kg/m2)	29.9 (27.6-31.8)	32 (29.9-34.3)	0,1509
Type 2 diabetes mellitus (n)	0/10	5/10	**0,0389**
Clinical lab values (normal range)			
ALP (30-135 U/L)	65 (61-73)	91 (71-114.5)	**0,0190**
g-GT (10 - 40 IU/l)	24 (16-31.2)	51.5 (42.2-57)	**0,0002**
ALT (0 - 50 IU/l)	30 (20.5-35.5)	73.5 (54.5-104.5)	**0,0002**
AST (0 - 35 IU/l)	25 (21-26)	47 (38-57.8)	**0,0011**
FPG (<5.6 mmol/l)	5 (4.8-5.3)	6.5 (5.6-8.2)	**0,0015**
HbA1c (%)	5.4 (5.3-5.5)	6.2 (5.9-7.6)	**0,0099**
HbA1c (mmol/mol)	36 (34-37)	44 (41-60)	**0,0099**
Triglycerides (<1.7 mmol/l)	1.4 (1-2)	1.5 (1.3-3)	0,2729
HDL cholesterol (>1.0 mmol/l)	1.2 (1-1.3)	1.2 (1-1.3)	0,8797
LDL cholesterol (< 3.0 mmol/l)	3 (2.6-3.3)	3.3 (2.7-3.9)	0,4961
Total cholesterol (1.5 - 6.5 mmol/l)	4.8 (4.5-5.6)	5.5 (4.9-5.8)	0,3644
Histological parameters (number)			
Steatosis grade score (0/1/2/3)		0/4/6/0	n/a
Lobular inflammation score (0/1/2)		0/6/4	n/a
Hepatocyte ballooning score (0/1/2)		0/8/2	n/a
Fibrosis state (0/1/2/3/4)		0/2/6/2/0	n/a
SAF-Activity (0/1/2/3/4)		0/0/5/4/1	n/a
NAFLD Activity Score		4 (3.2-5)	n/a

Table 2. Values are denoted as median (IQR1, IQR3). Continuous variables were tested using the Kruskal-Wallis test. Categorical variables were tested using the chi-squared test.

Together, our data show that the liver has a large capacity to metabolize ethanol as reflected by the concentration difference between peripheral and portal vein blood. In individuals with NAFLD, ethanol concentrations were 187 (IQR:17-516) times lower in peripheral compared to portal vein blood (**Figure 1**). And though taken weeks apart, portal and peripheral blood ethanol concentrations correlated significantly (**Extended Figure 4**). Transcriptomic analyses of the liver showed differences between NAFLD classes in processes including one carbon metabolism, PPAR signaling and apoptosis, as reported

in literature[6,35,36]. In addition, significant associations between post prandial ethanol concentrations with increased levels of mitochondrial encoded cytochrome B and ethanol induced epigenetic changes of the liver were observed (**Supplementary Table S3, Extended Figure 5**). Transcripts for *ADH1A* and *CYP2E1* were very highly expressed among all subjects and did not differ between groups. *ADH1A* and *CYP2E1* expression did not associate with ethanol concentrations suggesting constitutive expression and post-transcriptional regulation of these genes.

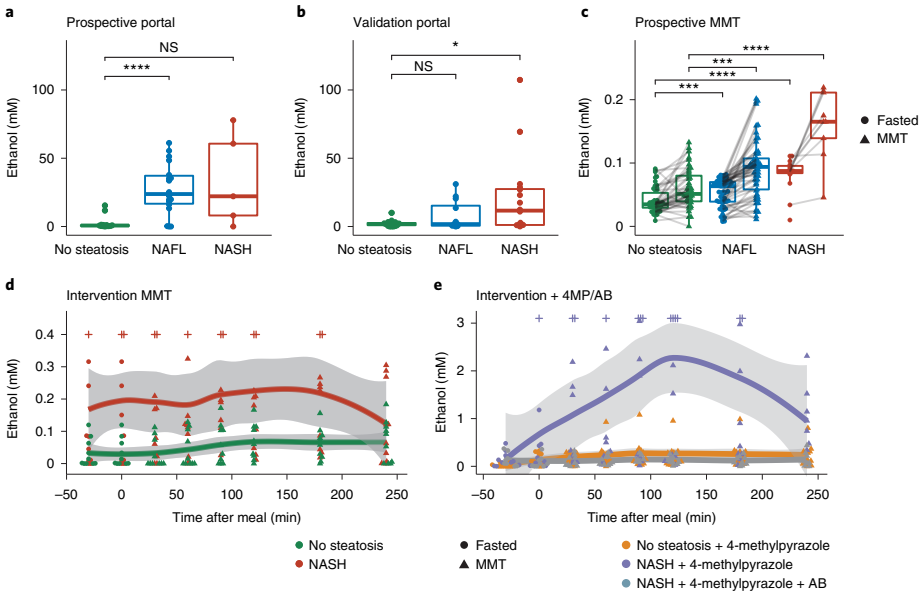


Figure 1. Ethanol concentrations in different blood compartments and studies. Ethanol concentrations in portal blood in the **a**, prospective cohort healthy (n=14, NAFL n=18, NASH n=5) and **b**, validation cohort (healthy n=22, NAFL n=12, NASH n=16). **c**, fasted and postprandial peripheral ethanol concentrations at baseline (fasting) and 120min into the mixed-meal test (healthy n=58,43, NAFL n=73,57, NASH n=15,9). **d**, ethanol concentrations during a standard mixed-meal test (D, Healthy n=10, NASH n=10) and **e**, during a mixed-meal test following 4-methylpyrazole infusion and antibiotics use (healthy n=10, NASH n=10, AB n=9). Box plots feature the median (center line), upper and lower quartiles (box limits), 1.5× the interquartile range (whiskers), points outside of boxplot range are outliers. Grey areas around the spline show the standard error. For visualization purpose, samples marked with the + in panels D and E were above the chosen axis limit. Significant differences were determined by two-tailed Mann-Whitney test (****P< 0.0001; ***P< 0.001; **P< 0.01; *P< 0.05)

Intervention with 4-methylpyrazole and broad-spectrum antibiotics

To determine whether microbial ethanol production is increased in patients with NAFLD, we designed an intervention study in which we bypassed the first-pass effect of the liver for ethanol using the selective alcohol dehydrogenase inhibitor 4-methylpyrazole. Twenty individuals were included (ten with NASH; ten age-, BMI- and sex-matched controls without NAFLD, **Table 2**) who, in line with the findings of the prospective study, showed increased

ethanol levels 120 minutes post prandially (0.050 ± 0.025; p=0.06). Ethanol levels were significantly higher in the NASH group (+0.136±0.064; p=0.047) compared to the control group (**Figure 1D**).

Infusion with 4-methylpyrazole prior to the mixed meal test strongly increased peripheral microbial ethanol levels (**Figure 1E**). In individuals without steatosis, ethanol concentrations increased at a rate of 0.09 (0.02-0.17) mM per hour and at a rate of 0.9 (0.07-1.70) mM per hour in the NASH group in the first 120 minutes after start of the infusion (**Figure 1E**). The decrease in microbial ethanol concentrations in the second half of the experiment is assumed to be a result of decreasing 4-methylpyrazole concentrations because the peak effect of 4-methylpryazole is between 1.5-2 hours after the infusion (**Supplementary table S5**). Next, nine out of the 10 individuals with NASH were given broad-spectrum oral antibiotics for one week to show that ethanol is indeed produced by gut bacteria. Antibiotics treatment completely depleted the gut microbiome, as evidenced by the fact that fecal microbial DNA could not be isolated (data not shown), which translated in a striking near-complete suppression of detectable ethanol during a mixed-meal test preceded by 4-methylpyrazole infusion. These findings support that overt ethanol production in NASH is likely to be gut microbiome driven (**Figure 1E).**

Gut microbiome analyses

Several bacterial species are known to produce relevant amounts of ethanol via the fermentation of dietary and non-dietary carbohydrates[37]. To identify the main microbial taxa contributing to overt ethanol production in NASH, we performed 16S rRNA gene amplicon sequencing on small intestinal tissues from individuals in the prospective cohort. The most abundant genus *Streptococcus* in the biopsies tended to be positively associated with NAFLD (p=0.018) (**Extended Figure 6A**), although the associations were not significant after correction for multiple testing. For a subset of 11 samples, for which portal ethanol concentrations were available, relative *Streptococcus* abundance was associated with blood ethanol concentrations but was not significant after multiple testing correction (R = 0.53, p = 0.078) (**Extended Figure 6**).

Fecal microbial alpha and beta diversity metrics (shotgun metagenomics) did not differ between the NAFLD classes in the prospective cohort, nor did it associate with ethanol levels (**Extended Figure 7 and 8**). Nevertheless, differential taxa abundance and correlation analyses revealed strong specific taxa associations with NAFLD classes (**Figure 2A**), and post prandial ethanol concentrations (**Figure 2B**), with the strongest positive correlations observed for *Streptococcus* and especially *Lactobacillus* species, both belonging to the order of lactic acid bacteria. Although fasting ethanol concentrations did not correlate with the fecal proportions of *Lactobacillaceae* (R = 0.097, p = 0.29) (**Figure 2C**), post prandial

ethanol concentrations correlated with this bacterial family (R = 0.42, p = 5.6e-06) (**Figure 2D, Supplementary Table S4**). The associations between lactic acid bacteria and post-prandial ethanol remained significant, when correcting for the use of microbiome-altering drugs including metformin, proton pump inhibitors and statins (p=0.0003). Proton-pump inhibitor use, a known risk factor for NAFLD[38], was associated with post prandial ethanol concentrations (+0.030±0.011; p=0.008) and the association was lactic acid bacteria mediated ($p<10^{-16}$). We also assessed the abundance of fungi in the metagenomes in the prospective cohort but detected only low abundances of *Saccharomyces cerevisiae* (able to produce ethanol[37]) that were slightly higher in the NAFL group but did not correlate with portal (R = 0.13, p = 0.52), fasted (R = 0.098, p = 0.3) or post prandial (R = 0.14, p = 0.17) ethanol concentrations (**Extended Figure 9**). No significant associations could be made between kegg orthologous functions and observed ethanol levels (**Supplementary Table S5**).

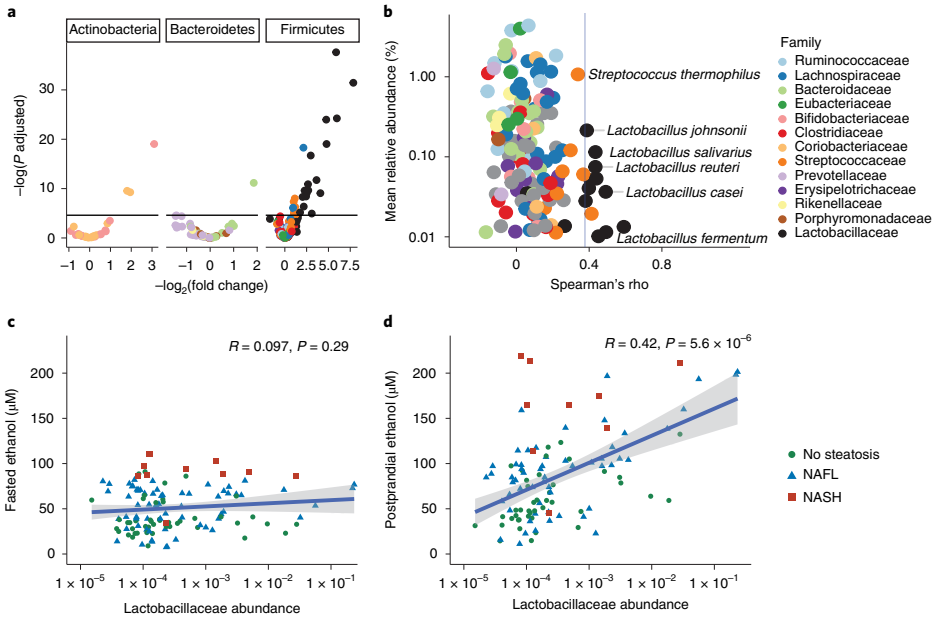


Figure 2. Gut microbiome analyses of the studies. a, differentially abundant taxa between healthy (n=47) and NALFD subjects (n=75) of the prospective cohort represented by a volcano plot. **b**, Spearman's Rho correlation coefficients of post-prandial peripheral ethanol concentrations and the most abundant species. The horizontal blue line marks the boundary for which the Benjamini & Hochberg adjusted p-value=0.05. Only taxa with mean abundances above 0.1% and belonging tothe top 13 families are shown. **c**, Lactobacillaceae abundance correlations with fasting (Healthy n=47: NALF n=64, NASH n=11) and **d**, post-prandial ethanol concentrations (healthy n=42: NALF n=57, NASH n=9). Lines in C and D represents a linear fit, with grey areas the standard error with a 0.95 confidence interval.

Diversity and dissimilarity measures in shotgun metagenomics data from fecal samples of individuals in the intervention cohort did not significantly differ between groups

(**Extended Figure 10A** and **B**). The microbiome composition and the rates of ethanol accumulation after 4-methylpyrazole infusion also did not correlate. Differential abundance analysis, however, revealed that several taxa from the order *Clostridiales* were decreased in the NASH group (**Extended Figure 10D, Supplementary Table S6**). Various taxa from multiple families were increased in the NASH group including *Lactobacillaceae* and *Streptococcaceae* (**Extended Figure 10D**). These results are in line with the microbiome data obtained from our prospective cohort and support our hypothesis that lactic acid bacteria might be implicated in NAFLD etiology via production of ethanol.

DISCUSSION

The putative role of gut microbial produced ethanol in the etiology of NAFLD has been proposed previously[1,2,9–11]. Gut microbial production of ethanol *per se*, however, has never been convincingly measured in individuals with and without NAFLD. Here, we show in two independent cohorts that ethanol concentrations in portal vein blood were significantly higher in individuals with NAFL and NASH compared to individuals with no hepatic steatosis. Peripheral ethanol concentrations could be induced during a standardized mixed meal test in all participants, with highest concentrations in those with more advanced disease (i.e., NASH). Inhibition of hepatic ethanol clearance using 4-methylpyrazole infusion implicated that individuals with NASH have an increased microbial capacity to produce ethanol compared to healthy individuals. This finding was further supported by the demonstration of completing abrogation of mixed-meal test-induced ethanol production in patients with NASH upon broad spectrum antibiotics treatment. Our data suggest that the livers of some individuals with NAFLD could be chronically exposed to increased amounts of ethanol of gut microbial origin.

Blood from the gastrointestinal tract drains directly into the portal vein and is enriched in microbial metabolites[39]. Because the portal circulation can only be sampled under invasive conditions such as abdominal surgery, it is rarely studied in diseases associated with the gut microbiome including NAFLD. Because the liver clears the vast majority of metabolites from the portal circulation (first pass effect), there hence is a black box in our knowledge on 'true' gut microbial metabolite load on the liver. This includes microbial produced ethanol.

Median ethanol concentrations in portal vein blood were low in individuals with no steatosis (mean 2.1 mM) whereas in individuals with NAFL and NASH median ethanol concentration ranged from 8.0 mM to 21.0 mM, respectively, which is in higher than the federal legal driving limit in the United States[40]. The lack of systemic alcohol misuse symptoms could be explained by the large ethanol clearance capacity of the liver, explaining

the 187 times lower peripheral ethanol concentrations in our cohort and implies that liver functional capacity to clear the portal influx of ethanol is intact. This assertion is supported by our RNA sequencing analyses which did not reveal differences in pathways of relevance for ethanol clearance despite the observation that *ADH* and *CYP2E1* were among the most abundantly expressed genes in the liver in all individuals independent of disease state.

Peripheral ethanol concentrations could be induced by administration of a standardized mixed meal. The post-prandial increase in plasma ethanol was most profound in the NAFL and NASH groups, which is in line with previous observations[1,2,9–11]. This suggests that ethanol concentrations are strongly linked to dietary stimuli, which is of particular relevance for individuals with obesity and who are at increased risk to develop NAFLD. Although peripheral ethanol concentrations were lower compared to portal ethanol concentrations, the concentrations in these compartments correlated significantly. We therefore speculate that peripheral ethanol concentrations, measured in mixed meal test settings, could be used as proxy for portal ethanol supply to the liver.

Microbial ethanol concentrations in plasma were undetectable during mixed meal test (preceded by 4-methylpyrazol infusion) in NASH participants treated with broad-spectrum antibiotics, in whom the gut microbiota was completely depleted. This observation indicates a causal role for gut bacteria in overt ethanol production in the individuals with NASH included in the present study. Both gram-negative and gram-positive bacteria are capable of producing ethanol in high concentrations[41–44]. The fecal gut microbiome has been shown to shift towards a more gram-negative community with disease progression (*e.g.*, from NAFL to NASH)[26]. Bacterial species belonging to the phylum Proteobacteria and family *Enterobacteriaceae* have been associated with fasting ethanol levels in NAFLD[1,45]. In our cohorts, next to Proteobacteria, species belonging to the lactic acid bacteria were higher in NAFLD and correlated significantly with ethanol in each blood compartment. Recently, *Klebsiella pneumoniae* was identified as a high ethanol producer in NAFLD[2]. *K. pneumoniae* contributed only little to the overall microbial load signature and was 20 times less abundant than the *Lactobacillaceae*. Furthermore, it did not correlate with ethanol levels in any compartment. Contributions of these lineages may vary depending on ethnicity, genetics, and other demographic factors[46,47].

There are several limitations of this study. The prospective study was limited by the use of tissue and plasma samples from individuals with severe obesity who underwent bariatric surgery, which might introduce bias in reproducibility of the findings in cohorts with less severe obesity and who have not had bariatric surgery. Furthermore, the cohort was sex-biased due to the skew towards females seeking surgical treatment for severe obesity[48]. Although *ADH* is constitutively active, *CYP2E1* can be induced by ethanol on transcriptional and protein levels[49]. Despite varying portal ethanol levels, in our study, *CYP2E1* expression

did not differ between individuals with and without NAFLD. We cannot rule out that the surgical procedure, when the liver is subjected to hemodynamics, stress and hormonal changes, altered (or blunted) *CYP2E1* expression. We used an enzyme-based kit to measure ethanol which is known to have activity towards other low molecular weight alcohols as well. Nevertheless, the kinetics of these side reactions are different because the turnover time is much slower comparing ethanol for instance with methanol. In 25% of the portal samples, we used an HPLC approach to show that the alcohol detected by the kit is in fact ethanol. In the intervention study we assessed hepatic steatosis in the healthy volunteers using an ultrasound and not with the gold standard (i.e., a liver biopsy), which could potentially falsely classify an individual as healthy controls. Although we obtained casual evidence that the gut microbiome can produce large amounts of ethanol, the impact on the disease course of NAFLD was not assessed. Further prospective and intervention studies in large human cohorts combined with dedicated mechanistic studies are required to obtain causal evidence that microbial produced ethanol affects NAFLD development. Also. metagenomic data analysis approaches were limited by multiple testing burden and were not fully exhaustive. In addition, despite the correlations, we have not identified the specific bacterial strains that produce ethanol. Although it has been shown that a multitude of bacterial strains can produce ethanol, including species belonging to lactic acid bacteria, the specific conditions that trigger this fundamental but dynamic pathway along the human digestive tract, require further research.

We conclude that the human gut microbiota produces large amounts of ethanol that might be clinically relevant for the pathogenesis of NAFLD. Ethanol production during a mixed meal test should be considered as a non-invasive diagnostic approach for the detection of high ethanol producing gut microbiomes and NAFLD risk. In our prospective cohort, high post prandial plasma ethanol concentrations correlated particularly with high relative fecal abundance of lactic acid bacteria. Clinical trials targeting the gut microbiome have not yielded any meaningful outcome in NAFLD thus far. To what extent persistent endogenous ethanol production is causally involved in the highly complex pathogenesis of NAFLD where a combination of environmental factors, genetic variants, obesity and disturbed lipid homeostasis interact, remain to be elucidated. Nevertheless, our findings suggest that further attention aiming to target the gut microbiota to reduce ethanol production and thereby lower additional risk for NAFLD is justified.

Acknowledgements

M.N. is supported by a ZONMW VICI grant 2020 [number 09150182010020]. H.H. is supported by a Senior Fellowship of the Dutch Diabetes Research Foundation (2019.82.004). A.S.M. is supported by Lilly/EFSD grant for young investigators. O.A. and V.T. are appointed

on a NNF GUTMMM grant 2016 NNF15OC0016798 (to MN, TWS and FB). The study reported here was additionally supported by Le Ducq consortium grant 17CVD01 to FB and MN. M.N, H.H. and A.K.G. are supported by a Dutch Heart Foundation CVON IN CONTROL-2 consortium grant. S.F. has a senior clinical research mandate from the Fund for Scientific Research (FWO) Flanders (1802154N). A.G.H. is supported by the Amsterdam UMC Fellowship, two TKI-PPP grants from Health~Holland and the Gilead Research Scholar grant.

Contributions

A.S., M.N., A.K., FB., V.G., conceived and designed the studies. - A.S.M., H.H., M.D., O.A., S.B., M.L.D.B, J.V., managed the clinical studies and collection of stool samples and clinical data. M.K., was responsible for the study medication. - A.S.M., O.A., A.G.H, M.T., were responsible for inclusion of participants. - E.D., J.W., A.V., C.D. B., L.V., S.F., delivered the samples of the validation cohort. - M.D., V.T., F.B., T.S., J.N. oversaw the processing of the data. - M.D., A.K., A.S.M., analyzed the data. A.S.M., M.D., H.H., U.B., A.K., M.N. wrote the manuscript, with input from all the authors

Competing interest statement

M.N. is in the scientific board of Caelus Pharmaceuticals, the Netherlands. F.B. is in the scientific board of Metabogen AB, Sweden. However, none of these are directly relevant for the current paper. S.F. has acted as advisor and/or lecturer for Roche, Gilead, Abbvie, Bayer, BMS, MSD, Janssen, Actelion, Astellas, Genfit, Inventiva, Intercept, Genentech, Galmed, Promethera, Coherus and NGM Bio. However, none of these are directly relevant for the current paper. A.G.H. has consulted for Novo Nordisk, Gilead, Amgen, Echosens and Julius Clinical. The remaining authors declare no competing interests'

REFERENCES

1. Zhu, L. *et al.* Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: A connection between endogenous alcohol and NASH. *Hepatology* **57**, 601–609 (2013).
2. Yuan, J. *et al.* Fatty Liver Disease Caused by High-Alcohol-Producing Klebsiella pneumoniae. *Cell Metab* **30**, 675-688.e7 (2019).
3. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
4. Farrell, G. C. & Larter, C. Z. Nonalcoholic fatty liver disease: From steatosis to cirrhosis. *Hepatology* **43**, S99–S112 (2006).
5. WK, S., V, T., SS, C. & AM, D. Similarities and differences in the pathogenesis of alcoholic and nonalcoholic steatohepatitis. *Semin Liver Dis* **29**, 200–210 (2009).
6. Diehl, A. M. & Day, C. Cause, Pathogenesis, and Treatment of Nonalcoholic Steatohepatitis. *New England Journal of Medicine* **377**, 2063–2072 (2017).
7. Thursz, M. *et al.* EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. *J Hepatol* **69**, 154–181 (2018).
8. Engstler, A. J. *et al.* Insulin resistance alters hepatic ethanol metabolism: studies in mice and children with non-alcoholic fatty liver disease. *Gut* **65**, 1–8 (2015).
9. Cope, K., Risby, T. & Diehl, A. M. Increased gastrointestinal ethanol production in obese mice: Implications for fatty liver disease pathogenesis. *Gastroenterology* **119**, 1340–1347 (2000).
10. de Medeiros, I. C. & de Lima, J. G. Is nonalcoholic fatty liver disease an endogenous alcoholic fatty liver disease? - A mechanistic hypothesis. *Med Hypotheses* **85**, 148–152 (2015).
11. Simic, M., Ajdukovic, N., Veselinovic, I., Mitrovic, M. & Djurendic-Brenesel, M. Endogenous ethanol production in patients with Diabetes Mellitus as a medicolegal problem. *Forensic Sci Int* **216**, 97–100 (2012).
12. Jeon S, C. R. Alcohol effects on hepatic lipid metabolism. *Journal of lipid research* vol. 61 470–479 (2020).
13. Cederbaum, A. I. ALCOHOL METABOLISM. *Clin Liver Dis* **16**, 667 (2012).
14. Wilson D.F, M. F. M. Ethanol metabolism: The good, the bad, and the ugly. *Med Hypotheses* **140**, (2020).
15. van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *J Intern Med* joim.13157 (2020) doi:10.1111/joim.13157.
16. Meijnikman, A. S. *et al.* Distinct differences in gut microbial composition and functional potential from lean to morbidly obese subjects. *J Intern Med* (2020) doi:10.1111/joim.13137.
17. Nascimbeni, F. *et al.* Clinical validation of the FLIP algorithm and the SAF score in patients with non-alcoholic fatty liver disease. *J Hepatol* (2020) doi:10.1016/j.jhep.2019.12.008.
18. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* (2005) doi:10.1002/hep.20701.
19. Utility and Appropriateness of the Fatty Liver Inhibition of Progression (FLIP) Algorithm and Steatosis, Activity, and Fibrosis (SAF) Score in the Evaluation of Biopsies of Nonalcoholic Fatty Liver Disease.
20. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* (2017) doi:10.1038/nbt.3960.
21. Karlsson, F. H., Nookaew, I. & Nielsen, J. Metagenomic Data Utilization and Analysis (MEDUSA) and Construction of a Global Gut Microbial Gene Catalogue. *PLoS Comput Biol* **10**, (2014).
22. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* (2012) doi:10.1038/nmeth.1923.
23. Chamberlain, S. A. & Szöcs, E. taxize : taxonomic search and retrieval in R [v2 ; ref status : indexed , http://f1000r.es/24v]. **191**, 1–28 (2014).
24. Soverini, M. *et al.* HumanMycobiomeScan: A new bioinformatics tool for the characterization of the fungal fraction in metagenomic samples. *BMC Genomics* **20**, 1–7 (2019).
25. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
26. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* 081257 (2016) doi:10.1101/081257.

27. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**, 5261–5267 (2007).

28. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590–D596 (2013).

29. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–20 (2014).

30. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525–527 (2016).

31. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at (2013).

32. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

33. Oksanen, J. *et al. Package 'vegan'. Community Ecology Package*. (2019).

34. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* vol. 57 289–300 Preprint at https://doi.org/10.2307/2346101 (1995).

35. Mardinoglu, A. *et al.* An Integrated Understanding of the Rapid Metabolic Benefits of a Carbohydrate-Restricted Diet on Hepatic Steatosis in Humans. *Cell Metab* **27**, (2018).

36. Francque, S. M. *et al.* A Randomized, Controlled Trial of the Pan-PPAR Agonist Lanifibranor in NASH. *New England Journal of Medicine* **385**, (2021).

37. Elshaghabee, F. *et al.* Ethanol Production by Selected Intestinal Microorganisms and Lactic Acid Bacteria Growing under Different Nutritional Conditions. *Front Microbiol* **7**, (2016).

38. Pyo, J. H. *et al.* Proton pump inhibitors use and the risk of fatty liver disease: A nationwide cohort study. *J Gastroenterol Hepatol* **36**, 1235–1243 (2021).

39. Krautkramer, K. A., Fan, J. & Bäckhed, F. Gut microbial metabolites as multi-kingdom intermediates. *Nature Reviews Microbiology 2020 19:2* **19**, 77–94 (2020).

40. Fell, J. C. & Scherer, M. Estimation of the Potential Effectiveness of Lowering the Blood Alcohol Concentration (BAC) Limit for Driving from 0.08 to 0.05 Grams per Deciliter in the United States. *Alcohol Clin Exp Res* **41**, (2017).

41. Nosova, T. *et al.* Aldehyde dehydrogenase activity and acetate production by aerobic bacteria representing the normal flora of human large intestine. *Alcohol Alcohol* **31**, 555–64 (1996).

42. Frantz, J. C. & McCallum, R. E. Growth yields and fermentation balance of Bacteroides fragilis cultured in glucose-enriched medium. *J Bacteriol* **137**, 1263–1270 (1979).

43. Amaretti, A. *et al.* Kinetics and metabolism of Bifidobacterium adolescentis MB 239 growing on glucose, galactose, lactose, and galactooligosaccharides. *Appl Environ Microbiol* **73**, 3637–3644 (2007).

44. Weimer, P. J. & Zeikus, J. G. Fermentation of cellulose and cellobiose by Clostridium thermocellum in the absence and presence of Methanobacterium thermoautotrophicum. *Appl Environ Microbiol* **33**, 289–297 (1977).

45. F, D. C. *et al.* Gut microbiota profiling of pediatric nonalcoholic fatty liver disease and obese patients unveiled by an integrated meta-omics-based approach. *Hepatology* **65**, 451–464 (2017).

46. Deschasaux, M. *et al.* Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat Med* **24**, (2018).

47. He, Y. *et al.* Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med* **24**, (2018).

48. Fuchs, H. F. *et al.* Benefits of bariatric surgery do not reach obese men. *Journal of Laparoendoscopic and Advanced Surgical Techniques* vol. 25 Preprint at https://doi.org/10.1089/lap.2014.0639 (2015).

49. Seitz, H. K. *et al.* Alcoholic liver disease. *Nat Rev Dis Primers* **4**, (2018).

Supplementary information available at:

Microbiome-derived ethanol in nonalcoholic fatty liver disease | Nature Medicine

9

Donor fecal microbiota transplantation alters gut microbiota and metabolites in obese individuals with steatohepatitis

Julia J. Witjes*, Loek P. Smits*, Ceyda T. Pekmez, Andrei Prodan, Abraham S. Meijnikman, Marian A. Troelstra, Kristien E.C. Bouter, Hilde Herrema, Evgeni Levin, Adriaan G. Holleboom, Maaike Winkelmeijer, Ulrich H. Beuers, Krijn van Lienden, Judith Aron-Wisnewky, Ville Mannisto, Jacques J. Bergman, Jurgen H. Runge, Aart J. Nederveen, Lars O. Dragsted, Prokopis Konstanti, Erwin G. Zoetendal, Willem de Vos, Joanne Verheij, Albert K. Groen, Max Nieuwdorp

*Shared first authors

Hepatology Communications 2020 Oct 7;4(11):1578-1590

ABSTRACT

Introduction

The intestinal microbiota has been linked to development and prevalence of steatohepatitis. Interestingly, steatohepatitis is significantly lower in individuals taking a plant-based, low animal protein diet which is thought to be mediated by gut microbiota. However, data on causality between these observations in humans is scarce. In this regard, fecal microbiota transplantation (FMT) using healthy donors is safe and is capable of changing microbial composition in human disease.

Design and Results

We thus performed a double-blind randomized controlled proof-of-principle study in which individuals with hepatic steatosis on ultrasound were randomized to two study arms; lean vegan donor (allogenic n=10) or own (autologous n=11) FMT, which were performed three times at eight-week intervals. A liver biopsy was performed at baseline and after 24 weeks in every subject to determine histopathology (NASH-CRN) classification and changes in hepatic gene expression based on RNA sequencing. Secondary outcome parameters were changes in intestinal microbiota composition and fasting plasma metabolomics. We observed a trend towards improved necro-inflammatory histology, and found significant changes in expression of hepatic genes involved in inflammation and lipid metabolism upon allogenic FMT. Intestinal microbial community structure changed upon allogenic FMT, which was associated with changes in plasma metabolites as well as markers of .

Conclusions

Allogenic FMT using lean vegan donors in individuals with hepatic steatosis shows an effect on intestinal microbiota composition, which associated with beneficial changes in plasma metabolites and markers of steatohepatitis.

INTRODUCTION

As a consequence of the pandemic spread of obesity and type 2 diabetes (T2DM), non-alcoholic fatty liver disease (NAFLD) is now recognized as the most prevalent chronic liver disease worldwide (1). NAFLD represents a spectrum of liver disease, with clinical and histological abnormalities ranging from simple steatosis (NAFL) to steatohepatitis (NASH), with the latter being diagnosed when the liver biopsy shows hepatocyte ballooning and inflammation, in addition to steatosis. The current estimated global prevalence of NAFLD is 25 – 30% and reaching staggering numbers up to 80% in individuals with metabolic syndrome and T2DM (1). Although it has been showed that individuals with NAFLD can progress towards NASH (2,3), the presence of steatosis has little prognostic value for disease development (4). Increasing evidence suggest that disease activity also known as the necro-inflammation score (i.e. inflammation and hepatocyte ballooning) independent of steatosis is clinically the most relevant parameter of NAFLD(5) (6). This relatively new concept in steatohepatitis describes and measures inflammation and liver cell injury and builds on the evidence that disease activity is highly associated with fibrosis progression (5) (6). In line, individuals with a high necro-inflammation score, thus high active inflammation, have a considerably higher risk of developing hepatic (cirrhosis, hepatocellular carcinoma, liver transplantation) and extrahepatic (mainly atherosclerotic cardiovascular) complications (1). As annual medical costs directly attributable to NAFLD keep increasing per year, this underscores the need of interventions to alleviate or even prevent an adverse disease course (1).

In search of potential new and effective treatment options, the gut microbiome has gained a lot of interest, mainly based on human observational studies and animal experiments. Indeed, alterations in gut microbial composition have frequently been observed in individuals with NAFLD (8–10). Accordingly, alterations in plasma metabolites derived from gut microbiota as well as from diet have been linked to NAFLD development (9). Compared to omnivorous diets, plant-based low animal protein diets, as practiced by vegans, are associated with reduced NAFLD incidence (11). Compared to omnivores, vegans have an altered gut microbiota composition (12) with concomitant alterations in plasma metabolites such as carnitine derivatives (13). This has previously been linked to a lower incidence of NAFLD in Chinese (14) and Western subjects (11). Although causality of these gut microbiota alterations on liver disease has been suggested in mice, in humans this remains to be elucidated. To find cause-and-effect relations between the gut microbiome and human disease in general, feces from affected individuals have been transplanted into rodents (15). Interestingly, in a recent systematic review, it was shown that 95% of published studies described the successful transfer of the pathological phenotype of human NAFLD into rodents, indicative of substantial publication bias as many studies were underpowered

(16). Combined with the complexity of causal relations, these findings suggest that this high success rate of inter-species transferable pathologies overestimates the role of the gut microbiome in human disease (16). Lessons from studies performing fecal microbiota transplantation (FMT) in humans have shown that FMT is relatively safe when performed in a clinical setting and capable of changing gut microbial composition with concomitant (modest) effects on human metabolism. For example, transfer of healthy donor feces was found to improve insulin sensitivity, alter short chain fatty acid (SCFA) production and affect plasma metabolite levels in individuals with metabolic syndrome (17). Nevertheless, not all FMT change metabolic traits or microbiota composition in treated individuals (17) and the effect seems to be modulated by donor's metabolic status as well as by the recipient's microbiota composition (18). This underscores the complexity of the relation between human diet, metabolism, composition and function of the gut microbiome in relation to cardiometabolic diseases and NAFLD. It also suggests that (diet specific) personal characteristics of both donor and acceptor determine the individual's response upon donor FMT (19). To date, dissecting causality of intestinal microbiota in NAFLD using FMT from donors on a plant-based, low animal-protein diet has not been performed. Therefore, the aim of our pilot randomized controlled trial study was to investigate a potential causal role of intestinal microbiota on NAFLD in humans.

METHODS

Design

This study was a single-center, double-blind, randomized controlled proof-of-principle pilot study comparing the effect of three eight-weekly lean vegan donor FMT versus autologous FMT on the severity of NAFLD, using liver biopsies in individuals with hepatic steatosis on ultrasound (**Supplemental figure 1**). The study was conducted in the Amsterdam University Medical Centers, location Academic Medical Center (AMC METC 2013_207), in compliance with the principles of the declaration of Helsinki and CONSORT guidelines. The protocol was reviewed and approved by the institutional review board of the AMC and was registered in the Dutch Trial Register (registration number NTR4339). All participants provided written informed consent.

Participants and donors

Caucasian, overweight, treatment-naïve, omnivorous individuals with hepatic steatosis on ultrasound were included. The main inclusion criteria were age 21-69 years, male or postmenopausal female, BMI > 25 kg/m2 with hepatic steatosis on previous ultrasound

with suspicion of NAFLD (based on elevated liver enzymes, impaired glucose tolerance, severity of steatosis on ultrasound). Exclusion criteria were any medication use, history of cardiovascular disease, T2DM, renal disease, cholecystectomy; compromised immunity; use of proton-pump inhibitors, antibiotics or anticoagulants in the past three months, any current use of medication, a history of moderate to heavy alcohol use (>12 grams per day), or other causes of liver disease besides NAFLD (*e.g.* hemochromatosis, auto-immune hepatitis, cirrhosis, hepatitis B or C, hemochromatosis, alpha-1 antitrypsin deficiency and alcoholic liver disease). None of the participants underwent bariatric surgery. Fecal donors were healthy, lean (BMI < 25 kg/m^2), treatment naïve, male or female Caucasian individuals on a stable (>3 months) plant-based low animal-protein (vegan) diet. They completed questionnaires on dietary and bowel habits, travel history, comorbidity including family history of diabetes mellitus and medication use. Donors were screened for the presence of infectious diseases as recommended (18,19).

Study visits

All participants were advised to retain their usual dietary habits during the study and were asked to fill out an online nutritional diary for the duration of one week before the baseline visit and the 24 weeks visit to monitor caloric intake including total calories, dietary carbohydrates, fat, proteins and fibers. Blood pressure, body weight and changes in health status were documented.

Intervention

All visits took place after an overnight fast with plasma samples taken and partly stored at -80°C for later analyses. Participants were randomized to treatments with either lean donor or autologous FMT performed according to the previously described procedure (19) at eight-week intervals (baseline gastroduodenoscopy whereas at 8 and 16 weeks a duodenal tube was placed by means of CORTRAK enteral access). This as we have previously observed that gut microbiota composition in the recipient is affected up to 8-12 weeks after donor FMT (18-19), we chose this time-window to ensure a stable donor gut microbiota composition over this 24 weeks period. The fecal samples received from the donors were collected approximately 6 hours before infusion into the recipients. Donors were not specifically matched with recipients based on histological or clinical characteristics. We preferred to use duodenal infusion for FMT administration over infusion via colonoscopy, because of the potential role of the duodenum in metabolism, combined with our established experience with, and the low complication rate of this method at our institution (18-20).

Randomization and blinding

Subjects were randomized using computerized block randomization, using blocks of four individuals. At the day of FMT, the independent physician provided the trial physician with the fecal material with the intended treatment from either the assigned donor or NAFLD individual. The study participants (*e.g.* vegan donors and NAFLD individuals) and all trial physicians (including all authors) were blinded for the treatment until completion of the trial.

Liver biopsy

Percutaneous liver biopsies were performed in the recruiting center on the basis of clinical indications according to local standard procedure. All histologic specimens were scored by a liver pathologist (J.V.) who was blinded to any other results. The NASH-CRN classification (20) was assessed with use of hematoxylin-eosin stained slides for steatosis, inflammation and ballooning and with a Sirius red stained slide for evaluation of fibrosis. The NAFLD activity score (NAS) is the unweighted sum of steatosis (0-3), lobular inflammation (0-3) and hepatocellular ballooning (0-2). RNA for RNA sequencing analysis was isolated using an RNA isolation protocol optimized for (very small) percutaneous liver biopsies directly frozen in liquid nitrogen after biopsy and stored at -80˚C (see supplemental data). RNA sequencing raw data (raw reads) were processed using Kallisto (v0.43.1) (21) to obtain gene counts. The R package tximport (22) and as a first step, the sequencing reads must be used as the basis for abundance quantification of transcriptomic features of interest, such as genes or transcripts. Several different quantification approaches have been proposed, ranging from simple counting of reads that overlap given genomic regions to more complex estimation of underlying transcript abundances. In this paper, we show that gene-level abundance estimates and statistical inference offer advantages over transcript-level analyses, in terms of performance and interpretability. We also illustrate that while the presence of differential isoform usage can lead to inflated false discovery rates in differential expression analyses on simple count matrices and transcript-level abundance estimates improve the performance in simulated data, the difference is relatively minor in several real data sets. Finally, we provide an R package (tximport was used to import gene counts into R (v3.4), where differential gene expression analysis was performed using DESeq2 (v1.16)(23). Differential gene expression was aimed at finding genes that showed a statistically significant interaction between the change in gene expression in time (between baseline and 24 weeks) and treatment allocation (autologous FMT versus allogenic vegan donor FMT). P-values for the interaction effects were adjusted using the Benjamin-Hochberg correction. Genes with adjusted p-values < 0.1 were considered significant (*i.e.* their expression levels changed differently in subjects that received autologous FMT compared to subjects that received allogenic vegan donor FMT).

Biochemistry

Glucose and C-reactive protein (CRP, Roche, Switzerland) were determined in fasted plasma samples. Also, alkaline phosphatase, gamma-GT, AST, ALT, total cholesterol, high density (HDLc) and low density (LDLc) cholesterol and triglycerides (TG) were determined in EDTA-containing fasted plasma samples using commercially available assays (Randox, Antrim, UK and DiaSys, Germany). All lipid analyses were performed using a Selectra autoanalyzer (Sopachem, The Netherlands). Low-density lipoprotein cholesterol (LDLC) was calculated using the Friedewald formula.

Plasma metabolites

Fasting plasma metabolites were measured at the University of Copenhagen. Plasma samples were centrifuged at 2000 x g for 15 min at 4°C from full blood mixed with EDTA, then stored at -80°C. The order of the samples was randomized within the analytical batch. Sample processing was performed at 4°C using an ice bath. The plasma samples were thawed on ice and subjected to protein precipitation using a 96-well Sirocco™ plate. 180 µl of solvent B (acetonitrile : methanol (50:50, v/v)) was added to plasma samples (40 µl) and spiked with an internal standard mixture of 7 compounds (10 µl) after which analyses were performed as previously described (24) (also see supplemental methods and **Supplemental Table 1**).

Fecal microbiota profiling

Fecal samples of donors and participants were taken at 0 and 24 weeks after initiation of study and analyzed for microbiota composition using 16S rRNA amplicon sequencing (25)"type":"article-journal","volume":"70"},"uris":["http://www.mendeley.com/documents/?uuid=12b779a4-f6bd-4635-bbc4-777bbc553114","http://www.mendeley.com/documents/?uuid=d7943a9f-f3c4-4f77-9338-4259dd9ac970"]}],"mendeley":{"formattedCitation":"(25. DNA extraction from fecal samples was performed using the repeated bead beating protocol as previously described (26). At baseline and 24 weeks, NAFLD individuals underwent gastro-duodenoscopy and uodenal biopsies were immediately collected in sterile tubes, snap-frozen in liquid nitrogen and stored at -80°C. DNA was isolated from duodenal biopsies using a slightly modified protocol and 16S sequencing was performed for small intestinal microbiota profiling as described previously (27) (also see supplemental methods)

Power and study endpoints

Based on previous intervention studies in NAFLD, for instance the study of Belfort et al. (28) in which treatment with pioglitazone in individuals with NASH led to a 54% reduction of steatosis compared to placebo, we performed a power analysis to calculate the number of participants necessary to detect a 25% reduction in our primary outcome parameter,

reversal of steatosis and necro-inflammation upon donor FMT. For a desired alpha of 0.05 and a desired power of 0.8, a sample size of 27 per arm was needed. Hence, 54 individuals were needed in total. A data safety and monitoring board (DSMB) was appointed for safety monitoring. The primary endpoint parameters of this study were histological change in NAFLD parameters including steatosis and hepatic necro-inflammatory activity (ballooning and lobular inflammation following NASH-CRN classification), without worsening of fibrosis in lean vegan donor versus autologous FMT treatment. To assess these outcome parameters, histopathological evaluation of a percutaneously obtained liver biopsy sample was performed at baseline and after 24 weeks (8 weeks after the third FMT) in combination with changes in hepatic gene expression (using RNA sequencing in the liver biopsy taken at baseline and after 24 weeks). Secondary outcome parameters comprised the change in intestinal microbiota composition between baseline and after 24 weeks. Other secondary outcome parameters included change in fasting plasma targeted metabolites, , plasma markers of fatty liver disease (ALT/AST) and inflammation (monocytes) at these time points.

Machine learning and follow-up statistical analyses

For baseline differences between groups, unpaired Student's t-test or the Mann-Whitney U tests were used dependent on the distribution of the data. Accordingly, data are expressed as mean ± the standard deviation or the median with interquartile range. The change in hepatic steatosis and hepatic necro-inflammatory activity (lobular inflammation and ballooning) following lean vegan donor FMT versus autologous FMT was tested using a Mann-Whitney U-test. Changes in plasma biochemistry derived outcome parameters between both treatment groups were tested using a student T-test or Mann-Whitney U-test, respectively for normal and non-parametrically distributed data. For correlation analyses, Spearman's rank test was used (as all parameters were non-parametric). A p-value < 0.05 was considered statistically significant. An Elastic Net machine learning classification algorithm in combination with a stability selection procedure (29) was used to identify biological features that changed differently between the two treatment groups as previously published (19) (also see supplemental data).

RESULTS

Between 2014 and 2017, 26 treatment-naïve obese individuals with metabolic syndrome and hepatic steatosis on ultrasound were included. In total, four of the included NAFLD individuals were excluded before randomization due to the diagnosis of new onset T2DM (n=3) or loss-of follow-up (n=1). After randomization, one individual had to be excluded due

to the diagnosis new onset T2DM. Due to slow recruitment, after 21 subjects were enrolled and completed the study, the trial was prematurely stopped. Baseline characteristics of the participants are shown in **Table 1**. Daily dietary intake, divided in four macronutrients and caloric content, did not significantly differ between the allogenic and autologous FMT recipients and remained stable over the course of the study (data not shown). Feces from four healthy lean vegan donors (two donated three times and five times respectively, the other two donated only once) were used for allogenic gut microbiota transfer to NAFLD individuals. The same donor was used for the three consecutive FMT's in each participant. There were no (serious) adverse events or adverse changes in plasma biochemistry and none of the study subjects used any medication (including no antibiotics) during the study.

	Autologous FMT (n=11)	Allogenic FMT (n=10)
Age, y	48.5 ± 10.2	51.2 ± 6.6
Male gender, %	96	86
BMI, kg/m2	31.5 ± 4.8	31.7 ± 3.5
Glucose, mmol/L	5.7 ± 0.5	5.8 ± 0.7
AST, IU/L	29.0 [26.5-33.0]	39.5 [37.0-49.5]
ALT, IU/L	48.1 ± 16.5	70.8 ± 23.4
ALP, IU/L	83.0 [54.0-120.5]	71.0 [58.8-76.8]
g-GT, IU/L	41.1 ± 21.4	45.1 ± 19.3
Cholesterol, mmol/L	5.8 ± 1.6	6.0 ± 0.8
HDL-C, mmol/L	1.2 [1.0-1.4]	1.2 [1.0-1.4]
LDL-C, mmol/L	4.0 ± 1.3	4.2 ± 0.7
Triglycerides, mmol/L	1.2 ± 0.6	1.4 ± 0.5
CRP, mg/mL	2.2 [0.8-4.3]	1.5 [0.9-3.2]
Leucocytes, 10^9/L	6.8 ± 1.8	5.8 ± 1.3
Monocytes, 10^9/L	0.56 ± 0.18	0.54 ± 0.18
Calories, kcal/day	1811.2 ± 376.3	2024.7 ± 499.3
Fat, g/day	68.6 ± 19.0	80.1 ± 19.6
Carbohydrates, g/day	191.8 ± 53.9	203.7 ± 64.3
Protein, g/day	82.9 ± 19.0	91.2 ± 27.1
Fiber, g/day	22.5 ± 6.2	18.4 ± 8.2
Steatosis, %	35.0 ± 20.7	34.1 ± 20.4
NAS score	2.45 ± 0.82	3.0 ± 0.94
Necroinflam. score	0.91 ± 0.30	1.4 ± 0.52
Fibrosis score	0.91 ± 0.70	1.2 ± 0.92

Table 1. Baseline characteristics of 21 individuals with biopsy-proven NAFLD. Data is expressed as mean ± standard deviation or median [interquartile range], depending on the distribution of the data. BMI: body mass index, ALP: alkaline phosphatase, g-GT: gamma glutamyl transferase, ALT: alanine aminotransferase, AST aspartate aminotransferase, HDLc: high-density lipoprotein cholesterol, LDLc: low-density lipoprotein cholesterol, CRP: C-reactive protein, NAS score: NAFLD activity score.

Primary outcomes

Liver histology and gene expression alterations after FMT

Analyzing paired liver biopsies for histology **(Supplemental table 3)**, we found no statistically significant change in the overall NAFLD activity score (NAS) **(Figure 1a)**, and steatosis grade **(Figure 1d).** We however did observe a trend towards improvement of the necro-inflammation score (comprising both lobular inflammation and hepatocellular ballooning) **(Figure 1b)** upon allogenic FMT approaching significance. Finally, fibrosis scores **(Figure 1c)** did not change over the period of 24 weeks in both groups.

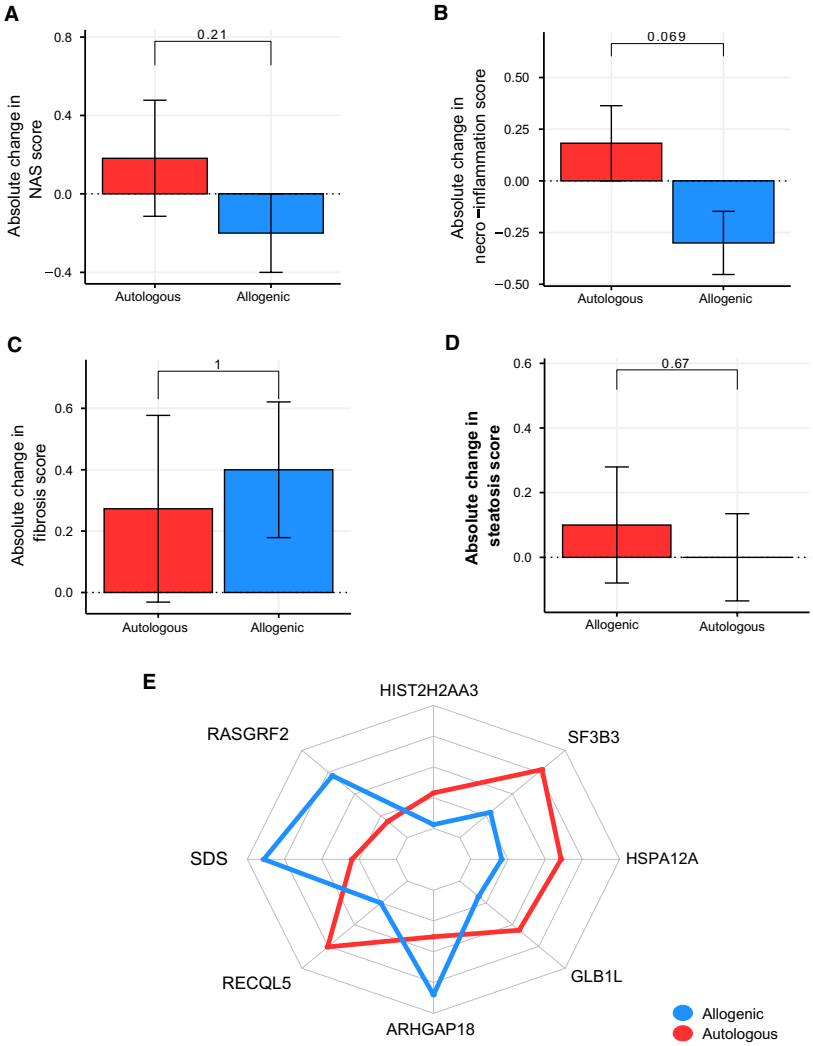


Figure 1. Changes in liver histology and gene expression. Error bars show standard errors of the means. **a)** NAFLD activity score (NAS-score); **b)** Necro-inflammation score; **c)** Fibrosis score; **d)** Steatosis score; **e)** Liver gene expression.

In line, there were significant changes in gene expression in liver biopsies in the allogenic FMT group, compared to the autologous FMT **(Figure 1e)**. For example, *ARHGAP18* expression, a protective gene that maintains endothelial cell alignment, increased upon allogenic FMT (p = 0.002). Furthermore, *serine dehydratase (SDS)* expression was significantly increased upon allogenic FMT (p = 0.049). SDS catalyzes the conversion of serine into pyruvate and ammonia and is found to be decreased during liver damage (30). In contrast, hepatic expression of *RECQL5* (p = 0.014), a gene that is implicated in DNA double strand break (DSB) repair (31) and is thus linked to the DNA damage response signaling pathway (32) and *SF3B3* (splicing factor 3b subunit 3) (p = 0.004), a gene promoting cell proliferation and known to be an early stage driver in the development of liver cancer (33,34) both increased upon autologous FMT.

Secondary outcomes

Biochemistry results 24 weeks after FMT

Upon FMT, the gamma-GT levels in the allogenic FMT group decreased (mean delta -6.4 ± 8.3, p = 0.038), while it remained unchanged in the autologous treated individuals (mean delta 0.7 ± 15.4, p = 0.883) (**Table 2**). Furthermore, ALT levels tended to decrease more in the allogenic treated group (mean delta -14.3 ± 24.6, p = 0.099) compared to the autologous treated group (mean delta -3.1 ± 20.2, p = 0.639).

	Autologous FMT – 24 weeks	P Value	Allogenic FMT – 24 weeks	P Value
Glucose, mmol/L	5.6 ± 0.8	0.241	5.8 ± 0.6	0.945
AST, IU/L	31.5 [18.8-41.3]	0.553	36.0 [29.0-42]	0.116
ALT, IU/L	46.6 ± 23.3	0.639	56.5 ± 19.2	0.099
ALP, IU/L	86.0 [66.8-112]	0.611	70.0 [57.3-83]	0.358
g-GT, IU/L	40.7 ± 28.5	0.883	38.7 ± 21.2	***0.038***
Cholesterol, mmol/L	5.5 ± 1.5	0.055	5.8 ± 0.8	0.139
HDL-C, mmol/L	1.2 [0.9-1.3]	0.280	1.1 [1.0-1.2]	0.308
LDL-C, mmol/L	3.7 ± 1.2	0.099	4.0 ± 0.8	0.378
Triglycerides, mmol/L	1.2 ± 0.6	0.796	1.4 ± 0.4	0.603
CRP, mg/mL	3.5 [0.6-6.3]	0.721	1.5 [0.7-4.4]	0.678
Leucocytes, 10^9/L	6.6 ± 1.7	0.643	6.0 ± 1.2	0.643
Monocytes, 10^9/L	0.53 ± 0.18	0.425	0.59 ± 0.25	0.460
Steatosis, %	30.5 ± 25.5	0.316	36.5 ± 25.3	0.527
NAS score	2.64 ± 1.36	0.553	2.8 ± 1.23	0.343
Necroinflam. Score	1.09 ± 0.54	0.341	1.10 ± 0.57	0.081
Fibrosis score	1.18 ± 0.75	0.391	1.60 ± 0.70	0.104

Table 2. Metabolic and histologic parameters after FMT treatment. Data is expressed as mean ± standard deviation or median [interquartile range], depending on the distribution of the data. BMI: body mass index, ALP: alkaline phosphatase, g-GT: gamma glutamyl transferase, ALT: alanine aminotransferase, AST aspartate aminotransferase, HDLc: high density lipoprotein cholesterol, LDLc: low density lipoprotein cholesterol, CRP: C-reactive protein, NAS score: NAFLD activity score.

FMT alters gut microbial composition

There was no difference in fecal microbiota alpha diversity (Shannon index) at baseline between NAFLD individuals and allogenic FMT donors (Shannon index: NAFLD 4.7 ± 0.4 vs donor 4.8 ± 0.1, ns.). Redundancy analysis showed a trend towards distinction in fecal microbiota composition between donors and NAFLD subjects (**Supplemental figure 2**) together with a significant difference of fiber intake in vegan donors (**Supplemental table 4**). Amongst others, bacteria related to *Prevotella* were associated with a plant-based diet, whereas several groups belonging to the *Lachnospiraceae* were related to the NAFLD individuals. Upon FMT, no significant changes in fecal microbiota diversity (Shannon index p = 0.84 for the allogenic FMT; p = 0.32 for the autologous FMT) were observed between baseline and week 24, however a change in gut microbiota composition, although not significant, was found upon allogenic FMT (**Supplemental figure 2**). Compared to autologous FMT, increases in fecal microbiota abundance upon allogenic FMT were seen in bacteria related to *Ruminococcus, Eubacterium hallii, Faecalibacterium* and *Prevotella copri* **(Figure 2a)**. In contrast, autologous FMT resulted in minor shifts in microbiota composition, and was primarily associated with changes in the abundance of bacteria related to *Lachnospiraceae*. There was no difference in duodenal microbiota diversity and composition before and after 24 weeks in either FMT group (data not shown).

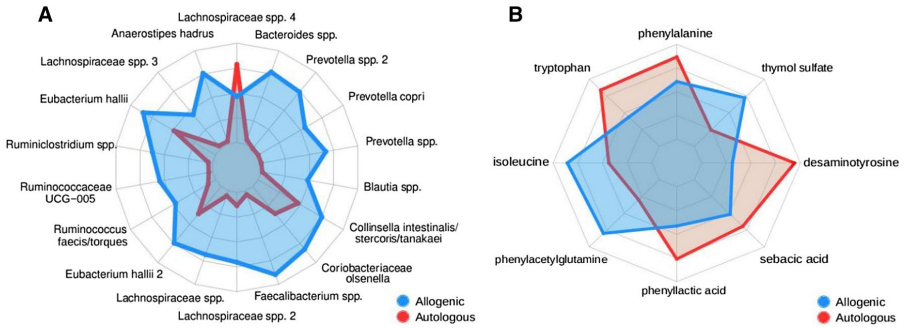


Figure 2. Radar plots of significantly altered biological features upon either autologous (red) or allogenic (blue) FMT; a) Fecal microbial strains; **b)** Plasma metabolites.

FMT alters plasma metabolite composition

Fasting plasma metabolites of both the autologous and the allogenic treated group significantly changed between baseline and 24 weeks after FMT (**Figure 2b**). Both plasma levels of the amino acids isoleucine (p = 0.039) and phenylacetylglutamine (p = 0.027) increased in plasma upon allogenic FMT (**supplemental Table 2**). In contrast, plasma phenyllactic acid, which is an adverse microbial product of aromatic amino acid metabolism,

was increased upon autologous FMT (p = 0.008). Also, plasma levels of desaminotyrosine, a microbial metabolite known to trigger type I interferon signaling (IFN1) were increased upon autologous FMT (p = 0.008). Finally, we found correlations between liver gene expression, fecal gut microbiota composition and plasma metabolite levels upon either autologous or allogenic FMT, as given in **Figure 3** and further described in the discussion.

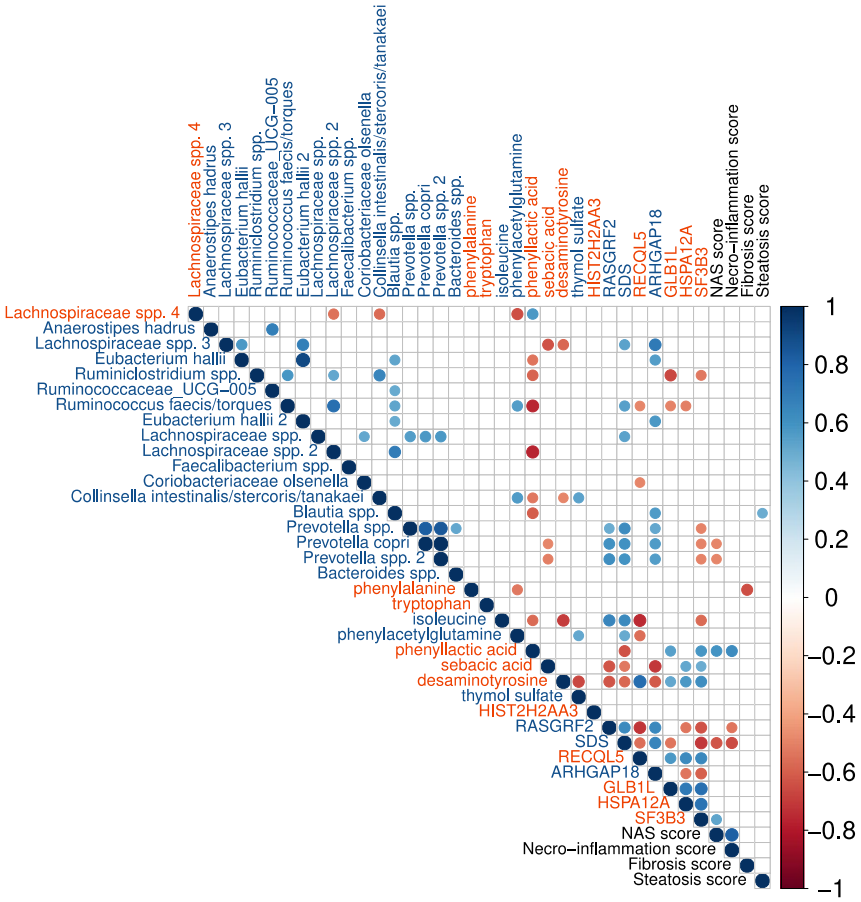


Figure 3. Correlation plot showing significant correlations between liver genes, fecal bacteria and plasma metabolite levels (Spearman's rho). Blue: increased in allogenic FMT/decreased in autologous FMT; red: increased in autologous FMT/decreased in allogenic FMT. Liver histology scores are included in black font.

DISCUSSION

We here show the effect of lean vegan donor (allogenic) versus own (autologous) fecal microbiota transplantation (FMT) on obese treatment naïve individuals with metabolic syndrome and biopsy-proven NAFLD, which are the subjects in whom one typically observes NAFLD. Although the present study was underpowered, allogenic FMT from vegan donors on a plant-based, low animal protein diet decreased the necro-inflammation score in paired liver biopsies. In addition, allogenic FMT showed an effect on intestinal microbiota composition, which was associated with both beneficial changes in plasma metabolites and the expression of liver genes involved in inflammation and lipid metabolism after donor FMT. Using differences in histology data in both treatment groups between baseline and 24 weeks after treatment, we calculated that 21 participants per treatment arm would have been needed to detect a significant beneficial effect of allogenic FMT in the necro-inflammation score, whereas 120 participants per group are needed to detect a significant difference on overall NAFLD activity score. Therefore, our study could serve as a blueprint for sample sizes and specific FMT donor selection of future microbiota-based intervention trials in liver biopsy-proven NAFLD individuals.

Interactions between changes in microbiota and liver genes/ histology upon FMT

As donor metabolic characteristics can be transferred by FMT (18) in this study we opted for FMT donors on a plant-based low animal protein diet, known to have less NAFLD (11). In line with a recent paper, assessing the effect of donor FMT on MR-spectroscopy-derived proton density fat fraction signal in NAFLD, we did not find changes in steatosis grade upon allogenic FMT as determined by liver biopsy (35), however we did find that liver necro-inflammation score improved, which was aligned by significant changes in several hepatic genes **(Figure 1).** In this regard, the *ARHGAP18* gene is a protective gene that maintains endothelial cell alignment and loss of *ARHGAP18* may predispose to atherosclerosis development (36). Liver endothelial cells play a pivotal role in maintaining liver homeostasis and endothelial cell dysfunction (i.e. loss of fenestrations) occurs early in the pathogenesis of NAFLD, promoting steatosis, inflammation and liver fibrosis (37). Moreover, a mutual relation between gut microbiota composition and liver endothelial cell fenestration has been described (38)there is a loss of fenestrations. LSECs are uniquely exposed to gut-derived dietary and microbial substrates delivered by the portal circulation to the liver. Here we studied the effect of 25 diets varying in content of macronutrients and energy on LSEC fenestrations using the Geometric Framework method in a large cohort of mice aged 15 mo. Macronutrient distribution rather than total food or energy intake was associated with changes in fenestrations. Porosity and frequency were inversely associated

with dietary fat intake, while fenestration diameter was inversely associated with protein or carbohydrate intake. Fenestrations were also linked to diet-induced changes in gut microbiome, with increased fenestrations associated with higher abundance of Firmicutes and reduced abundance of Bacteroidetes. Diet-induced changes in levels of several fatty acids (C16:0, C19:0, and C20:4, with a positive relation between the abundance of *Firmicutes* and endothelial integrity. Upon allogenic FMT, ARHGAP18 was positively correlated with *E. hallii*, suggesting a protective role of this bacterial species in maintaining liver endothelial cell function. Moreover, *ARHGAP18* was inversely correlated with plasma desaminotyrosine levels, a microbial metabolite that is known to trigger type I interferon signaling (IFN1) (39). Recently it was shown that a high-fat diet induces an IFN1 response, which via metabolically activated intrahepatic T cell pathogenicity, results in NAFLD progression (40).

Finally, we observed that after allogenic FMT, hepatic *serine dehydratase (SDS)* expression was significantly increased and showed an inverse relation with necro-inflammation and steatosis in liver histology **(Figure 3)**. In human liver biopsies, *SDS* resides predominantly in the perivenous region of the hepatocyte (41) and expression decreased during liver damage (30). *SDS* catalyzes the conversion of serine into pyruvate and ammonia and previous studies have linked these metabolites to specific microbiota composition, like *P. copri* abundance (42). With regard to the latter, in our study NAS scores were inversely related with *P. copri,* which is in line with data reported by Boursier *et al* (10), but conflict with other publications that have linked an increased abundance of this species to NAFLD (43). Hepatocyte injury and inflammatory cell infiltration in the perivenous (efferent) areas of the liver are a hallmark of NASH and this area is involved in glycolysis, lipogenesis, ureagenesis from ammonia, and biotransformation of plasma compounds, including metabolites (44). In line, plasma phenyllactic acid levels, which is a microbial product of aromatic amino acid metabolism and already linked to NAFLD in humans (45), were inversely related to hepatic *SDS* expression upon allogenic FMT. As phenyllactic acid is produced by lactic acid bacteria, the observed inverse relation between phenyllactic acid and *E. hallii* is interesting to note, as this bacterial species can use lactic acid for butyrate production and has beneficial metabolic effects in humans (46). Taken together, our data point towards a beneficial role of vegan donor FMT on prevention of NAFLD by reducing specific gut microbiota derived plasma metabolites that are toxic to the liver.

Interactions between changes in plasma metabolites and liver genes/ histology upon FMT

The linear correlation of the SDS gene with isoleucine and phenylacetylglutamine (both increased in plasma upon allogenic FMT) is of interest, as a recent paper using genome-scale models indicated that individuals with NAFLD have altered metabolism of these amino acids (47). Alterations in circulating amino acids and branched chain amino acids are often

explained to be the result of impaired amino acid metabolism linked to insulin resistance, especially in the muscle (48). Only recently it has been revealed that the gut microbiome is associated with plasma metabolite alterations of amino acids and branched chain amino acids(42). Our results strengthen the observation that the gut microbiome contributes to plasma amino acids and branched chain amino acid composition. The inverse relation between plasma phenylalanine with the fibrosis score corroborates with older literature that phenylalanine metabolism is decreased in individuals with hepatopathy, with plasma phenylalanine values decreasing with relatively early liver fibrosis (49).

However, only recently it was shown that metabolic pathways involved in the biosynthesis of phenylalanine are increased in the gut microbiome of individuals with NAFLD (9)we combine shotgun sequencing of fecal metagenomes with molecular phenomics (hepatic transcriptome and plasma and urine metabolomes. Despite the enhanced capacity of the gut microbiome to produce phenylalanine, plasma levels are dependent on the liver to catabolize this metabolite. Under normal circumstances and homeostasis, phenylalanine is converted in the liver to tyrosine and when the liver is not capable of metabolizing phenylalanine, phenyllactic acid and/or phenylacetylglutamine is produced, products that are negatively (phenylacetylglutamine) (9) we combine shotgun sequencing of fecal metagenomes with molecular phenomics (hepatic transcriptome and plasma and urine metabolomes or positively (phenyllactic acid) (45) associated with NAFLD according to recent findings. Although phenylalanine and downstream metabolites are highlighted in this study, it is most likely part of a much broader multifactorial process, only partly orchestrated by the gut microbiome. Thus, further research to prove causality of the gut microbiome in the development of NAFLD and NASH is warranted.

Study limitations

First, due to slow recruitment, our study was underpowered as we prematurely had to terminate our trial. Second, although we observed no significant differences in baseline liver histology (NAFL-NASH classifications) between autologous and allogenic FMT groups **(Table 2)**, differences in baseline AST and ALT levels were observed. As participants were randomized this difference occurred by chance but could have influenced the outcome of our study. However, if this would be the case, this baseline difference would have led to an underestimation of the allogenic donor effect as all parameters were increased in that group. A third limitation is the use of untargeted analysis of hepatic gene expression, since the genes found to be differentially expressed are not classical markers of hepatic inflammation in NAFLD. Another limitation is our choice of FMT donors, as donor metabolic characteristics can be transferred by FMT (18). We therefore chose FMT donors consuming a plant-based low animal protein diet who are known to have less NAFLD (11). Yet, we did not compare

the effect of lean FMT from donors on either an omnivorous or plant-based diet. Finally, our population is relatively healthy and does not fully reflect the typical NAFLD individual who is often affected by multiple diseases besides NAFLD and thus treated with multiple medications. This is reflected by the relatively low necro-inflammation score found. However, numerous medications, not limited to proton pump inhibitors and antibiotics, dramatically alter gut microbial composition, even with inter-individual differences (50,51). In order to demonstrate causality between the gut microbiome and the development of NAFLD, the use of medication was an exclusion criterion for participation in this study. Bearing these limitations in mind, the question to what extent the gut microbiome plays a significant role in the development of NAFLD and especially in individuals with a more progressive form (i.e. higher necro-inflammation score) remains to be answered.

In conclusion, our study shows that repetitive allogenic donor FMT in individuals with NAFLD affects hepatic gene expression and plasma metabolites involved in inflammation and lipid metabolism, highlighting the crosstalk between gut microbiota composition and NAFLD. Therefore, our work does not only further validate previous associative studies on the relation between NAFLD and gut microbiota, but also provides future sample sizes for microbiota-based intervention trials aimed at treating NAFLD in humans.

Acknowledgments

We acknowledge Ineke Heikamp-de Jong for support in the fecal sample DNA isolation. H.R. Buller and B.A. Hutten are acknowledged as DSMB members. We acknowledge Cecilie Fryland Appeldorff for UHPLC-MS analyses. Finally, we respectfully acknowledge our participants who selflessly applied themselves to help completing this burdensome project.

Author contributions

A.K.G. and M.N. designed the study. J.W. and L.P.S. performed the research. C.T.P., A.S.M., M.A.T., K.E.B., J.W.,H.H., A.G.H., M.W., U.H.B., K.vL., J.A.W., V.M., J.J.B., J.H.R., A.J.N., L.O.D., P.K., E.G.Z. and J.V. provided the analytical tools. J.W., L.P.S., A.P. and E.L. performed the statistical analysis. J.W., L.P.S., A.P., W.M.D.V., A.K.G. and M.N. drafted the paper. All authors critically reviewed the manuscript.

Funding

This trial was funded by a CVON INCONTROL grant (2012-03 and 2018.27). WMDV was also supported by the Netherlands Organization for Scientific Research (Spinoza Award). JAW is supported by a Bettencourt Schueller Foundation grant. VM is supported by a Sigrid Jusélius Fellowship grant. MN is supported by a personal ZONMW-VIDI grant 2013 [016.146.327]. The work was also partly supported by an IMI 2 LITMUS grant (777377)

and a Le Ducq consortium grant (17CVD01). CTP and MN were supported by the HDHL-JPI DINAMIC project (the Innovation Fund Denmark grant # 5195-00001B) and LOD was supported by a Semper Ardens grant from the Carlsberg Foundation.

Conflicts of interest

MN and WdV are in the Scientific Advisory Board of Caelus Pharmaceuticals, the Netherlands; MN is in the Scientific Advisory Board of Kaleido, USA; WdV is in the Scientific Advisory Board of A-Mansia, Belgium. None of these are directly relevant to the current paper. There are no patents, products in development or marketed products to declare. The other authors declare no competing financial interests.

REFERENCES

1. Younossi Z, Tacke F, Arrese M, Chander Sharma B, Mostafa I, Bugianesi E, et al. Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. Hepatology [Internet]. 2019 Jun;69(6):2672–82. Available from: http://doi.wiley.com/10.1002/hep.30251
2. Pais R, Charlotte F, Fedchuk L, Bedossa P, Lebray P, Poynard T, et al. A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. J Hepatol [Internet]. 2013;59(3):550–6. Available from: http://dx.doi.org/10.1016/j.jhep.2013.04.027
3. McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: Implications for prognosis and clinical management. J Hepatol [Internet]. 2015 May;62(5):1148–55. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168827814008836
4. Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology [Internet]. 2015 May;61(5):1547–54. Available from: http://doi.wiley.com/10.1002/hep.27368
5. Ratziu V. Back to Byzance: Querelles byzantines over NASH and fibrosis. Journal of Hepatology. 2017.
6. Kleiner DE, Brunt EM, Wilson LA, Behling C, Guy C, Contos M, et al. Association of Histologic Disease Activity With Progression of Nonalcoholic Fatty Liver Disease. JAMA Netw open. 2019 Oct;2(10):e1912565.
7. Fazel Y, Koenig AB, Sayiner M, Goodman ZD, Younossi ZM. Epidemiology and natural history of non-alcoholic fatty liver disease. Metabolism [Internet]. 2016;65(8):1017–25. Available from: http://dx.doi.org/10.1016/j.metabol.2016.01.012
8. Loomba R, Seguritan V, Li W, Long T, Klitgord N, Bhatt A, et al. Gut Microbiome-Based Metagenomic Signature for Non-invasive Detection of Advanced Fibrosis in Human Nonalcoholic Fatty Liver Disease. Cell Metab [Internet]. 2019 Sep;30(3):607. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1550413119304310
9. Hoyles L, Fernández-Real JM, Federici M, Serino M, Abbott J, Charpentier J, et al. Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. Nat Med. 2018;24(7):1070–80.
10. Boursier J, Mueller O, Barret M, Machado M, Fizanne L, Araujo-Perez F, et al. The severity of nonalcoholic fatty liver disease is associated with gut dysbiosis and shift in the metabolic function of the gut microbiota. Hepatology [Internet]. 2016 Mar;63(3):764–75. Available from: http://doi.wiley.com/10.1002/hep.28356
11. Mazidi M, Kengne AP. Higher adherence to plant-based diets are associated with lower likelihood of fatty liver. Clin Nutr. 2019 Aug;38(4):1672–7.
12. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. Gut. 2016 Jan;65(1):63–72.
13. Koeth RA, Lam-Galvez BR, Kirsop J, Wang Z, Levison BS, Gu X, et al. l-Carnitine in omnivorous diets induces an atherogenic gut microbial pathway in humans. J Clin Invest. 2018 Dec;129(1):373–87.
14. Chen Y, Liu Y, Zhou R, Chen X, Wang C, Tan X, et al. Associations of gut-flora-dependent metabolite trimethylamine-N-oxide, betaine and choline with non-alcoholic fatty liver disease in adults. Sci Rep. 2016 May;6(1):19076.
15. Meijnikman AS, Gerdes VE, Nieuwdorp M, Herrema H. Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. Endocr Rev [Internet]. 2018 Apr 1;39(2):133–53. Available from: https://academic.oup.com/edrv/article/39/2/133/4772276
16. Walter J, Armet AM, Finlay BB, Shanahan F. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. Cell [Internet]. 2020 Jan;180(2):221–32. Available from: https://linkinghub.elsevier.com/retrieve/pii/S009286741931387X
17. Vrieze A, Van Nood E, Holleman F, Salojärvi J, Kootte RS, Bartelsman JFWMWM, et al. No Title. Gastroenterology. 2012 Oct;143(4):913-916.e7.
18. de Groot P, Scheithauer T, Bakker GJ, Prodan A, Levin E, Khan MT, et al. Donor metabolic characteristics drive effects of faecal microbiota transplantation on recipient insulin sensitivity, energy expenditure and intestinal transit time. Gut [Internet]. 2020 Mar;69(3):502–12. Available from: http://gut.bmj.com/lookup/doi/10.1136/gutjnl-2019-318320

19. Smits LP, Kootte RS, Levin E, Prodan A, Fuentes S, Zoetendal EG, et al. Effect of vegan fecal microbiota transplantation on carnitine- and choline-derived trimethylamine-N-oxide production and vascular inflammation in patients with metabolic syndrome. J Am Heart Assoc. 2018;7(7).

20. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology [Internet]. 2005 Jun;41(6):1313–21. Available from: http://doi.wiley.com/10.1002/hep.20701

21. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol [Internet]. 2016 May 4;34(5):525–7. Available from: http://www.nature.com/articles/nbt.3519

22. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research [Internet]. 2015 Dec 30;4:1521. Available from: https://f1000research.com/articles/4-1521/v1

23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol [Internet]. 2014 Dec 5;15(12):550. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8

24. Barri T, Holmer-Jensen J, Hermansen K, Dragsted LO. Metabolic fingerprinting of high-fat plasma samples processed by centrifugation- and filtration-based protein precipitation delineates significant differences in metabolite information coverage. Anal Chim Acta [Internet]. 2012 Mar;718:47–57. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0003267012000062

25. Taylor M, Wood HM, Halloran SP, Quirke P. Examining the potential use and long-term stability of guaiac faecal occult blood test cards for microbial DNA 16S rRNA sequencing. J Clin Pathol [Internet]. 2017 Jul;70(7):600–6. Available from: http://jcp.bmj.com/lookup/doi/10.1136/jclinpath-2016-204165

26. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M, Kekkonen RA, et al. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J Microbiol Methods. 2010;81(2):127–34.

27. Zoetendal, E., Heilig, H., Klaassens E. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat Protoc. 2006;1(2):870–873.

28. Belfort R, Harrison SA, Brown K, Darland C, Finch J, Hardies J, et al. A Placebo-Controlled Trial of Pioglitazone in Subjects with Nonalcoholic Steatohepatitis. N Engl J Med [Internet]. 2006 Nov 30;355(22):2297–307. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa060326

29. Meinshausen N, Bühlmann P. Stability selection. J R Stat Soc Ser B (Statistical Methodol [Internet]. 2010 Jul 5;72(4):417–73. Available from: http://doi.wiley.com/10.1111/j.1467-9868.2010.00740.x

30. López-Flores I, Barroso JB, Valderrama R, Esteban FJ, Martínez-Lara E, Luque F, et al. Serine dehydratase expression decreases in rat livers injured by chronic thioacetamide ingestion. Mol Cell Biochem. 2005;268(1–2):33–43.

31. Liao W-Q. Recql5 protects against lipopolysaccharide/D-galactosamine-induced liver injury in mice. World J Gastroenterol [Internet]. 2015;21(36):10375. Available from: http://www.wjgnet.com/1007-9327/full/v21/i36/10375.htm

32. Ciccia A, Elledge SJ. The DNA Damage Response: Making It Safe to Play with Knives. Mol Cell [Internet]. 2010 Oct;40(2):179–204. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1097276510007471

33. Shen Q, Nam SW. SF3B4 as an early-stage diagnostic marker and driver of hepatocellular carcinoma. BMB Rep [Internet]. 2018 Feb 28;51(2):57–8. Available from: http://www.bmbreports.org/journal/view.html?doi=10.5483/BMBRep.2018.51.2.021

34. Mas VR, Fisher RA, Archer KJ, Maluf DG. Proteomics and liver fibrosis: identifying markers of fibrogenesis. Expert Rev Proteomics [Internet]. 2009 Aug 9;6(4):421–31. Available from: http://www.tandfonline.com/doi/full/10.1586/epr.09.59

35. Craven L, Rahman A, Nair Parvathy S, Beaton M, Silverman J, Qumosani K, et al. Allogenic Fecal Microbiota Transplantation in Patients With Nonalcoholic Fatty Liver Disease Improves Abnormal Small Intestinal Permeability. Am J Gastroenterol [Internet]. 2020 May 15;Publish Ah. Available from: https://journals.lww.com/10.14309/ajg.0000000000000661

36. Lay AJ, Coleman PR, Formaz-Preston A, Ting KK, Roediger B, Weninger W, et al. ARHGAP18: A Flow-Responsive Gene That Regulates Endothelial Cell Alignment and Protects Against Atherosclerosis. J Am Heart Assoc [Internet]. 2019 Jan 22;8(2). Available from: https://www.ahajournals.org/doi/10.1161/JAHA.118.010057

37. Hammoutene A, Rautou P-E. Role of liver sinusoidal endothelial cells in non-alcoholic fatty liver disease. J Hepatol [Internet]. 2019 Jun;70(6):1278–91. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168827819301278

38. Cogger VC, Mohamad M, Solon-Biet SM, Senior AM, Warren A, O'Reilly JN, et al. Dietary macronutrients and the aging liver sinusoidal endothelial cell. Am J Physiol Circ Physiol [Internet]. 2016 May 1;310(9):H1064–70. Available from: https://www.physiology.org/doi/10.1152/ajpheart.00949.2015

39. Steed AL, Christophi GP, Kaiko GE, Sun L, Goodwin VM, Jain U, et al. The microbial metabolite desaminotyrosine protects from influenza through type I interferon. Science (80-) [Internet]. 2017 Aug 4;357(6350):498–502. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.aam5336

40. Ghazarian M, Revelo XS, Nøhr MK, Luck H, Zeng K, Lei H, et al. Type I interferon responses drive intrahepatic T cells to promote metabolic syndrome. Sci Immunol [Internet]. 2017 Apr 21;2(10):eaai7616. Available from: https://immunology.sciencemag.org/lookup/doi/10.1126/sciimmunol.aai7616

41. Kashii T, Gomi T, Oya T, Ishii Y, Oda H, Maruyama M, et al. Some biochemical and histochemical properties of human liver serine dehydratase. Int J Biochem Cell Biol [Internet]. 2005 Mar;37(3):574–89. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1357272504003152

42. Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BAH, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature [Internet]. 2016 Jul 13;535(7612):376–81. Available from: http://www.nature.com/articles/nature18646

43. Schwimmer JB, Johnson JS, Angeles JE, Behling C, Belt PH, Borecki I, et al. Microbiome Signatures Associated With Steatohepatitis and Moderate to Severe Fibrosis in Children With Nonalcoholic Fatty Liver Disease. Gastroenterology [Internet]. 2019 Oct;157(4):1109–22. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0016508519410378

44. Abdelmalek MF, Diehl AM. Nonalcoholic Fatty Liver Disease as a Complication of Insulin Resistance. Med Clin North Am [Internet]. 2007 Nov;91(6):1125–49. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0025712507000740

45. Caussy C, Hsu C, Lo M-T, Liu A, Bettencourt R, Ajmera VH, et al. Link between gut-microbiome derived metabolite and shared gene-effects with hepatic steatosis and fibrosis in NAFLD. Hepatology [Internet]. 2018 Sep;68(3):918–32. Available from: http://doi.wiley.com/10.1002/hep.29892

46. Gilijamse PW, Hartstra A V., Levin E, Wortelboer K, Serlie MJ, Ackermans MT, et al. Treatment with Anaerobutyricum soehngenii: a pilot study of safety and dose–response effects on glucose metabolism in human subjects with metabolic syndrome. npj Biofilms Microbiomes [Internet]. 2020 Dec 27;6(1):16. Available from: http://www.nature.com/articles/s41522-020-0127-0

47. Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat Commun [Internet]. 2014 May 14;5(1):3083. Available from: http://www.nature.com/articles/ncomms4083

48. Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A Branched-Chain Amino Acid-Related Metabolic Signature that Differentiates Obese and Lean Humans and Contributes to Insulin Resistance. Cell Metab [Internet]. 2009 Apr;9(4):311–26. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1550413109000400

49. Ishii Y, Suzuki S, Kohno T, Aoki M, Kohno T, Ito A, et al. L-[1- 13 C] phenylalanine breath test reflects histological changes in the liver. J Surg Res [Internet]. 2003 Oct;114(2):120–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022480403002786

50. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, et al. Extensive impact of non-antibiotic drugs on human gut bacteria. Nature [Internet]. 2018 Mar 19;555(7698):623–8. Available from: https://www.nature.com/articles/nature25979

51. Zaneveld JR, McMinds R, Vega Thurber R. Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat Microbiol [Internet]. 2017 Sep 24;2(9):17121. Available from: http://www.nature.com/articles/nmicrobiol201712

Supplemental information available at:

Donor Fecal Microbiota Transplantation Alters Gut Microbiota and Metabolites in Obese Individuals With Steatohepatitis - Witjes - 2020 - Hepatology Communications - Wiley Online Library

PART 3

CELLULAR SENESCENCE,

an old but new player in metabolic disease

10

Evaluating causality of cellular senescence in Non-Alcoholic Fatty Liver Disease

Abraham Stijn Meijnikman, Hilde Herrema, Torsten Scheithauer, Jeffrey Kroon, Max Nieuwdorp, Albert Kornelis Groen

Jhep Reports, 2021 May 1;3(4):100301

ABSTRACT

Cellular senescence is a state of irreversible cell cycle arrest with important physiological functions. However, cellular senescence is also a hallmark of ageing and has been associated with several pathological conditions. A wide range of factors including genotoxic stress, mitogens and inflammatory cytokines can induce senescence. Phenotypically, senescent cells are characterized by short telomeres, an enlarged nuclear area and damaged genomic and mitochondrial DNA. Secretion of proinflammatory proteins, also known as senescence-associated secretory phenotype, is a characteristic of senescent cells that is mainly held accountable for their disease-inducing properties.

In the past decade, cellular senescence gained significant interest due to its putative role in the development of Non-Alcoholic Fatty Liver Disease (NAFLD) and the progression towards Non-Alcoholic Steatohepatitis (NASH). Until recently, it was suggested that hepatocyte cellular senescence is a mere consequence of the metabolic dysregulation and inflammatory phenomena in fatty liver disease. However, recent work in rodents has suggested that senescence may be a causal factor in NAFLD development. Although causality is yet to be established in humans, current evidence suggests that targeting of senescent cells has novel treatment potential for NAFLD

We aim to provide insight in the quality of the evidence supportive of a causal role of cellular senescence in the development of NAFLD in rodents and humans. We will elaborate on key cellular and molecular features of senescence and discuss the efficacy and safety of novel senolytic drugs to treat or even prevent this disease.

KEY POINTS

- Cellular senescence has been put forward as an augmenting factor in the development of NAFLD
- Senescent cells exhibit the following four hallmarks 1) prolonged and generally irreversible cell cycle arrest, 2) macromolecular damage, 3) secretory features and 4) deregulated metabolism which are all present in hepatocytes of both humans and rodents with NAFLD
- Data from studies in rodents and humans have shown that NAFLD is accompanied by an increase in senescent cells in the liver, and that the number of senescent cells is associated with disease progression
- Under normal circumstance, around 3-7% of hepatocytes are senescent and this percentage can increase to 50-100% in end stage liver disease
- Only a few markers reliably detect senescent cells at this moment in time and novel, non-invasive analyses tools are much needed to better understand the role of senescent cells in NAFLD
- Causal evidence of cellular senescence in the development of NAFLD originates from studies in rodents
- Targeting senescence has emerged as an attractive therapeutic target for NAFLD since senescence might be involved in the full spectrum of the disease
- Senolytic drugs can be administrated intermittently, thereby minimizing potential toxic effects and increasing adherence in the individual often affected by multiple morbidities and thus treated with multiple medications

INTRODUCTION

Accompanying the obesity pandemic, Non-Alcoholic Fatty Liver Disease (NAFLD) is rapidly increasing with prevalence exceeding 80% in morbidly obese individuals[1]. NAFLD represents a spectrum of liver diseases with clinical and histological abnormalities ranging from Non-Alcoholic Fatty Liver (NAFL) in case of isolated steatosis to Non-Alcoholic steatohepatitis (NASH), fibrotic NASH, advanced fibrosis, liver cirrhosis, and hepatocellular carcinoma (HCC)[2]. Accumulation of fat in hepatocytes has long been considered a relatively benign condition[3,4]. However, around 30% of individuals with NAFL progress to well-defined NASH with clinically significant fibrosis[5–7]. Advanced forms of NAFLD often require a liver transplantation and are the main cause of liver-related deaths[8]. The rapidly growing prevalence of NAFLD and lack of effective treatment options to tackle this potentially debilitating disease, will further increase the obesity-related burden on public health and economies. In order to develop appropriate, non-invasive diagnostic methods and treatment options, it is critical to deeply investigate the complex pathophysiology of NAFLD.

The underlying mechanisms that govern hepatic lipid accumulation and the predisposition to inflammation and fibrosis are complex and multifactorial. In the past decades, a multitude of disease-inducing factors have been unveiled resulting in the multi-hit hypothesis, which integrates parallel and synergistically operating disease promoting factors[2,9]. Insulin resistance[10], adipocyte dysfunction[11], genetic variants[12], bile acid metabolism[13], the gut microbiome[14], and lipotoxicity[15] are intensely studied players in NAFLD pathophysiology. These players are unified by the metabolic dysregulation accompanying obesity. Metabolic dysregulation refers to a complex range of metabolic alterations often induced by insulin resistance. Insulin resistance for example increases circulating FFA levels via reduced insulin-mediated suppression of lipolysis in the visceral adipose tissue compartment[16]. This increases hepatocellular influx of FFAs which may subsequently be stored as triglycerides. Increased fat storage in the liver is strongly linked to reduced hepatic insulin sensitivity and a consequential increase in hepatic gluconeogenesis, a major contributor to the hyperglycemia observed in diseases associated with NAFLD[17]. Moreover, de novo lipogenesis is increased due to the constant high levels of insulin, producing even more triglycerides and further enhancing hepatic gluconeogenesis[18,19]. Thus, hepatic insulin resistance in individuals with NAFLD is considered to be limited to the pathway involving suppression of hepatic glucose production and not the lipogenic pathway, which is referred to as selective insulin resistance[20]. Accumulating evidence obtained in the past decades revealed that this pathogenic paradox plays a pivotal role in the development of NAFLD[21,22].

Triglyceride accumulation is not hepatotoxic per se and could even represent a defensive mechanism to balance FFAs excess. However, high levels of free fatty acids (FFA),

free cholesterol and other lipid metabolites can lead to lipotoxicity[15]. Lipotoxicity causes mitochondrial dysfunction, resulting in the formation of reactive oxygen species, endoplasmatic reticulum (ER) stress, inflammation and cell damage[15,23]. As a consequence of the overload of FFAs, proinflammatory pathways such as c-jun terminal kinase (JNK) via apoptosis signal-regulating kinase 1 (ASK-1), and nuclear factor kappa-light-chain-enhancer of activated B-cells (NF-κB) are activated leading to hepatic inflammation and eventually fibrosis[24,25]. Expansion of subcutaneous and visceral adipose tissue compartments in obesity leads to hypoxia-induced hypersecretion of adipocytokines such as tumor necrosis factor (TNF)α, interleukin (IL) 6 and monocyte chemoattractant protein-1 (MCP-1/CCL2) by the adipocytes[26,27]. In addition, the inflammatory immune cells that accumulate in adipose tissue of individuals with obesity further perpetuates the low-grade inflammatory state[26,27]. These proinflammatory mediators are secreted into the circulation and contribute to activation of inflammatory signaling pathways in the liver, thereby contributing to development and progression of NAFLD[26,27].

Recently, cellular senescence has been put forward as an augmenting factor in the progression of NAFLD. Cellular senescence is one of the hallmarks of aging and is defined as a stable arrest of the cell cycle coupled to specific phenotypic changes[28]. Senescent cells secrete a collection of proteins called the senescence-associated secretory phenotype (SASP)[29,30]. This pro-inflammatory secretome has been suggested to drive age-related tissue dysfunction. Interestingly, metabolic dysregulation is thought to favor cellular senescence in several tissues involved in the pathogenesis of NAFLD such as the liver, pancreas and adipose tissue, further perpetuating metabolic dysregulation. The number of senescent cells is particularly increased in adipose tissues and liver of individuals with obesity[31–33]. Hepatocytic senescence has been shown to impair mitochondrial β-oxidation[34]. Concomitantly, SASP components are abundantly present in adipose tissues of individuals with obesity and may promote insulin resistance and inflammation[32]. Senescence has also been linked to the hyperinsulinemic state often observed in individuals with obesity. Senescence in pancreatic beta cells induces greater glucose uptake and mitochondrial activity, leading to increased insulin secretion[35]. This observation suggests that beta cell senescence enhances the insulin secreting capacity of the pancreas. This is in contrast to the general dogma that cellular senescence deteriorates cell and organ function[36]. Indeed, animal studies have shown that clearance of senescent cells regains tissue function including liver metabolism[34]. Targeting senescence might therefore be an interesting future therapeutic option to tackle cardiometabolic diseases, including NAFLD.

In this review, we aim to provide insight in the quality of the evidence that supports a causal role of cellular senescence in the development of NAFLD in rodents and humans. We will elaborate on defining and discussing key cellular and molecular features of senescence.

Finally, we will discuss the effectiveness and safety of novel senolytic (senescence destroying) drugs to treat this disease.

Hallmarks of cellular senescence

Cellular senescence, originally described by Hayflick and Moorhead[37], is a cellular state implicated in various physiological processes. Senescent cells exhibit the following four hallmarks 1) prolonged and generally irreversible cell cycle arrest, 2) macromolecular damage, 3) secretory features and 4) deregulated metabolism (**Figure 1)**[38,39]. Senescence is driven by a variety of factors such as genotoxic stress, mitogens and inflammatory cytokines. Also, metabolic factors including high glucose levels, ceramides, fatty acids, prostanoids and reactive oxygen species (ROS) are capable of inducing cellular senescence. Furthermore, signals originating from senescent cells have been shown to be able to transfer the senescent phenotype to neighboring cells[40]. In addition, senescence is also linked to the age-associated loss of the regenerative capacity in the liver after severe liver injury[41,42]. Following acute liver damage, senescence occurs in hepatocytes as well as nonparenchymal cells in the livers of adult but not young mice, leading to impaired regeneration. Recently, Ritschka et al.[43] showed that treatment with the senolytic drug agent ABT-737 (a BCL-2 family inhibitor) decreased the senescence markers in hepatocytes and reduced inflammation, which was associated with an improvement in liver function and regeneration following partial hepatectomy in adult mice.

Cell cycle withdrawal and macromolecular damage

These senescence-inducing signals activate transcriptional cascades which culminates in the activation of the cyclin-dependent kinase inhibitors p21 or p16, resulting in an irreversible cell cycle arrest[44]. This will eventually lead to specific phenotypic changes such as telomere shortening, nuclear area enlargement and genomic and mitochondrial DNA damage[45]. There are two major mechanisms of cellular senescence; one is replicative senescence and the other is stress-induced premature[44,46]. Replicative senescence depends on telomere shortening or erosion, predominantly upon aging, whereas stress-induced premature senescence is telomere-independent and refers to intracellular or environmental stress factors leading to macromolecular damage (i.e. DNA damage, protein damage, lipid damage)[44,46,47]. Both mechanisms induce a complex multigenic pathway known as DNA damage response (DDR)[48]. DDR can inhibit the cell cycle progression and prevent the propagation of corrupted genetic information to neighboring cells[48]. Some factors involved in the DDR, such as the phosphorylation of histone H2AX (yH2AX), the associated proteins, including MDC1, 53BP1 and the activated form of the kinase ataxia telangiectasia mutated (ATM) accumulate at sites of DNA damage. The factors form cytologically detectable nuclear foci

and mark the individual sites of DNA damage and subsequently contribute to checkpoint enforcement and cell cycle arrest, until damage has been repaired. If DNA damage persists, the tumor suppressor p53 will be phosphorylated via activation of ATM and stimulates the expression of p21, an essential mediator of senescence-associated cell cycle arrest[49]. After the early activation of p21, p16 is activated and is also suggested to play a role in several types of senescence. P16 inhibits cyclin-dependent kinase-4 (CDK4) and cyclin-dependent kinase-6 (CDK6), thus maintaining the senescence phenotype[50]. Activation of either p21 or p16 results in the inhibition of retinoblastoma factor (Rb) phosphorylation, allowing it to bind to the E2F transcription factor which prevents cell division[51].

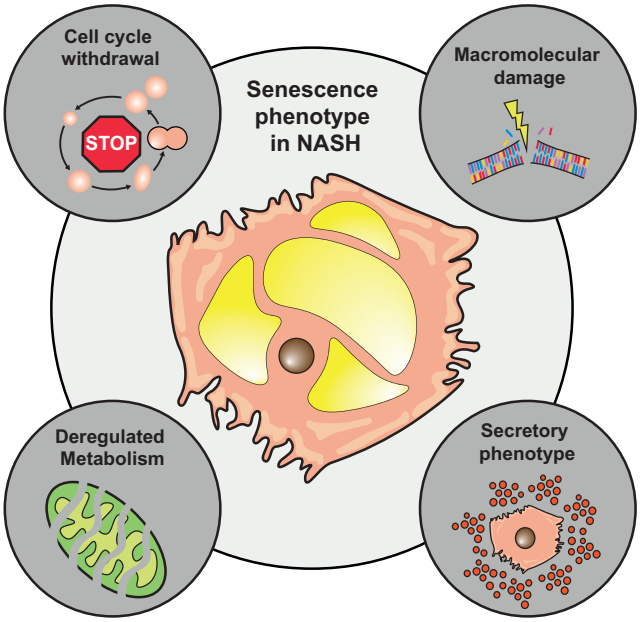


Figure 1. senescent cells in general and in NASH exhibit the following four hallmarks 1) prolonged and generally irreversible cell cycle arrest, 2) macromolecular damage, 3) secretory features and **4)** deregulated metabolism.

Senescence-associated secretory phenotype

Another hallmark of senescence is the SASP[29,30,40]. This proinflammatory secretome is a hallmark of senescent cells and contributes to tissue dysfunction in autocrine and paracrine fashion[40]. Interestingly, the SASP stimulates the immune system to clear senescent cells but can also reinforce or even maintain the senescent cell state[52–54]. Furthermore, it has been suggested that the SASP contributes to persistent chronic inflammation often found in cardiometabolic diseases including NASH and can explain some of the deleterious pro-aging effects of senescent cells[55,56]. The SASP is regulated via various mechanisms. For

example, remodeling of enhancer regions of genes results in changes of the senescent cell phenotype and induces qualitative and quantitative changes in their secretome. In addition, transcription factors such as GATA4 (an upstream regulator of NF-kB) and mammalian target of rapamycin (mTOR) as well as p38MAPK signaling pathways have been strongly implicated in regulation of the SASP[57–59]. Interestingly, GATA4 has been reported to be involved in autophagy, which is a highly regulated cellular program involved in recycling of intracellular proteins and damaged/non-functional organelles[60]. GATA4 is degraded by p62-mediated selective autophagy under normal circumstances. Interestingly, during senescence this regulation is suppressed and thus may initiate and maintains SASP facilitating senescence[58]. Recently, GATA6 has been suggested to be involved in the induction of senescence. When GATA6 accumulation is not decreased by autophagy, the expression of p53 and p16 is enhanced, while knocking down GATA6 reduces the upregulation of p53 and p16, and thereby hepatic senescence[61]. *In vitro* data showed that the SASP is produced in a p16-independent manner as a result of DDR-dependent and independent signaling through p38MAPK and NF-kB. Whereas the early SASP is dominated by growth factors such as TGFbeta which triggers senescence in an autocrine fashion[57,62], a switch towards a more pro-inflammatory SASP is established through the activation of NOTCH1, where secreted and membrane bound IL-1a acts in an autocrine fashion to reinforce the production of IL-6 and IL-8[63,64].

Dysregulated metabolism

Accumulating evidence suggests a bidirectional link between cellular senescence and mitochondria[65]. Senescent cells are capable of deregulating metabolism by altering mitochondrial function, dynamics and morphology. In the early stages of senescence, deterioration of mitochondrial oxidative phosphorylation increases production of ROS[66,67]. ROS can maintain and enhance senescence through feedback loops replenishing the DDR[68,69]. Of interest, mitochondrial DNA is highly vulnerable to ROS due to its close proximity to the generation site and damage to mitochondrial DNA further impairs oxidative phosphorylation. Several DNA repair mechanisms exist within a cell to restore DNA integrity and while these pathways have been extensively studied in the nucleus, the current knowledge and evidence for DNA repair pathways located in the mitochondria is more limited[70]. Furthermore, mitochondrial ROS accelerates telomere shortening and triggers senescence in a paracrine fashion[69,71]. *In vitro* data have shown that senescent cells induce considerable metabolic changes on the cellular level, related to mitochondrial metabolites (i.e. decrease in NAD+/NADH ratio and tricarboxylic acid (TCA) cycle metabolites[69,72]. Also, changes in mitochondrial dynamics such as biogenesis, fusion, fission and mitophagy have been described in senescent cells[68,73]. Interestingly, recently mitochondrial dysfunction was shown to induce a distinct type of

senescence termed MiDAS (mitochondrial dysfunction-associated senescence), as a result of a decreased NAD+/NADH ratio[74]. The authors showed that altered AMP/ATP and ADP/ATP ratios activate AMPK which may induce senescence by phosphorylating p53 or stabilizing p16[74].

The above-described hallmarks of cellular senescence are observed in hepatocytes of both humans and rodents with NAFLD. Nevertheless, it has been suggested that hepatocytic senescence is a mere consequence of the metabolic dysregulation and inflammatory phenomena observed in fatty liver disease instead of a causal player. This chicken-egg situation can be clarified in large, prospective studies, which will provide insight in the timeline of disease development linked to the presence of cellular senescence. A causal contribution (*i.e.,* cellular senescence as driving factor for disease development) can only be substantiated from results of highly targeted intervention studies. To successfully identify, characterize and pharmacologically eliminate senescent cells, one of the major limitations of the field needs to be overcome: robust, cell-and pathway-specific biomarkers for cellular senescence should be developed. Driven by the non-specificity of many current senescence markers and the existence of distinct senescence programs, the scientific community has struggled to identify universal and unequivocal signatures characterizing the senescence state[75].

Markers and detection of cellular senescence

Development and optimization of sensitive and specific assays to track senescent cells are challenged by the complex and cell-specific senescence phenotype. Of importance, numerous non-senescent cells, especially proinflammatory cells such as macrophages, precancerous and cancer cells, share features with senescent cells and impair specificity of currently available assays[76]. Hence, only a few markers reliably detect senescent cells at this moment in time and novel (combined) analysis tools are urgently needed to better understand the role of senescent cells in NAFLD.

The first tool to successfully identify senescent cells arose from observations that senescent cells display β-galactosidase enzymatic activity at pH 6, whereas more common β-galactosidase isoforms show peak enzymatic activity at pH 4–4.5. This is referred as senescence-associated beta-galactosidase (SA- β gal)[77,78]. Shortly thereafter, the cyclin-dependent kinase inhibitor p16 that serves as a master regulator of cell cycle arrest, was discovered to play a role in senescence[79]. In the past decade, numerous other senescence markers such as increased cell size and intracellular protein content, accumulation of lipofuscin, increased expression of p21, epigenetic profiles and SASP factors have been identified and linked to distinct senescence pathways[40,75]. An overview of currently used markers of senescence is provided in Table 1.

Unfortunately, these markers have their limitations. For example, SA- β gal activity can be high in macrophages and even p16, which is considered to be one of the most specific senescence markers, is also expressed in certain non-senescent cells[76,80]. Moreover, p16 is not expressed by all senescent cells. To overcome these challenges, a multi-marker approach based on immunohistochemistry and quantitative polymerase chain reaction (qPCR) or transcriptomic analyses was proposed to identify senescent cells[75]. Cells would first be screened for SA- β gal or lipofuscin staining. Initial senescence leads would then be verified by additional markers such as p16 or p21 and further specified into specific types of senescence by characterizing SASP or DDR (**Figure 2**)[75].

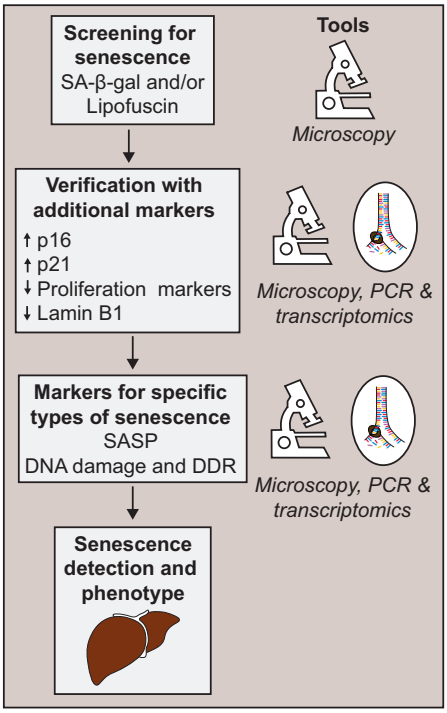


Figure 2. The proposed multi-marker workflow approach[75]. Cells would first be screened for senescence-associated beta-galactosidase (SA- β gal) or lipofuscin staining. Initial senescence leads would then be verified by additional markers such as p16 or p21 and further specified into specific types of senescence by characterizing Senescence Associated Secretory Phenotype (SASP) or DNA damage response (DDR). For the detection of the senescent cells several tools can be used such as immunohistochemistry, quantitative polymerase chain reaction (qPCR) or transcriptomic analyses.

This multi-marker approach enables the detection of senescent cells in various experimental settings and tissues. The first clinical trials targeting cellular senescence in humans are currently performed or all already finished with promising results[81,82]. Therefore, it is of critical importance to develop, implement, test, and harmonize methods

and standard operating protocols (SOPs) for translational early phase trials of agents that target fundamental aging processes. New, effective and low-cost assays are needed to detect and trace senescence in blood, cells and biopsies of the targeted organ for use in clinical trials. Circulating microvesicles originating from senescent cells, senescence-specific micro-RNAs or epigenetic profiles are currently evaluated as novel composite assays to detect senescence[83]. Of interest, recently, Saif et al.[84] were capable to measure the accumulation of lipofuscin in the liver non-invasively using near infrared (NIR) and shortwave infrared (SWIR) autofluorescence. This technique might serve as a diagnostic medical tool or can be used in clinical trials targeting senescence in the liver. Also, there are several studies ongoing aiming to discover novel markers using "omics" techniques to quantify various macromolecules, even at the single cell level to include intrapopulation variability[75].

Senescent cell hallmark	Class	Markers
Cell cycle arrest	Lack of DNA synthesis	BrdU, EdU
	Lack of proliferation	Ki67
	Activation of p16-pRB axis	p16INK4a, pRB, phospho-pRb
	Activation of p53-p21 axis	p21, p53, phospho-p53, DEC1 (BHLHB2), PPP1A
Structural changes	Morphology, cell size	Morphology, cell size
	Increased lysosomal compartment and activity	SA-β-galactosidase, SA-α-Fucosidase, Lipofuscin
	DNA damage	γH2AX, 53BPI, Rad17, ATR, ATM, MDC1, TIF.
	Telomere shortening	Telomeres
	SAHFs formation	DAPI/Hoechst 33342, HIRA, H3K9-methylation, PML bodies, HP1-gamma
	Nuclear membrane	Lamin B1
Pro-survival	Apoptosis exclusion	Annexin V, Cleaved PARP, Cleaved caspase 2/3/9, TUNEL staining
SASP	Chemokines	IL-8; GRO-a, -b, -g; MCP-2; MCP-4; MIP-1a; MIP-3a; HCC-4; eotaxin; eotaxin-3; TECK; ENA-78; I-309; I-TAC
	Growth factors; regulators	Amphiregulin; epiregulin; heregulin; EGF; bFGF; HGF; KGF (FGF7); VEGF; angiogenin; SCF; SDF-1; TGFb; PIGF; NGF; IGFBP-2, -3, -4, -6, -7
	Insoluble factors	Amphiregulin; epiregulin; heregulin; EGF; bFGF; HGF; KGF (FGF7); VEGF; angiogenin; SCF; SDF-1; TGFb; PIGF; NGF; IGFBP-2, -3, -4, -6, -7
	Interleukins	IL-6; IL-7; IL-1; IL-1b; IL-13; IL-15
	Non-protein molecules	PGE2; nitric oxide; ROS
	Other inflammatory molecules	GM-CSE; G-CSE; IFN-g; BLC; MIF
	Proteases and regulators	MMP-1, -3, -10, -12, -13, -14; TIMP-1; TIMP-2; PAI-1, -2; tPA; uPA; cathepsin B
	Receptors; ligands	ICAM-1, -3; OPG; sTNFRI; sTNFRII; TRAIL-R3; Fas; uPAR; SGP130; EGF-R

Table 1. Overview of senescence markers.

Insight obtained from studies in rodents of a role of senescence in NAFLD

Hepatocytic senescence can induce remarkable changes in tissue homeostasis and microenvironment of the liver. Under normal circumstances, hepatocytes are considered to be reversed post-mitotic cells that preserve their proliferating potential. Yet, natural aging correlates with loss of proliferating potential, functionality and thus regenerative capacity of the hepatocyte[85]. Of interest, cellular senescence is considered one of the hallmarks of aging[28] and indeed aging itself is a major risk factor for NAFLD development and progression[44]. Increased oxidative stress and age-related mitochondrial dysfunction have been shown to contribute to NAFLD development in old mice fed a high-fat diet (HFD)[86]. Indeed when young and old mice were treated with the same profibrotic regimen, old mice developed more severe fibrosis in their liver compared to their younger counterparts[87,88]. Also, aging is associated with upregulation of CDK4 in the liver. CDK4 phosphorylates CCAAT- enhancer-binding protein (C/EBPα) which facilitates formation of C/EBPa-p300 complexes leading to NAFLD in the presence of a metabolic driver (*i.e.* overnutrition). *Vice versa*, pharmacological inhibition of CDK4 reduces NAFLD[89,90].

Age-independent hepatocytic senescence has also been described in rodents. In 2000, Rudolph and DePinho revealed that progressive and repetitive liver damage in mice induces hepatocytic cellular senescence and subsequently cirrhosis[91]. Other studies in rodents have shown a clear relation between senescence, NAFLD and liver fibrosis. Obesity-prone rats fed a HFD developed NAFLD after 13 weeks with a concomitant increase in hepatic gene expression of p16 and p21 compared to age-matched lean controls. Subsequently, the increased p16 and p21 resulted in a significant decrease in the phosphorylation of retinoblastoma protein (Rb), thereby inducing cell cycle arrest[92]. P53 deficient mice fed a methionine- and choline-deficient diet (a widely used diet to induce NAFLD in rodents), had slower disease progression compared to wild type mice[93].

Although previous studies have provided insight in the putative role of cellular senescence in the development of NAFLD, data pointing towards causality was only recently published by Ogrodnik et al[34]. First, NAFLD was associated with several markers of senescence in hepatocytes such as increased senescence-associated damage foci, as determined by the presence of yH2AX, increased senescence-associated distention of satellites and larger nuclear areas[34]. Second, hepatocytic senescence was shown to impair hepatic mitochondrial β-oxidation, thereby hindering fatty acid elimination and promoting triglyceride accumulation[34]. Finally, a causal role between hepatocytic senescence and hepatic steatosis was unraveled using INK-ATTACK mice and via the administration of a senolytic drug cocktail. INK-ATTACK transgenic mice (INK-linked apoptosis through targeted activation of caspase) contain an inducible suicide gene in the *CDKNA2* locus, which encodes p16, a key molecule in senescent cells[94]. By using this elegant rodent model, it is possible to selectively

eliminate p16-expressing cells *in vivo* following the administration of a specific molecule that dimerizes the FKBP-CASP8 fusion protein and induces apoptosis. After clearing of p16-expressing cells systemically hepatocytic senescence decreased accompanied by the amelioration of hepatic steatosis. Also the administration of the senolytic drug cocktail Dasatinib and Quercetin, to *db/db* mice reduced the number of senescent cells and clearance of triglycerides from the hepatocytes[34]. However, it is important to note that the authors did not account for the possible pleiotropic effects of the senolytic drug cocktail. Moreover, to what extent these results can be translated to the entire spectrum of NAFLD (*i.e.* hepatocyte ballooning, lobular inflammation and fibrosis) remains to be investigated.

Epigenetic modifications have been observed in senescence cells. These modifications resemble a DNA methylation profile similar to those observed in cancer and aging[95,96]. A global loss of DNA methylation at CpG sites is characteristic of replicative senescence. This loss is followed by a focal increase in DNA methylation at certain CpG islands[96,97]. Senescence-associated DNA methylation patterns have been shown to alter expression of genes typically involved in hepatic lipid metabolism in mice fed a choline- and folate-deficient diet[98]. Alterations in the methylome profile of hepatocytes could therefore determine the severity of NAFLD. Individual histone modifications are also altered during senescence and NAFLD[97]. Certain histone modifications such as elevation of H4K20me3 and H3K9me3 are in fact crucial for induction of proliferation arrest[97,99] whereas elevation of H3K27ac in gene enhancers promotes the SASP[75]. Increased expression of p21 was associated with increased acetylation of both histone H3 and H4 and decreased trimethylation of H3K27 at the p21 promotor was observed[92].

Senescence beyond the hepatocyte

Importantly, senescence in the liver is not limited to hepatocytes. Obesity associated senescence has for example been observed in hepatic stellate cells (HSC)[100]. In healthy livers, HSCs are in quiescent state and are activated following liver injury and play an important role in liver fibrosis. Indeed, liver fibrosis is accompanied with excessive deposition of extracellular matrix by the HSC. Interestingly, when HSC become senescent, they could limit the extent of fibrosis. Indeed, HSC deficient in the key senescence genes p53 or Rb pathways continued to proliferate and contributed to excessive extracellular matrix deposition. These findings suggest that senescence in HSC could be beneficial. Moreover, senescent HSC secrete matrix metalloproteases that digest matrix metalloproteins (MMP's) and collagens such as CollA. In contrast, Yoshimoto and colleagues[100] showed in a mouse model that the gut microbial metabolite deoxycholic acid (DCA), a metabolite that has been associated with insulin resistance and NAFLD provokes the SASP phenotype in HSCs[13,101]. This phenotype subsequently facilitates the development of HCC via the secretion of inflammatory and

tumor promoting factors[100]. Moreover, fibroblasts and non-tumoral HSCs demonstrated increased expression of senescence and SASP markers in NASH-related HCC compared to other chronic liver induced HCC[102]. However, patients with NASH-HCC were significantly older, had a higher BMI and more metabolic diseases such as diabetes compared to patients with other chronic liver induced HCC, thus these data should be interpreted with caution. Therefore, more in-depth studies are needed to understand the balance between fibrogenic and non-fibrogenic SASP in senescent HSC.

Senescence has also been observed in cholangiocytes and may have potential deleterious effects in biliary diseases such as PBC and PSC[103,104]. In addition, cholangiocyte senescence has also been demonstrated in other chronic parenchymal liver diseases including NAFLD[104,105]. Recently, an *in vivo* model was introduced to study the detailed mechanism of cholangiocyte mediated biliary senescence in the liver[106]. The activation of senescence in cholangiocytes induced profound alterations in cellular and signaling microenvironment, resulting in the deposition of collagen and TGFbeta production and the induction of senescence in neighboring cholangiocytes and hepatocytes[106].

Liver sinusoid endothelial cells (LSECs) are fenestrated endothelial cells that form the lining of the hepatic sinusoids. The structure and function of LSEC's change upon aging, which in turn impact liver functions. Age induced morphological changes in LSECs are characterized by defenestration (defined by the decrease in the number and size of fenestrae), endothelial thickening, and basal lamina and collagen deposition in LSECs and have been described in rodents as well as in humans. Of recent, it was shown that senescence markers increase in older mice, followed by an enhanced ability to clear macromolecular waste[107]. However, this enhanced ability rapidly declines with further aging probably due to increased endothelial thickness and senescence induced silencing of scavenger receptors and endocytosis genes. Of importance, age-dependent changes in LSEC's were recently confirmed in liver from aged humans underscoring that aging and senescence is accompanied by significant liver sinusoidal dysregulation[108].

Kupffer cells (KC), the resident macrophages of the liver are located within the lumen of the liver sinusoids. KC's are the key detector of commensal or pathogenic microbial signals, danger signals, and tumor cells moving through the hepatic circulation and produce soluble cell mediators such as *TNFα* and IL-6 as part of the innate immune response. While there have been many studies on the effects of aging and senescence on macrophages, the effect on KC's has not been well characterized on a cellular and molecular level. The effects of aging on macrophages include reduced phagocytosis and autophagy and increased production of cytokines such as IL-6, suggesting that KCs might be one source of elevated IL-6 that is characteristic of old age[108].

Although the abovementioned studies imply that senescence is variously involved in NAFLD pathogenesis and progression, it is of interest whether data from animal studies can be translated to humans. In the next paragraph, we will elaborate on the evidence of a role of cellular senescence in NAFLD in humans.

Clinical evidence linking ageing and senescence processes to NAFLD

Under normal circumstance, around 3-7% of hepatocytes are senescent. This percentage can increase to 50-100% in end stage liver disease[109,110]. As mentioned above, senescence in the liver can have protective as well as deleterious effects on liver function and metabolism. Interestingly, hepatocytic senescence is considered to act as a protective mechanism against the development of hepatocellular carcinoma (HCC)[111]. Data obtained from studies in humans revealed that several hallmarks of cellular senescence are present in biopsies of individuals with NAFLD and that the number of senescent cells increases with disease progression. Relative nuclear size of hepatocytes in individuals with NAFLD was significantly larger than the normal value of healthy controls, independent of telomere length[112]. Interestingly, telomere length correlated negatively with nuclear size in both individuals with NAFLD as well as in healthy controls, while the average nuclear size of the hepatocyte correlated with age only in the healthy controls. This suggests that nuclear enlargement proceeds independently of age in individuals with NAFLD. Other studies also showed that average telomere length in livers of individuals with NAFLD is shorter compared to age-matched healthy controls[112,113]. Moreover, in a longitudinal study of six years, it was shown that individuals who developed NAFLD had shorter telomeres in peripheral blood leukocytes at the end of the follow-up period compared to the individuals who did not develop NAFLD[114]. Despite this observation, individuals who developed NAFLD were metabolically already more challenged compared to the individuals without NAFLD. Also, Laish et al.[115] observed shorter telomeres in peripheral lymphocytes accompanied with a higher expression of telomerase reverse transcriptase messenger RNA compared to healthy controls. To what extent the telomere length in peripheral blood corresponds to telomere length in liver cells remain to be investigated. Nevertheless, these results support a role for telomere dysfunction in the development of NAFLD.

Bearing this evidence in mind, one might ask the question: how do telomeres signal senescence? It has been hypothesized that a protein complex that shapes and safeguards human telomeres, also called the “shelterin”, destabilizes with each cell division[116,117]. This destabilization results in exposure of the telomere, which is subsequently recognized as a double-strand DNA break. This triggers recruitment of proteins belonging to the DDR such as ATM and H2A.X and Rad 17[118]. As mentioned above, the DDR activates transcription factors such as p53, which is a positive regulator of p21. Both *in vivo* and *in vitro* studies

have shown that p21 plays a key role in telomere-induced senescence[47,119]. Although the majority of the studies provide evidence that senescence is a result of telomere shortening, several other reports now suggest that telomere dysfunction can also occur in a length-independent manner[47,49,120,121]. For example, chronic mild inflammation is able to induce telomere gene damage in hepatocytes and enterocytes of the small intestine, irrespectively of telomere length[120,121]. Moreover, it has been suggested that DNA damage is more likely to occur at long telomeres as they represent a more abundant target for lesion formation which can explain length-independent DDR activation[49]. Interestingly, the link between telomere-induced senescence and the p16 pathway is less clear and the link between p16 expression and NAFLD in humans is not that robust compared to p21[33].

Several studies revealed a link between DNA damage in hepatocytes, hepatocytic senescence and NAFLD[33,113] Aravinthan and colleagues[113], first showed that there is increased DNA damage in livers of individuals with NAFLD which increased with disease progression (*i.e.* NASH and NAFLD with advanced fibrosis). Moreover, by using paired biopsies from 35 individuals, hepatocyte p21 expression was shown to increase with disease progression, whereas subjects with disease improvement, had decreased expression of hepatocyte p21. Thus, hepatocytic senescence is a marker for disease progression. Also, hepatocytic senescence was positively correlated with progression of liver fibrosis[110].

Recently, a prediction model based on epigenetic DNA methylation was introduced to measure human chronological and biological age. Using this model, it is possible to predict normal aging rate based on methylation patterns[122–125]. One of the algorithms for this model is the so-called Horvath Clock, which represents an epigenetic profile comprising methylation levels of 353 CpG dinucleotide sites[123]. Of these 353 CpG sites, 193 positively correlate with age when hypermethylated whereas 160 negatively correlate with age when hypomethylated. To illustrate the time frame of the Horvath clock, the DNA methylation score of embryonic stem cells is approximately zero and increases rapidly during normal development. The validity of peripheral DNA methylation to accurately predict chronological age of different tissues including the liver has been confirmed in multiple studies[123,126,127]. Moreover, the intrinsic rate of the epigenetic clock can be altered by metabolic diseases. Obesity for example is able to alter the epigenetic clock for the liver, but not of other tissues[127]. Recently, Loomba and colleagues have shown by using the Horvath Clock that individuals with NASH demonstrate significant acceleration in their biological age[126]. An enrichment analyses of the genes associated with differentially methylated CpG islands revealed significant enrichment of senescence pathways such as p53 signaling, suggesting that, in line with other reports, a specific pattern of DNA methylation is another senescence marker associated with NAFLD and its progression to NASH[126,128]. Murphy et al showed that individuals with NAFLD with mild fibrosis could be distinguished from

individuals with NAFLD with advanced fibrosis based on different methylation patterns[129]. Individuals with advanced fibrosis had more hypomethylated genes in their liver biopsies resulting in overexpression of tissue repair genes, whereas metabolism associated genes were hypermethylated resulting in down regulation of these genes. Another study found that alterations in methylation patterns in genes involved in the cell cycle are closely related to oxidative DNA damage in the liver of individuals with NAFLD[130]. Collectively, these data indicate that NAFLD may induce altered methylation profiles in a plethora of cells, apart from the hepatocytes, even in peripheral blood cells. However, in this view, senescence is a consequence of the metabolic dysregulation and inflammatory phenomena occurring within the liver instead of a causal player.

Treatment options for targeting senescence

Usually, NAFLD is accompanied by other obesity-induced age-related diseases. This inevitably leads to polypharmacy because most treatment strategies are disease specific. Unfortunately, polypharmacy can lead to adverse events, unpredictable drugs interactions and poor adherence[83]. Bearing in mind that senescent cells are present in several metabolic diseases, targeting senescence cells has therefore emerged as an attractive therapeutic strategy to simultaneously treat these diseases.

Targeting senescence could be performed via inhibiting the SASP or by selectively eliminating senescent cells using senolytics. SASP inhibitors, also known as senomorphics, target signaling pathways that are involved in the regulation or exacerbation of the SASP such as target of rapamycin complex 1 (mTORC1), JAK1/JAK2, STAT3, and mitochondrial dysfunction[83]. Senomorphics include rapamycin, ruxolitinib, glucocorticoids and metformin[131–133]. However, most of the senomorphics do not reduce the entire range of SASP factors and have many other effects in addition to being a senomorphic. Disentangling effects on age-related phenotypes due to SASP modulation from other off-target age related processes is therefore challenging. In addition, senomorphics would need to be administrated continuously to maintain SASP suppression, which limit the applicability.

Senolytic agents were first discovered in 2015. Although already in 1995 it was discovered that senescent cells are resistant to apoptosis[134], the authors hypothesized that senescent cells depend on senescence-associated antiapoptotic pathways (SCAPs), which permit senescent cells to survive their own SASP[135]. Using a combination of bioinformatic tools and *in vitro* RNA interference studies, it was verified that senescent cells rely on the SCAPs. Hence, SCAPs were identified as the Achilles heel of senescent cells. Since this discovery, considerable progress has been made in identifying small molecules, peptides and antibodies that selectively induce apoptosis in senescent cells. The combination of Dasatinib, which is an FDA-approved tyrosine kinase inhibitor and the antioxidant quercitin, which is a flavonol

present in many fruits and vegetables, successfully induce apoptosis in senescent cells *in vitro* and rodent models[134]. According to *in vitro* data, a brief disruption of pro-survival pathways is adequate to[83,136] suggest that senolytics could be administrated intermittently, which reduced the risks of adverse effects compared to continuous treatment.

As mentioned previously, administration of Dasatinib plus quercitin successfully eliminated both adipocyte and hepatocytic senescence and decreased lipid accumulation[34]. So far, no studies in individuals with NAFLD using Dasatinib plus quercitin have been performed. However, several clinical trials addressing the efficacy of Dasatinib plus quercetin to treat metabolic diseases are currently ongoing[137]. Combining results from two human clinical trials using Dasatinib plus quercitin for the treatment of diabetic chronic kidney disease and idiopathic pulmonary fibrosis -two diseases characterized by accumulation of senescent cells- revealed that eliminating senescent cells, which was shown by analyzing adipose tissue, improved clinical outcomes[138,139]. These clinical investigations have proven that the risks of using Dasatinib plus quercitin in combination were minimal in relation to the clinical benefits. However, several senolytic drugs, including Dasatinib, have been used for cancer treatment and often exhibit adverse effects such as nausea, vomiting, diarrhea and skin rashes when taken continuously. Moreover, senolytics also have other effects. For example, the antioxidant quercetin inhibits ferroptosis and can decrease inflammation and lipid metabolism, all pathways that are associated with NAFLD[81,82]. Although it has been suggested that senolytics could be administrated intermittently, thereby reducing the risks of adverse events, large clinical trials are needed to define the benefits and potential risks of these drugs.

CONCLUSIONS

Data from studies in rodents and humans have shown that NAFLD is accompanied by an increase in senescent cells in the liver, and that the number of senescent cells is associated with a more advanced disease state . Despite the strong associations between senescence and NAFLD in humans and the work derived from *in vitro* studies and rodents, it remains to be investigated if hepatic senescence is a mere consequence of the metabolic dysregulation and inflammatory phenomena in NAFLD or a causal player in the development of this disease. Although a causal role of cellular senescence has to be further substantiated first and subsequently established in humans, this pathophysiological process holds great potential bearing in mind that there is currently no effective treatment for NAFLD. Targeting senescence has emerged as an attractive therapeutic target for NAFLD since senescence might be involved in the full spectrum of the disease (*i.e.* from early steatosis

to liver cirrhosis). Moreover, senolytic drugs can be administrated intermittently, thereby minimizing potential toxic effects and increasing adherence in the individual often affected by multiple morbidities and thus treated with multiple medications. Nevertheless, clinical trials conducted in individuals with NAFLD using senolytics have not been performed. Such trials are needed to better define the benefits and potential risks of these drugs. To increase efficacy and accuracy of these clinical trials, new or composite assays are needed, and development of these assays should be a top priority for the field.

REFERENCES

1. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
2. Tilg, H., Adolph, T. E. & Moschen, A. R. Multiple Parallel Hits Hypothesis in Nonalcoholic Fatty Liver Disease: Revisited After a Decade. *Hepatology* Preprint at https://doi.org/10.1002/hep.31518 (2020).
3. Ekstedt, M., Franzén, L. E., Mathiesen, U. L. & Kechagias, S. Low clinical relevance of the nonalcoholic fatty liver disease activity score (NAS) in predicting fibrosis progression. *Scand J Gastroenterol* (2012) doi:10.3109/00365521.2011.634024.
4. Ekstedt, M. *et al.* Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* (2015) doi:10.1002/hep.27368.
5. Pais, R. *et al.* A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. *Journal of Hepatology* Preprint at https://doi.org/10.1016/j.jhep.2013.04.027 (2013).
6. McPherson, S. *et al.* Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: Implications for prognosis and clinical management. *J Hepatol* (2015) doi:10.1016/j.jhep.2014.11.034.
7. Farrell, G. C. & Larter, C. Z. Nonalcoholic fatty liver disease: From steatosis to cirrhosis. *Hepatology* **43**, S99–S112 (2006).
8. Younossi, Z. *et al.* Global Perspectives on Non-alcoholic Fatty Liver Disease and Non-alcoholic Steatohepatitis. *Hepatology* (2018) doi:10.1002/hep.30251.
9. Buzzetti, E., Pinzani, M. & Tsochatzis, E. A. The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism* (2016) doi:10.1016/j.metabol.2015.12.012.
10. Cusi, K. Nonalcoholic steatohepatitis in nonobese patients: Not so different after all. *Hepatology* Preprint at https://doi.org/10.1002/hep.28839 (2017).
11. Marjot, T., Moolla, A., Cobbold, J. F., Hodson, L. & Tomlinson, J. W. Nonalcoholic Fatty Liver Disease in Adults: Current Concepts in Etiology, Outcomes, and Management. *Endocr Rev* (2020) doi:10.1210/endrev/bnz009.
12. Eslam, M., Valenti, L. & Romeo, S. Genetics and epigenetics of NAFLD and NASH: Clinical impact. *J Hepatol* **68**, 268–279 (2018).
13. Arab, J. P., Karpen, S. J., Dawson, P. A., Arrese, M. & Trauner, M. Bile Acids and Nonalcoholic Fatty Liver Disease: Molecular Insights and Therapeutic Perspectives. (2016) doi:10.1002/hep.28709.
14. Tripathi, A. *et al.* The gut-liver axis and the intersection with the microbiome. *Nature Reviews Gastroenterology and Hepatology* Preprint at https://doi.org/10.1038/s41575-018-0011-z (2018).
15. Samuel, V. T. & Shulman, G. I. Nonalcoholic Fatty Liver Disease as a Nexus of Metabolic and Hepatic Diseases. *Cell Metab* **27**, 22–41 (2018).
16. Donnelly, K. L. *et al.* Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *Journal of Clinical Investigation* (2005) doi:10.1172/jci200523621.
17. Maeda Júnior, A. S. *et al.* Cafeteria diet feeding in young rats leads to hepatic steatosis and increased gluconeogenesis under fatty acids and glucagon influence. *Nutrients* **10**, (2018).
18. Hafizi Abu Bakar, M. *et al.* Mitochondrial dysfunction as a central event for mechanisms underlying insulin resistance: The roles of long chain fatty acids. *Diabetes/Metabolism Research and Reviews* Preprint at https://doi.org/10.1002/dmrr.2601 (2015).
19. Musso, G., Cassader, M., Paschetta, E. & Gambino, R. Bioactive Lipid Species and Metabolic Pathways in Progression and Resolution of Nonalcoholic Steatohepatitis. *Gastroenterology* vol. 155 282-302.e8 Preprint at https://doi.org/10.1053/j.gastro.2018.06.031 (2018).
20. Brown, M. S. & Goldstein, J. L. Selective versus Total Insulin Resistance: A Pathogenic Paradox. *Cell Metab* **7**, 95–96 (2008).
21. Honma, M. *et al.* Selective insulin resistance with differential expressions of IRS-1 and IRS-2 in human NAFLD livers. *Int J Obes* (2018) doi:10.1038/s41366-018-0062-9.
22. Besse-Patin, A. *et al.* PGC1A regulates the IRS1:IRS2 ratio during fasting to influence hepatic metabolism downstream of insulin. *Proc Natl Acad Sci U S A* (2019) doi:10.1073/pnas.1815150116.

23. Lebeaupin, C. *et al.* Endoplasmic reticulum stress signalling and the pathogenesis of non-alcoholic fatty liver disease. *Journal of Hepatology* vol. 69 927–947 Preprint at https://doi.org/10.1016/j.jhep.2018.06.008 (2018).

24. Schattenberg, J. M. *et al.* JNK1 but not JNK2 promotes the development of steatohepatitis in mice. *Hepatology* (2006) doi:10.1002/hep.20999.

25. Pikarsky, E. *et al.* NF-κB functions as a tumour promoter in inflammation-associated cancer. *Nature* (2004) doi:10.1038/nature02924.

26. Fernandez-Real, J.-M. Circulating Interleukin 6 Levels, Blood Pressure, and Insulin Sensitivity in Apparently Healthy Men and Women. *Journal of Clinical Endocrinology & Metabolism* (2001) doi:10.1210/jc.86.3.1154.

27. Du Plessis, J. *et al.* Association of Adipose Tissue Inflammation with Histologic Severity of Nonalcoholic Fatty Liver Disease. *Gastroenterology* **149**, 635e14-648.e14 (2015).

28. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* Preprint at https://doi.org/10.1016/j.cell.2013.05.039 (2013).

29. Ogrodnik, M., Salmonowicz, H. & Gladyshev, V. N. Integrating cellular senescence with the concept of damage accumulation in aging: Relevance for clearance of senescent cells. *Aging Cell* Preprint at https://doi.org/10.1111/acel.12841 (2019).

30. Minamino, T. *et al.* A crucial role for adipose tissue p53 in the regulation of insulin resistance. *Nat Med* (2009) doi:10.1038/nm.2014.

31. Palmer, A. K. *et al.* Targeting senescent cells alleviates obesity-induced metabolic dysfunction. *Aging Cell* (2019) doi:10.1111/acel.12950.

32. Gustafson, B., Nerstedt, A. & Smith, U. Reduced subcutaneous adipogenesis in human hypertrophic obesity is linked to senescent precursor cells. *Nat Commun* (2019) doi:10.1038/s41467-019-10688-x.

33. Papatheodoridi, A.-M., Chrysavgis, L., Koutsilieris, M. & Chatzigeorgiou, A. The role of senescence in the development of non-alcoholic fatty liver disease and progression to non-alcoholic steatohepatitis. *Hepatology* (2019) doi:10.1002/hep.30834.

34. Ogrodnik, M. *et al.* Cellular senescence drives age-dependent hepatic steatosis. *Nat Commun* (2017) doi:10.1038/ncomms15691.

35. Helman, A. *et al.* p16 Ink4a-induced senescence of pancreatic beta cells enhances insulin secretion. *Nat Med* (2016) doi:10.1038/nm.4054.

36. Van Deursen, J. M. The role of senescent cells in ageing. *Nature* Preprint at https://doi.org/10.1038/nature13193 (2014).

37. Hayflick, L. The limited in vitro lifetime of human diploid cell strains. *Exp Cell Res* (1965) doi:10.1016/0014-4827(65)90211-9.

38. Gorgoulis, V. *et al.* Cellular Senescence: Defining a Path Forward. *Cell* Preprint at https://doi.org/10.1016/j.cell.2019.10.005 (2019).

39. Tchkonia, T., Zhu, Y., Van Deursen, J., Campisi, J. & Kirkland, J. L. Cellular senescence and the senescent secretory phenotype: Therapeutic opportunities. *Journal of Clinical Investigation* Preprint at https://doi.org/10.1172/JCI64098 (2013).

40. Coppé, J.-P., Desprez, P.-Y., Krtolica, A. & Campisi, J. The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression. *Annual Review of Pathology: Mechanisms of Disease* (2010) doi:10.1146/annurev-pathol-121808-102144.

41. Krizhanovsky, V. *et al.* Senescence of Activated Stellate Cells Limits Liver Fibrosis. *Cell* (2008) doi:10.1016/j.cell.2008.06.049.

42. Bird, T. G. *et al.* TGFβ inhibition restores a regenerative response in acute liver injury by suppressing paracrine senescence. *Sci Transl Med* (2018) doi:10.1126/scitranslmed.aan1230.

43. Ritschka, B. *et al.* The senotherapeutic drug ABT-737 disrupts aberrant p21 expression to restore liver regeneration in adult mice. *Genes Dev* (2020) doi:10.1101/gad.332643.119.

44. Aravinthan, A. Cellular senescence: a hitchhiker's guide. *Hum Cell* (2015) doi:10.1007/s13577-015-0110-x.

45. Childs, B. G., Durik, M., Baker, D. J. & Van Deursen, J. M. Cellular senescence in aging and age-related disease: From mechanisms to therapy. *Nature Medicine* Preprint at https://doi.org/10.1038/nm.4000 (2015).

46. Campisi, J. & D'Adda Di Fagagna, F. Cellular senescence: When bad things happen to good cells. *Nature Reviews Molecular Cell Biology* Preprint at https://doi.org/10.1038/nrm2233 (2007).

47. Victorelli, S. & Passos, J. F. Telomeres and Cell Senescence - Size Matters Not. *EBioMedicine* Preprint at https://doi.org/10.1016/j.ebiom.2017.03.027 (2017).

48. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* Preprint at https://doi.org/10.1038/nature08467 (2009).

49. Fumagalli, M. *et al.* Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation. *Nat Cell Biol* (2012) doi:10.1038/ncb2466.

50. Dulić, V., Beney, G.-E., Frebourg, G., Drullinger, L. F. & Stein, G. H. Uncoupling between Phenotypic Senescence and Cell Cycle Arrest in Aging p21-Deficient Fibroblasts. *Mol Cell Biol* (2000) doi:10.1128/mcb.20.18.6741-6754.2000.

51. Rubin, S. M., Gall, A. L., Zheng, N. & Pavletich, N. P. Structure of the Rb C-terminal domain bound to E2F1-DP1: A mechanism for phosphorylation-induced E2F release. *Cell* (2005) doi:10.1016/j.cell.2005.09.044.

52. Acosta, J. C. *et al.* Chemokine Signaling via the CXCR2 Receptor Reinforces Senescence. *Cell* (2008) doi:10.1016/j.cell.2008.03.038.

53. Wajapeyee, N., Serra, R. W., Zhu, X., Mahalingam, M. & Green, M. R. Oncogenic BRAF Induces Senescence and Apoptosis through Pathways Mediated by the Secreted Protein IGFBP7. *Cell* (2008) doi:10.1016/j.cell.2007.12.032.

54. Childs, B. G., Baker, D. J., Kirkland, J. L., Campisi, J. & Deursen, J. M. Senescence and apoptosis: dueling or complementary cell fates? *EMBO Rep* (2014) doi:10.15252/embr.201439245.

55. Palmer, A. K., Gustafson, B., Kirkland, J. L. & Smith, U. Cellular senescence: at the nexus between ageing and diabetes. *Diabetologia* (2019) doi:10.1007/s00125-019-4934-x.

56. Franceschi, C. & Campisi, J. Chronic inflammation (Inflammaging) and its potential contribution to age-associated diseases. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* Preprint at https://doi.org/10.1093/gerona/glu057 (2014).

57. Freund, A., Patil, C. K. & Campisi, J. P38MAPK is a novel DNA damage response-independent regulator of the senescence-associated secretory phenotype. *EMBO Journal* (2011) doi:10.1038/emboj.2011.69.

58. Kang, C. *et al.* The DNA damage response induces inflammation and senescence by inhibiting autophagy of GATA4. *Science (1979)* (2015) doi:10.1126/science.aaa5612.

59. Ito, Y., Hoare, M. & Narita, M. Spatial and Temporal Control of Senescence. *Trends in Cell Biology* Preprint at https://doi.org/10.1016/j.tcb.2017.07.004 (2017).

60. Allaire, M., Rautou, P. E., Codogno, P. & Lotersztajn, S. Autophagy in liver diseases: Time for translation? *Journal of Hepatology* Preprint at https://doi.org/10.1016/j.jhep.2019.01.026 (2019).

61. Zhang, Z. *et al.* Interaction between autophagy and senescence is required for dihydroartemisinin to alleviate liver fibrosis. *Cell Death Dis* (2017) doi:10.1038/cddis.2017.255.

62. Rodier, F. *et al.* DNA-SCARS: Distinct nuclear structures that sustain damage-induced senescence growth arrest and inflammatory cytokine secretion. *J Cell Sci* (2011) doi:10.1242/jcs.071340.

63. Hoare, M. *et al.* NOTCH1 mediates a switch between two distinct secretomes during senescence. *Nat Cell Biol* (2016) doi:10.1038/ncb3397.

64. Orjalo, A. V., Bhaumik, D., Gengler, B. K., Scott, G. K. & Campisi, J. Cell surface-bound IL-1α is an upstream regulator of the senescence-associated IL-6/IL-8 cytokine network. *Proc Natl Acad Sci U S A* (2009) doi:10.1073/pnas.0905299106.

65. Birch, J. & Passos, J. F. Targeting the SASP to combat ageing: Mitochondria as possible intracellular allies? *BioEssays* (2017) doi:10.1002/bies.201600235.

66. Yoon, G. *et al.* Iron chelation-induced senescence-like growth arrest in hepatocyte cell lines: Association of transforming growth factor β1 (TGF-β1)-mediated p27Kip1 expression. *Biochemical Journal* (2002) doi:10.1042/BJ20011445.

67. Yoon, Y. S., Byun, H. O., Cho, H., Kim, B. K. & Yoon, G. Complex II Defect via Down-regulation of Iron-Sulfur Subunit Induces Mitochondrial Dysfunction and Cell Cycle Delay in Iron Chelation-induced Senescence-associated Growth Arrest. *Journal of Biological Chemistry* (2003) doi:10.1074/jbc.M308489200.

68. Passos, J. F. *et al.* Feedback between p21 and reactive oxygen production is necessary for cell senescence. *Mol Syst Biol* (2010) doi:10.1038/msb.2010.5.

69. Passos, J. F. *et al.* Mitochondrial dysfunction accounts for the stochastic heterogeneity in telomere-dependent senescence. *PLoS Biol* (2007) doi:10.1371/journal.pbio.0050110.

70. Saki, M. & Prakash, A. DNA damage related crosstalk between the nucleus and mitochondria. *Free Radical Biology and Medicine* Preprint at https://doi.org/10.1016/j.freeradbiomed.2016.11.050 (2017).

71. Petersen, S., Saretzki, G. & Von Zglinicki, T. Preferential accumulation of single-stranded regions in telomeres of human fibroblasts. *Exp Cell Res* (1998) doi:10.1006/excr.1997.3893.

72. Kaplon, J. *et al.* A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature* (2013) doi:10.1038/nature12154.

73. Tai, H. *et al.* Autophagy impairment with lysosomal and mitochondrial dysfunction is an important characteristic of oxidative stress-induced senescence. *Autophagy* (2017) doi:10.1080/15548627.2016.1247143.

74. Wiley, C. D. *et al.* Mitochondrial dysfunction induces senescence with a distinct secretory phenotype. *Cell Metab* (2016) doi:10.1016/j.cmet.2015.11.011.

75. Hernandez-Segura, A., Nehme, J. & Demaria, M. Hallmarks of Cellular Senescence. *Trends in Cell Biology* Preprint at https://doi.org/10.1016/j.tcb.2018.02.001 (2018).

76. Hall, B. M. *et al.* Aging of mice is associated with p16(Ink4a)- and β-galactosidasepositive macrophage accumulation that can be induced in young mice by senescent cells. *Aging* (2016) doi:10.18632/aging.100991.

77. Kurz, D. J., Decary, S., Hong, Y. & Erusalimsky, J. D. Senescence-associated β-galactosidase reflects an increase in lysosomal mass during replicative ageing of human endothelial cells. *J Cell Sci* (2000).

78. Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA-βgal) activity, a biomarker of senescent cells in culture and in vivo. *Nat Protoc* (2009) doi:10.1038/nprot.2009.191.

79. Alcorta, D. A. *et al.* Involvement of the cyclin-dependent kinase inhibitor p16 (INK4a) in replicative senescence of normal human fibroblasts. *Proc Natl Acad Sci U S A* **93**, 13742–13747 (1996).

80. Sharpless, N. E. & Sherr, C. J. Forging a signature of in vivo senescence (Nature Reviews Cancer (2015) 15:(397-408)). *Nature Reviews Cancer* Preprint at https://doi.org/10.1038/nrc3988 (2015).

81. Wang, Z. X. *et al.* Quercetin induces p53-independent cancer cell death through lysosome activation by the transcription factor EB and Reactive Oxygen Species-dependent ferroptosis. *Br J Pharmacol* (2020) doi:10.1111/bph.15350.

82. Yang, H. *et al.* Quercetin improves nonalcoholic fatty liver by ameliorating inflammation, oxidative stress, and lipid metabolism in db/db mice. *Phytotherapy Research* (2019) doi:10.1002/ptr.6486.

83. Khosla, S., Farr, J. N., Tchkonia, T. & Kirkland, J. L. The role of cellular senescence in ageing and endocrine disease. *Nature Reviews Endocrinology* Preprint at https://doi.org/10.1038/s41574-020-0335-y (2020).

84. Saif, M. *et al.* Non-invasive monitoring of chronic liver disease via near-infrared and shortwave-infrared imaging of endogenous lipofuscin. *Nat Biomed Eng* (2020) doi:10.1038/s41551-020-0569-y.

85. Cieslak, K. P., Baur, O., Verheij, J., Bennink, R. J. & van Gulik, T. M. Liver function declines with increased age. *HPB* (2016) doi:10.1016/j.hpb.2016.05.011.

86. Lohr, K. *et al.* Reduced mitochondrial mass and function add to age-related susceptibility toward diet-induced fatty liver in C57BL/6J mice. *Physiol Rep* (2016) doi:10.14814/phy2.12988.

87. Collins, B. H. *et al.* Association of age-dependent liver injury and fibrosis with immune cell populations. *Liver International* (2013) doi:10.1111/liv.12202.

88. Delire, B. *et al.* Aging enhances liver fibrotic response in mice through hampering extracellular matrix remodeling. *Aging* (2017) doi:10.18632/aging.101124.

89. Nguyen, P. *et al.* Elimination of Age-Associated Hepatic Steatosis and Correction of Aging Phenotype by Inhibition of cdk4-C/EBPα-p300 Axis. *Cell Rep* (2018) doi:10.1016/j.celrep.2018.07.014.

90. Jin, J. *et al.* Activation of CDK4 Triggers Development of Non-alcoholic Fatty Liver Disease. *Cell Rep* (2016) doi:10.1016/j.celrep.2016.06.019.

91. Rudolph, K. L., Chang, S., Millard, M., Schreiber-Agus, N. & DePinho, R. A. Inhibition of experimental liver cirrhosis in mice by telomerase gene delivery. *Science (1979)* (2000) doi:10.1126/science.287.5456.1253.

92. Zhang, X., Zhou, D., Strakovsky, R., Zhang, Y. & Pan, Y. X. Hepatic cellular senescence pathway genes are induced through histone modifications in a diet-induced obese rat model. *Am J Physiol Gastrointest Liver Physiol* (2012) doi:10.1152/ajpgi.00032.2011.

93. Tomita, K. *et al.* P53/p66Shc-mediated signaling contributes to the progression of non-alcoholic steatohepatitis in humans and mice. *J Hepatol* (2012) doi:10.1016/j.jhep.2012.05.013.

94. Childs, B. G. *et al.* Senescent cells: An emerging target for diseases of ageing. *Nature Reviews Drug Discovery* Preprint at https://doi.org/10.1038/nrd.2017.116 (2017).

95. Xie, W. *et al.* DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk. *Cancer Cell* (2018) doi:10.1016/j.ccell.2018.01.008.

96. Cruickshanks, H. A. *et al.* Senescent Cells Harbour Features of the Cancer Epigenome. *Nat Cell Biol* (2013) doi:10.1038/ncb2879.

97. Cheng, L. Q., Zhang, Z. Q., Chen, H. Z. & Liu, D. P. Epigenetic regulation in cell senescence. *Journal of Molecular Medicine* Preprint at https://doi.org/10.1007/s00109-017-1581-x (2017).

98. Tryndyak, V. P. *et al.* Status of hepatic DNA methylome predetermines and modulates the severity of non-alcoholic fatty liver injury in mice. *BMC Genomics* (2016) doi:10.1186/s12864-016-2617-2.

99. Salama, R., Sadaie, M., Hoare, M. & Narita, M. Cellular senescence and its effector programs. *Genes and Development* Preprint at https://doi.org/10.1101/gad.235184.113 (2014).

100. Yoshimoto, S. *et al.* Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97–101 (2013).

101. Cariou, B. *et al.* Fasting plasma chenodeoxycholic acid and cholic acid concentrations are inversely correlated with insulin sensitivity in adults. *Nutr Metab (Lond)* **8**, (2011).

102. Lee, J. S. *et al.* Tumor stroma with senescence-Associated secretory phenotype in steatohepatitic hepatocellular carcinoma. *PLoS One* (2017) doi:10.1371/journal.pone.0171922.

103. Harada, K. *et al.* Increased expression of WAF1 in intrahepatic bile ducts in primary biliary cirrhosis relates to apoptosis. *J Hepatol* (2001) doi:10.1016/S0168-8278(00)00075-1.

104. Sasaki, M. *et al.* Bile ductular cells undergoing cellular senescence increase in chronic liver diseases along with fibrous progression. *Am J Clin Pathol* (2010) doi:10.1309/AJCPWMX47TREYWZG.

105. Aravinthan, A. D. & Alexander, G. J. M. Senescence in chronic liver disease: Is the future in aging? *Journal of Hepatology* Preprint at https://doi.org/10.1016/j.jhep.2016.05.030 (2016).

106. Ferreira-Gonzalez, S. *et al.* Paracrine cellular senescence exacerbates biliary injury and impairs regeneration. *Nat Commun* (2018) doi:10.1038/s41467-018-03299-5.

107. Grosse, L. & Bulavin, D. V. LSEC model of aging. *Aging* (2020) doi:10.18632/aging.103492.

108. Maeso-Díaz, R. *et al.* Effects of aging on liver microcirculatory function and sinusoidal phenotype. *Aging Cell* (2018) doi:10.1111/acel.12829.

109. Paradis, V. *et al.* Replicative senescence in normal liver, chronic hepatitis C, and hepatocellular carcinomas. *Hum Pathol* (2001) doi:10.1053/hupa.2001.22747.

110. WIEMANN, S. U. *et al.* Hepatocyte telomere shortening and senescence are general markers of human liver cirrhosis. *The FASEB Journal* (2002) doi:10.1096/fj.01-0977com.

111. Xue, W. *et al.* Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* (2007) doi:10.1038/nature05529.

112. Nakajima, T. *et al.* Nuclear size measurement is a simple method for the assessment of hepatocellular aging in non-alcoholic fatty liver disease: Comparison with telomere-specific quantitative FISH and p21 immunohistochemistry. *Pathol Int* (2010) doi:10.1111/j.1440-1827.2009.02504.x.

113. Aravinthan, A. *et al.* Hepatocyte senescence predicts progression in non-alcohol-related fatty liver disease. *J Hepatol* (2013) doi:10.1016/j.jhep.2012.10.031.

114. Ping, F. *et al.* Deoxyribonucleic acid telomere length shortening can predict the incidence of non-alcoholic fatty liver disease in patients with type 2 diabetes mellitus. *J Diabetes Investig* (2017) doi:10.1111/jdi.12555.

115. Laish, I. *et al.* Telomere Dysfunction in Nonalcoholic Fatty Liver Disease and Cryptogenic Cirrhosis. *Cytogenet Genome Res* (2017) doi:10.1159/000454654.

116. De Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development* Preprint at https://doi.org/10.1101/gad.1346005 (2005).

117. Griffith, J. D. *et al.* Mammalian telomeres end in a large duplex loop. *Cell* (1999) doi:10.1016/S0092-8674(00)80760-6.

118. Takai, H., Smogorzewska, A. & De Lange, T. DNA damage foci at dysfunctional telomeres. *Current Biology* (2003) doi:10.1016/S0960-9822(03)00542-6.

119. Choudhury, A. R. *et al.* Cdkn1a deletion improves stem cell function and lifespan of mice with dysfunctional telomeres without accelerating cancer formation. *Nat Genet* (2007) doi:10.1038/ng1937.

120. Jurk, D. *et al.* Chronic inflammation induces telomere dysfunction and accelerates ageing in mice. *Nat Commun* (2014) doi:10.1038/ncomms5172.

121. Hewitt, G. *et al.* Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence. *Nat Commun* (2012) doi:10.1038/ncomms1708.

122. Hannum, G. *et al.* Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol Cell* (2013) doi:10.1016/j.molcel.2012.10.016.

123. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* (2013) doi:10.1186/gb-2013-14-10-r115.

124. Koch, C. M. & Wagner, W. Epigenetic-aging-signature to determine age in different tissues. *Aging* (2011) doi:10.18632/aging.100395.

125. Rakyan, V. K. *et al.* Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* (2010) doi:10.1101/gr.103101.109.

126. Loomba, R. *et al.* DNA methylation signatures reflect aging in patients with nonalcoholic steatohepatitis. *JCI Insight* (2018) doi:10.1172/jci.insight.96685.

127. Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci U S A* (2014) doi:10.1073/pnas.1412759111.

128. Hardy, T. *et al.* Plasma DNA methylation: A potential biomarker for stratification of liver fibrosis in non-alcoholic fatty liver disease. *Gut* (2017) doi:10.1136/gutjnl-2016-311526.

129. Murphy, S. K. *et al.* Relationship between methylome and transcriptome in patients with nonalcoholic fatty liver disease. *Gastroenterology* (2013) doi:10.1053/j.gastro.2013.07.047.

130. Nishida, N. *et al.* Unique features associated with hepatic oxidative DNA damage and DNA methylation in non-alcoholic fatty liver disease. *Journal of Gastroenterology and Hepatology (Australia)* (2016) doi:10.1111/jgh.13318.

131. Laberge, R. M. *et al.* Glucocorticoids suppress selected components of the senescence-associated secretory phenotype. *Aging Cell* (2012) doi:10.1111/j.1474-9726.2012.00818.x.

132. Moiseeva, O. *et al.* Metformin inhibits the senescence-associated secretory phenotype by interfering with IKK/NF-κB activation. *Aging Cell* (2013) doi:10.1111/acel.12075.

133. Singh, M. *et al.* Effect of Low-Dose Rapamycin on Senescence Markers and Physical Functioning in Older Adults with Coronary Artery Disease: Results of a Pilot Study. *J Frailty Aging* (2016) doi:10.14283/jfa.2016.112.

134. Wang, E. Senescent Human Fibroblasts Resist Programmed Cell Death, and Failure to Suppress bell Is Involved. *Cancer Res* (1995).

135. Zhu, Y. *et al.* The achilles' heel of senescent cells: From transcriptome to senolytic drugs. *Aging Cell* (2015) doi:10.1111/acel.12344.

136. Kirkland, J. L. & Tchkonia, T. Cellular Senescence: A Translational Perspective. *EBioMedicine* Preprint at https://doi.org/10.1016/j.ebiom.2017.04.013 (2017).

137. Kirkland, J. L. & Tchkonia, T. Senolytic Drugs: From Discovery to Translation. *J Intern Med* joim.13141 (2020) doi:10.1111/joim.13141.

138. Hickson, L. T. J. *et al.* Senolytics decrease senescent cells in humans: Preliminary report from a clinical trial of Dasatinib plus Quercetin in individuals with diabetic kidney disease. *EBioMedicine* (2019) doi:10.1016/j.ebiom.2019.08.069.

139. Justice, J. N. *et al.* Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study. *EBioMedicine* (2019) doi:10.1016/j.ebiom.2018.12.052.

11

Hyperinsulinemia is highly associated with markers of hepatocytic senescence in two independent cohorts

Abraham S. Meijnikman*, Casper C. van Olden*, Omrum Aydin, Hilde Herrema, Dorota Kaminska, Dimitra Lappa, Ville Männistö, Valentina Tremaroli, Louise E. Olofsson, Maurits de Brauw, Arnold van de Laar, Joanne Verheij, Victor E. A. Gerdes, Thue W. Schwartz, Jens Nielsen, Fredrik Bäckhed, Päivi Pajukanta, Jussi Pihlajamäki, Tamar Tchkonia, James L. Kirkland, Folkert Kuipers, Max Nieuwdorp, Albert K. Groen

*both authors contributed equally

Diabetes, 2022 Sep 1;71(9):1929-1936

ABSTRACT

Cellular senescence is an essentially irreversible growth arrest that occurs in response to various cellular stressors and may contribute to development of type 2 diabetes and Non-Alcoholic Fatty Liver Disease (NAFLD). Here, we investigated whether chronically elevated insulin levels are associated with cellular senescence in the human liver. In 107 individuals undergoing bariatric surgery, hepatic senescence markers were assessed by immunohistochemistry as well as transcriptomics. A subset of 180 participants from the ongoing Finnish Kuopio OBesity Surgery (KOBS) study was used as validation cohort. We found plasma insulin to be highly associated with various markers of cellular senescence in liver tissue. The liver transcriptome of individuals with high insulin revealed significant upregulation of several genes associated with senescence: *p21, TGFβ, PI3K, HLA-G, IL8, p38, Ras,* and *E2F*. Insulin was associated with hepatic senescence independently of NAFLD and plasma glucose. By using transcriptomic data from the KOBS study, we could validate the specific association of insulin with senescence in the liver. Our results support an important role for hyperinsulinemia in induction of cellular senescence in the liver. These findings emphasize the importance of lowering insulin levels in obese individuals with insulin resistance.

INTRODUCTION

Cellular senescence is one of the hallmarks of aging[1]. It is defined as a stable arrest of the cell cycle coupled to specific phenotypic changes[1]. Senescent cells can secrete a collection of proteins and other factors termed the senescence-associated secretory phenotype (SASP)[2,3]. It is now generally accepted that cellular senescence contributes to aging phenotypes and accumulating evidence shows that senescence is associated with age-related diseases, such as type 2 diabetes and Non-Alcoholic Fatty Liver Disease (NAFLD)[4–6]. Although causality is yet to be established in humans, current evidence suggests that targeting of senescent cells has novel treatment potential for the treatment of several age-related diseases, including type 2 diabetes and NAFLD[7,8].

In search of mechanisms that drive cellular senescence, the focus has been on factors associated with insulin resistance such as chronic inflammation and elevated glucose and lipid levels[9]. Although, studies investigating the role of insulin itself in the development of cellular senescence are scarce, a causal role in inducing cellular senescence in adipose tissue was recently established[10]. Since hyperinsulinemia is one of the shared key features of aging, obesity, type 2 diabetes, as well as NAFLD[11–13], we hypothesized that chronically elevated insulin levels may be associated with cellular senescence in liver tissue in humans, independent of NAFLD.

In this study, we investigated the relationship between plasma insulin levels and markers of senescence in a cohort of obese patients scheduled for bariatric surgery. The main results were validated in an independent cohort.

MATERIALS AND METHODS

Participants were recruited from our bariatric surgery cohort as previously described[14]. In brief, 107 individuals underwent a complete metabolic work-up prior to their bariatric surgery procedure between September 2016 and the end of 2018. Within two months before surgery, a two-hour mixed meal test (MMT) was performed. The MMT consisted of two Nutridrink compact 125ml (Nutricia®), in total containing 23.3 grams fat, 74.3 grams carbohydrates and 24.0 grams protein. Blood samples were drawn at baseline and then 10, 20, 30, 60, 90, and 120 minutes after ingestion of Nutridrink. In a subset of 23 individuals, samples from both portal and peripheral venous blood were drawn on the day of surgery. All samples were stored at -80°C until further processing. The study was performed in accordance with the Declaration of Helsinki and was approved by the local Ethics Committee (approval code: NL55755.018.15). All participants provided written informed consent.

Diabetes definitions

Normal glucose tolerance and prediabetes definitions were in accordance with the American Diabetes Association (ADA) criteria[15]. Type 2 diabetes was defined as individuals who fulfilled the ADA criteria for diabetes, those who were treated with glucose-lowering agents, and/or who had a history of T2DM.

Liver and biopsies and histology

Biopsies were taken from segment three or five of the liver. All biopsies were snap frozen in liquid nitrogen and stored at -80°C. Paraffin-embedded histological sections were stained with Haematoxylin-Eosin and Sirius red and scored according to the Steatosis Activity and Fibrosis score (SAF) score[16] by the Dutch Liver Pathology Panel. NAFLD was categorized into NAFL when steatosis was present in >5% of hepatocytes alone or with mild inflammation but without ballooning or NASH when steatosis was present in >5% of hepatocytes and if ballooning and inflammation were both present in the biopsy. p21, BCL-2 and p53 immunohistochemistry was performed on formalin-fixed sections using Mouse anti-p21, (Zymed, 18-0401 Clone EA10), DAKO/M0887 clone 124 subclass IgG1 and Thermo Scientific/ Ms 738-P clone DO-7+BP53-12 subclass IgG2a/2b, respectively, both with secondary staining with Benchmark Ultra, Ventana. As positive controls for p21 and p53 tonsil and p53 positive tumor with overexpression (**supplemental Figure 1**). Ten blinded, consecutive, non-overlapping fields were acquired at × 400 magnification and counted manually.

Transcriptomics

RNA from the liver biopsies was extracted using the TriPure Isolation Reagent (Roche). The extracted RNA was purified using RNeasy MinElute spin columns and libraries for RNA sequencing were prepared using a r-RNA depletion method and sequenced at NOVOGENE. After ribosomal RNA depletion the RNA was fragmented using fragmentation buffer. Double-stranded cDNA was then synthesized using mRNA template and random hexamer primers (for the first strand), followed by second-strand synthesis buffer, dNTPs, RNAse H, and DNA polymerase I for the second strand. After a series of terminal repair, A ligation, and sequencing adaptor ligation, the cDNA library was size-selected and PCR-enriched. Library quality control was performed using Qubit 2.0 (library concentration), Agilent 2100 (insert size), and qPCR (for precise effective library concentration).

Validation cohort

A subset of 180 participants with liver mRNA sequencing data available from the ongoing Finnish Kuopio OBesity Surgery (KOBS) study[17] were included in the validation cohort. All participants provided informed consent and the study protocol was approved by the local

ethics committee. Liver histology was evaluated by one experienced liver pathologist and the study population was divided into those with normal liver, those with NAFLD, and those with NASH, as described above for the discovery cohort. RNA-sequencing of strand-specific Ribo-Zero libraries was performed on the HiSeq 2500 (Illumina) with 50 bp paired-end reads. The reads mapped to exons (GRCh38 assembly, release 29) were counted with the Rsubread R package and normalized with the TMM method using edgeR. Normalized read counts were converted to log2-counts-per-million (logCPM). Expression data were corrected for previously identified technical cofactors[18].

Statistics

Data are expressed as mean ± SD for normally distributed variables or as median (interquartile range) when distributions were skewed. The normal distribution of continuous variables was assessed using the Kolmogorov-Smirnov method. To gain normality, variables with skewed distribution were log transformed. When the distribution was not normal, a Wilcoxon test was performed. For more than two groups, Kruskal-Wallis test was performed. Correlation tests were performed using Pearson correlation coefficient and p-values were calculated using t-test when the distribution was normal and Spearman's rank-order correlation and p-values *via* the asymptotic *t* approximation when the distribution was not. Two-tailed significance was set at 0.05. Regression analyses were performed using multiple linear regression. For baseline variables only participants with complete baseline variables were included. For missing data in the mixed meal test, missing values were imputed using predictive mean matching using the MICE package (version 3.11.0). Transcriptomic data were analysed using the following methods: abundances of transcripts were quantified using Kallisto. Counts were normalized and differential expression calculated using DESeq2. For analyses of senescence genes, subjects were stratified in two groups of quintiles. These groups were based on baseline insulin values. Differential expressions were calculated between the highest and lowest quintiles. Analysis of genes was limited to genes in the KEGG senescence pathway. P-values of differentially expressed genes were adjusted using the Benjamini and Hochberg method and significance for false discovery was set at $p<0.05$

Data and Resource

Data available upon request.

RESULTS

The recruited participants were stratified based on glucose tolerance parameters according to ADA criteria[15]. Thirty-five individuals had normal glucose tolerance (NGT), 48 had pre-diabetes, while 24 had type 2 diabetes. The clinical characteristics of the 107 individuals selected are summarized in **Supplemental Table 1.**

Variables	No steatosis n = 22	NAFL n = 24	NASH = 10
Sex (male/female)	3/19	10/14	2/8
Age (years)	43 [34 - 47]	49 [44 - 53]	51 [46 - 56]
BMI (kg/m^2)	39 [37 - 43]	40 [37 - 42]	40 [39 - 43]
T2DM (%)	1 (5)	5 (21)	3 (30)
ALP (30-135 U/L)	88 [78 - 100]	79 [65 - 101]	72 [66 - 80]
g-GT (10-40 IU/l)	22 [18 - 26]	32 [24 - 48]	19 [16 - 37]
ALAT (0-50 IU/l)	24 [19 - 31]	38 [25 - 50]	33 [24 - 49]
ASAT (0-35 IU/l)	23 [20 - 26]	26 [22 - 30]	28 [26 - 34]
Ferritin (24-336 ug/l)	99 [37 - 130]	144 [61 - 215]	126 [82 - 192]
FPG (<5.6 mmol/l)	5.3 [5.0 - 5.5]	5.6 [5.1 - 6.6]	5.8 [5.6 - 6.7]
HbA1c (<5.6%)	5.7 [5.4 - 5.8]	5.8 [5.4 - 6.0]	6.0 [5.7 - 6.4]
HbA1c (mmol/mol)	39 [36 - 40]	40 [40 - 42]	42 [39 - 46.4]
Fasting insulin (18-48 pmol/l)	62 [48 - 81]	113 [71 - 141]	108 [82 - 155]
Total cholesterol (1.5-6.5 mmol/l)	4.7 [4.0 - 5.8]	4.8 [4.0 - 5.6]	5.1 [4.3 - 5.2]
Triglycerides (<1.7 mmol/l)	1.0 [0.8 - 1.4]	1.2 [1.0 - 1.7]	1.2 [1.0 - 1.3]

Table 1: Comparison between individuals with no steatosis, Non-Alcoholic Fatty Liver (NAFL), and Non-Alcoholic Steatohepatitis (NASH). Data are expressed as median [interquartile range] and parentheses after variable name, depict units and, if applicable, reference values. BMI: body mass index, AF: alkaline phosphatase, g-GT: gamma glutamyl transferase, ALAT: alanine aminotransferase, ASAT: aspartate aminotransferase, FPG: fasting plasma glucose, HDL: high-density lipoprotein, LDL: low-density lipoprotein.

Liver biopsies of 56 patients were available for histological analysis. The robust senescence biomarker, p21 (cyclin-dependent kinase inhibitor 1A) was used as a first screening marker: p21 protein levels were quantified by immunohistochemistry (**Figure 1A-C)**. Next, we quantified p53 to ensure robustness of the hepatocytic senescence signature (**Figure 1D-F)**. BCL-2 (anti-apoptosis marker) was also quantified (**supplemental Figure 2**). To assess insulin resistance, we analysed MMT data, in which glucose, insulin, and triglycerides were measured at seven time points. Hepatic p21 expression correlated significantly with glucose area under the curve (AUC) ($r = 0.33$, $p = 0.009$), peripheral insulin ($r = 0.49$, $p < 0.001$), and insulin AUC ($r = 0.46$, $p < 0.001$) during the MMT (**Figure 1G-I)**. In line with this, BCL-2 significantly correlated with glucose AUC ($r = 0.54$, $p = 0.033$), peripheral insulin ($r = 0.58$, $p = 0.018$), and insulin AUC ($r = 0.53$, $p = 0.036$) (**supplemental Figure 3)**. For p53, a significant correlation was observed with peripheral insulin ($r = 0.52$, $p = 0.039$) (**Supplemental Figure 3)**. Fasting glucose did not correlate with BCL-2 or p53 expression. Bearing in mind

that the pancreas drains into the portal vein, highest concentrations of insulin are likely to be found in this blood compartment. In a subset of 23 individuals, we collected portal vein blood and observed a strong correlation between p21 expression and insulin levels in the portal vein ($r = 0.74$, $p < 0.001$) (**Figure 1J**).

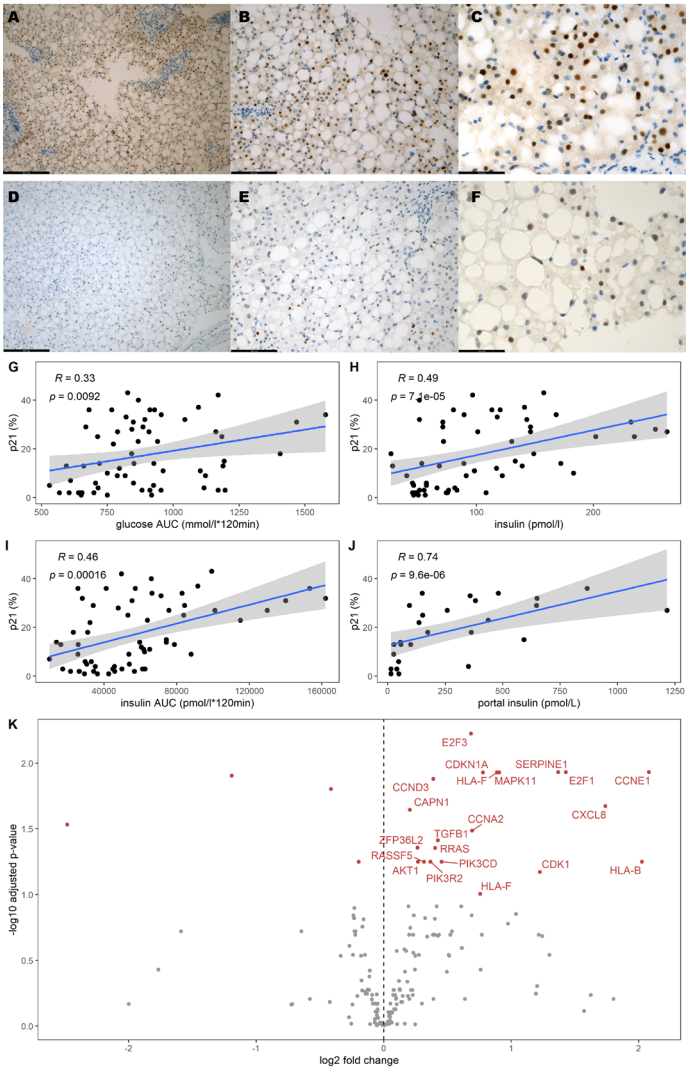


Figure 1. A-C: Representative images of immunohistochemistry staining of liver tissue with p21 with respectively 100-, 200- and 400-times magnification. **D-F:** Representative images of immunohistochemistry staining of liver tissue with p53 with respectively 100-, 200- and 400-times magnification. **G-J:** Scatterplot depicting the percentage of p21 on the y-axis and respectively, glucose AUC, peripheral fasting insulin, insulin AUC, and portal insulin on the x-axis. The blue line is linear regression prediction line, grey is the 95% confidence interval of the regression, rho is the Spearman correlation coefficient, and p-value is the significance level. **K:** Volcano plot showing genes in the KEGG senescence pathway identified by RNA sequencing of liver tissue biopsies, respectively, red genes are those that are significantly upregulated in patients with highest quintile peripheral insulin. RNA sequencing data were adjusted for age and individuals with type 2 diabetes were excluded from this analysis.

To further investigate the relationship between insulin and hepatocytic senescence, we performed RNA sequencing in liver biopsies derived from the 107 individuals. Transcriptomic analyses revealed differential expression in senescence-related genes between groups stratified for peripheral fasting insulin (**Figure 1K**). The key cell cycle inhibitors *CDKN1A* (coding for p21) and *MAPK11* as well as genes encoding SASP factors, including *TGFB1* and *CXCL8,* were upregulated in the high insulin group. In addition, the transcription factors *E2F1* and *E2F3*, which control progression from the G1 to S phase of the cell cycle[19], were higher in the high insulin group.

Next, we investigated if cellular senescence is associated with NAFLD. According to the SAF score, 22 individuals had no steatosis, 24 had NAFL, and 10 had NASH (**Table 1**). p21 protein expression was significantly higher in individuals with NASH compared to NAFL and individuals without steatosis (**Figure 2A)**. According to the MMT data, fasting insulin and insulin AUC increased with disease progression substantiating that hepatic insulin resistance plays an important role in NAFLD pathogenesis **(Figure 2B, C)**. This is further demonstrated by the increased concentration of plasma triglycerides in individuals with NAFL and NASH (**Figure 2D**), which may indicate a loss of suppression of Apolipoprotein B expression by insulin and hence increased VLDL-triglyceride production[20]. Interestingly, in patients with high portal vein insulin, increased expression of p21 was observed suggesting that insulin is associated with senescence independent of NAFLD.

To further investigate interrelations among hyperinsulinemia, senescence, and NAFLD, we performed regression analysis**.** The regression model predicting p21 percentage, using either fasting peripheral insulin or the NAFLD classification, had an adjusted R^2 of 0.18, $F(1,58) = 14$, $p<0.001$ and 0.39, $F(1,61) = 40$, $p < 0.001$, respectively. When using both insulin and NAFLD classification to predict p21 percentage, the adjusted R^2 was 0.45, $F(2,57) = 24.8$, $p<0.001$ and both variables were significant predictors, indicating that insulin, independently from NAFLD, is associated with hepatocytic senescence. Neither fasting glucose or glucose AUC were significant predictors when combined with fasting peripheral insulin and NAFLD. Adding insulin AUC to the model with NAFLD and fasting peripheral insulin did not improve the R^2 or the Akaike information criterion.

To validate our findings, we analyzed data from 180 individuals from an independent cohort from whom clinical data, liver histology, and liver transcriptomics data were available (**supplemental Table 2**). In line with outcomes from our study, hepatic *CDKN1A (p21)* expression correlated with insulin resistance and NAFLD in the KOBS study ($r = 0.38$, $p <0.001$) (**Figure 3**). Moreover, stratifying individuals from the validation cohort by insulin levels revealed a strong cellular senescence signature. In addition, we used the same regression model and found that insulin was associated with hepatocytic senescence independently of NAFLD. Predicting *CDKNIA* using fasting peripheral insulin and NAFL classification had

an adjusted R^2 of 0.10, $F(1,115) = 14$ ($p<0.001$) and 0.18, $F(1,119) = 28$ ($p<0.001$), respectively. The combined model had an adjusted R^2 of 0.22, $F(2,114) = 17$ ($p<0.001$), with both variables being significant predictors, validating the findings in our other cohort.

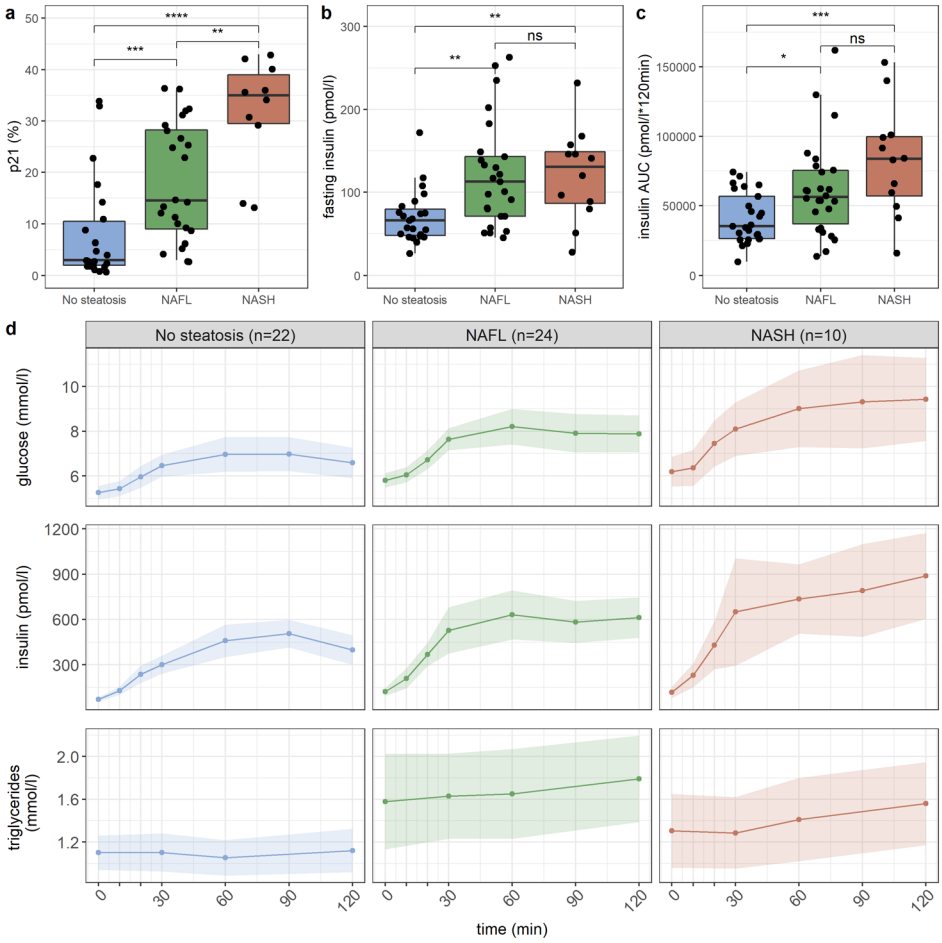


Figure 2. A-C: Boxplots showing p21 percentage, peripheral fasting insulin, and insulin AUC on the y-axis and the different categories of NAFL displayed on the x-axis (n=56). **D:** Glucose, insulin, and triglyceride, excursions during a 2-hour mixed meal tolerance test comparing individuals with no steatosis, with Non-Alcoholic Fatty Liver (NAFL), and Non-Alcoholic Steatohepatitis (NASH). Data are means +/- 95% confidence intervals.

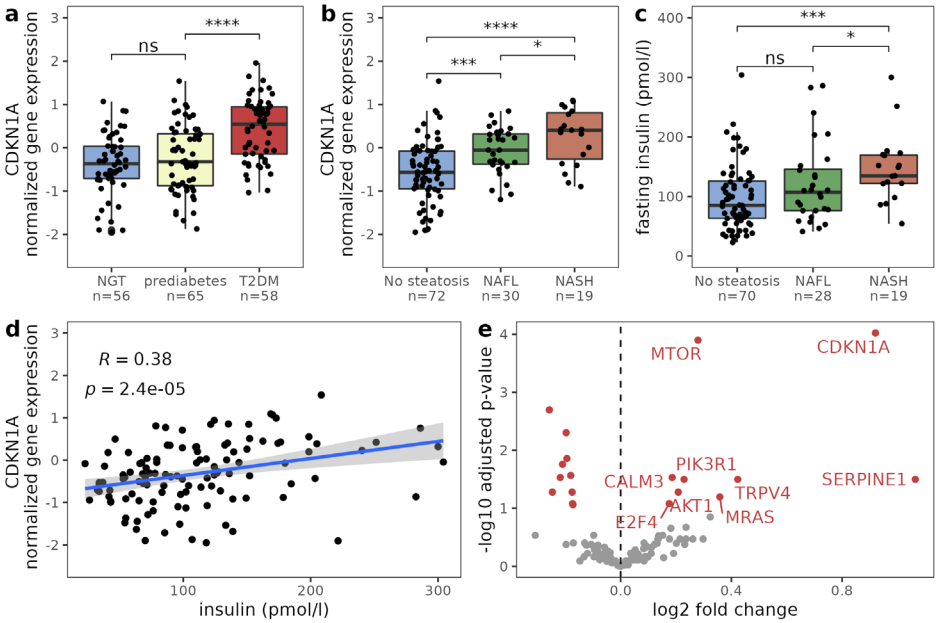


Figure 3. Results from the validation cohort (n=180). A: Hepatic CDKN1A expression in individuals with normal glucose tolerance, impaired fasting glucose, or type 2 diabetes (T2DM) **B:** Boxplot of hepatic CDKN1A expression in individuals with no steatosis, Non-Alcoholic Fatty Liver (NAFL), and Non-Alcoholic Steatohepatitis (NASH). **C:** Peripheral fasting insulin concentrations in individuals without NAFL, NAFL, and NASH: individuals with T2DM were excluded from this analysis. **D:** Scatterplot displaying the relation between peripheral insulin and CDKN1A expression. Blue line is linear prediction regression lines, grey area is 95% confidence intervals. Rho is Spearman correlation coefficient; p is the significance level; individuals with T2DM were excluded from this analysis. **E:** Volcano plot showing genes in the KEGG senescence pathway identified by RNA sequencing of liver biopsies. Genes in red indicate that the gene is significantly upregulated in individuals with fasting insulin levels in the highest quintile group. Adjusted for age; and individuals with type 2 diabetes were excluded from this analysis.

DISCUSSION

The major finding of this study is the strong association between plasma insulin and the presence of senescence markers in liver tissue derived from individuals undergoing bariatric surgery, suggesting that insulin might play a role in the induction of senescence.

Cellular senescence has been implicated in the development of type 2 diabetes and NAFLD[21]. It has long been considered to be a consequence and not the cause of hepatic steatosis. However, studies in mice revealed a causal role for cellular senescence in development of hepatic steatosis[7,8]. Removal of the senescent cells induced attenuation of hepatic lipid accumulation. Indeed, our data show a clear association between plasma insulin levels and hepatic steatosis (Figure 2B,C). Furthermore, our data also open the possibility that the role of insulin may be more complex than just its role in controlling lipid accumulation. Protein expression of the senescence markers p21, p53, and BCL-2 in the liver increased with plasma

insulin concentration. Moreover, at the transcriptional level, expression of the cell cycle regulators *CDKN1A* and *MAPK11* as well as multiple genes encoding SASP proteins were enhanced in individuals with high plasma insulin levels. Interestingly, multiple regression analysis showed that insulin correlated with senescence independently from NAFLD indicating that insulin-induced senescence indeed may precede NAFLD. In line with this hypothesis, we noted that individuals without NAFLD with a relatively high expression of p21 also had high concentrations of insulin in their portal blood. Although our results are associative, we speculate that insulin rather than glucose might be an important factor for inducing cellular senescence in hepatocytes. To validate our results, we performed similar regression analyses in 180 individuals from the ongoing KOBS study[17]. In line with data from our study, hepatic *CDKN1A (p21)* expression correlated with insulin resistance and NAFLD in the KOBS study (r = 0.38, p <0.001). Also in this data set, multiple regression analysis revealed that insulin correlated independently from NAFLD and glucose with hepatic senescence markers. This suggests that, at least partly, insulin-induced senescence is upstream in the pathogenesis of NAFLD. This is in line with current evidence showing that insulin resistance precedes NAFLD[22]. Also, our studies are in line with the recent paper showing that insulin induces senescence in adipose tissue[10].

Limitations

We note that the analyses of human data sets in our study have some limitations. Here, we used liver and plasma samples obtained from individuals who underwent bariatric surgery, which may introduce bias due to pre-operative weight-loss. However, individuals who had lost more than 3% of their body weight in the month prior to surgery or more than 5% during the six months before surgery were excluded. Last, we used transcriptomics data to validate our findings. To what extent the senescence signature found in the transcriptomics data reflect a "true" senescent signature remains to answered.

In summary, we have shown by using two independent cohorts that plasma insulin levels correlate with markers of cellular senescence in liver tissue. Further studies are needed to investigate to what extent insulin is causally involved in inducing senescence in the liver.

Acknowledgments

We thank the nursing staff and all participating individuals for making this study possible. We also thank Manuela Krämer and Robert Jakubowicz for the preparation of the samples. The authors wish to acknowledge CSC - IT Center for Science, Finland, for computational resources and pathologist Vesa Kärjä for histological characterization in KOBS cohort. We also thank the Dutch Liver Pathology Panel members: Carolien M. Bronkhorst, Michael Doukas, Paul Drillenburg, Arantza Farina, Natascha N.T. Goemaere, Mieke Jonker, Miangela

M. Lacle, and Iryna Samarski. AG is the **guarantor** of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Authors' contributions

AM, CvO, OA, MdB, and AvdL performed the patient visits and were involved in the recruitment of individuals. AM, and CvO, performed the data analyses. VT, and HH were responsible for sample processing. DK analyzed the RNA-seq data and performed the statistical analyses of the KOBS validation cohort. VM was responsible for the clinical data and interpretation of the KOBS. JPI is the PI of the KOBS cohort. FK, TT, JLK, VG, MN, and AG designed the study and supervised all parts of the project. AM, CvO, HH, and AG drafted the manuscript. All authors provided support and constructive criticism throughout the project and approved the final version of the article.

Declaration of interest

M. N. is on the Scientific Advisory Board of Caelus Health, The Netherlands. F.B. is on the board of directors of Metabogen AB, Sweden. However, none of these possible conflicts of interest bear direct relation to the outcomes of this specific study. T.T. and J.L.K. have a financial interest related to this research: patents on senolytic drugs and their uses are held by Mayo Clinic. This research has been reviewed by the Mayo Clinic Conflict of Interest Review Board and was conducted in compliance with Mayo Clinic Conflict of Interest policies. All other authors declare that they have no competing interests.

Financial support

This study was funded by Leducq consortium grant CVD01 and Novo Nordisk Foundation (NNF15OC0016798). DK was supported by the Academy of Finland grant (contract 316458). The Kuopio Obesity Surgery Study (PI J.P.) was supported by Kuopio University Hospital Project grant (EVO/VTR grants 2005-2019) and the Academy of Finland grant (Contract no. 138,006). A.S.M is supported by EFSD/Lilly. FK is supported by the Noaber Foundation. JLK and TT were supported by National Institute of Health grants AG013925, AG062413, and the Translational Geroscience Network (AG061456), Robert and Arlene Kogod, the Connor Group, Robert J. and Theresa W. Ryan, and the Ted Nash Long Life and Noaber Foundations.

REFERENCES

1. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* Preprint at https://doi.org/10.1016/j.cell.2013.05.039 (2013).
2. Ogrodnik, M., Salmonowicz, H. & Gladyshev, V. N. Integrating cellular senescence with the concept of damage accumulation in aging: Relevance for clearance of senescent cells. *Aging Cell* Preprint at https://doi.org/10.1111/acel.12841 (2019).
3. Minamino, T. *et al.* A crucial role for adipose tissue p53 in the regulation of insulin resistance. *Nat Med* **15**, 1082–1087 (2009).
4. Palmer, A. K. *et al.* Targeting senescent cells alleviates obesity-induced metabolic dysfunction. *Aging Cell* (2019) doi:10.1111/acel.12950.
5. Gustafson, B., Nerstedt, A. & Smith, U. Reduced subcutaneous adipogenesis in human hypertrophic obesity is linked to senescent precursor cells. *Nat Commun* (2019) doi:10.1038/s41467-019-10688-x.
6. Aravinthan, A. *et al.* Hepatocyte senescence predicts progression in non-alcohol-related fatty liver disease. *J Hepatol* (2013) doi:10.1016/j.jhep.2012.10.031.
7. Meijnikman, A. S. *et al.* Evaluating causality of cellular senescence in Non-Alcoholic Fatty Liver Disease. *JHEP Reports* **0**, 100301 (2021).
8. Ogrodnik, M. *et al.* Cellular senescence drives age-dependent hepatic steatosis. *Nat Commun* (2017) doi:10.1038/ncomms15691.
9. Palmer, A. K., Gustafson, B., Kirkland, J. L. & Smith, U. Cellular senescence: at the nexus between ageing and diabetes. *Diabetologia* (2019) doi:10.1007/s00125-019-4934-x.
10. Li, Q. *et al.* Obesity and hyperinsulinemia drive adipocytes to activate a cell cycle program and senesce. *Nat Med* **27**, (2021).
11. Gumbiner, B. *et al.* Effects of Aging on Insulin Secretion. *Diabetes* (1989) doi:10.2337/diab.38.12.1549.
12. Arab, J. P., Arrese, M. & Trauner, M. Recent Insights into the Pathogenesis of Nonalcoholic Fatty Liver Disease. *Annual Review of Pathology: Mechanisms of Disease* (2018) doi:10.1146/annurev-pathol-020117-043617.
13. Corkey, B. E. Banting lecture 2011: hyperinsulinemia: cause or consequence? *Diabetes* **61**, 4–13 (2012).
14. van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *J Intern Med* joim.13157 (2020) doi:10.1111/joim.13157.
15. American Diabetes Association. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care* **43**, S14–S31 (2020).
16. Utility and Appropriateness of the Fatty Liver Inhibition of Progression (FLIP) Algorithm and Steatosis, Activity, and Fibrosis (SAF) Score in the Evaluation of Biopsies of Nonalcoholic Fatty Liver Disease.
17. Simonen, M. *et al.* Desmosterol in human nonalcoholic steatohepatitis. *Hepatology* (2013) doi:10.1002/hep.26342.
18. Männistö, V. *et al.* Protein Phosphatase 1 Regulatory Subunit 3B Genotype at rs4240624 Has a Major Effect on Gallbladder Bile Composition. *Hepatol Commun* **5**, 2021 (2020).
19. Rubin, S. M., Gall, A. L., Zheng, N. & Pavletich, N. P. Structure of the Rb C-terminal domain bound to E2F1-DP1: A mechanism for phosphorylation-induced E2F release. *Cell* (2005) doi:10.1016/j.cell.2005.09.044.
20. Wiegman, C. H. *et al.* Hepatic VLDL production in ob/ob mice is not stimulated by massive de novo lipogenesis but is less sensitive to the suppressive effects of insulin. *Diabetes* (2003) doi:10.2337/diabetes.52.5.1081.
21. Papatheodoridi, A.-M., Chrysavgis, L., Koutsilieris, M. & Chatzigeorgiou, A. The role of senescence in the development of non-alcoholic fatty liver disease and progression to non-alcoholic steatohepatitis. *Hepatology* (2019) doi:10.1002/hep.30834.
22. Khan, R. S., Bril, F., Cusi, K. & Newsome, P. N. Modulation of Insulin Resistance in Nonalcoholic Fatty Liver Disease. *Hepatology* **70**, 711–724 (2019).

Supplemental material available at:

https://diabetesjournals.figshare.com/articles/figure/Hyperinsulinemia_is_highly_associated_with_markers_of_hepatocytic_senescence_in_two_independent_cohorts/20051795

12

Adipose tissue senescence is associated with increased postprandial plasma insulin concentrations in bariatric surgery patients

Casper C. van Olden*, Abraham S. Meijnikman*, Omrum Aydin, Hilde Herrema, Maurits de Brauw, Arnold van de Laar, Victor E. A. Gerdes, Thue W. Schwartz, Jens Nielsen, Fredrik Bäckhed, Folkert Kuipers, Max Nieuwdorp, Albert K. Groen

*both authors contributed equally

Submitted after revisions

ABSTRACT

Insulin resistance is a key feature of type 2 diabetes and is detectable in early stages of disease progression, i.e., before blood glucose elevation. Recent work has convincingly shown that adipocyte senescence is a driver of peripheral insulin resistance. The emergence of senescent adipocytes is linked to chronic hyperinsulinemia and is modulated by metformin. These mechanistic insights are derived from cell culture and animal experiments but supporting data in humans are scarce. Here we report a strong association between plasma insulin levels after mixed meal test and levels of p16 protein, a marker of cellular senescence, in mesenteric adipose tissue (MAT) of 103 individuals from our bariatric surgery cohort. P16 protein levels were increased irrespective of glucose concentrations. Several genes associated with senescence: *CHKN2A, E2F3, TGFB1, MAPK12* and *CHEK2*, were upregulated in MAT of individuals with high insulin levels. Furthermore, we observed that p16 protein levels were significantly lower in individuals using metformin. These results confirm the recently proposed link between hyperinsulinemia and cellular senescence in adipocytes and demonstrate that senescent adipocytes are present in MAT in insulin resistance prior to the onset of type 2 diabetes in humans.

INTRODUCTION

It is now widely accepted that cellular senescence contributes to several aging phenotypes. The numbers of senescent cells increase in several organs and tissues with aging while genetic or pharmacological clearance of senescent cells can alleviate multiple age-related diseases and increase health span in animal models[1–3]. Despite this clear link, aging and cellular senescence are not synonymous: cells can enter a senescent state regardless of organismal age. This process is driven by a variety of factors, such as genotoxic stress, mitogens and inflammatory cytokines. Currently available evidence supports the idea that anabolic signalling *via* the insulin/insulin-like growth factor 1 (IGF-1) pathway accelerates aging. In fact, the insulin/IGF-1 signalling pathway is the most conserved aging-controlling pathway in evolution[4,5]. Genetic downregulation of the insulin/IGF-1 signalling pathway has been shown to extend lifespan in worms, flies and mice[5–7]. Only recently, it was shown that reducing circulating insulin levels improves metabolic homeostasis, health span and longevity in mice[8].

In obesity, the number of senescent cells is particularly increased in adipose tissues[9,10]. Recent work establishes that clearance of cells high in *p16* and *p21* expression alleviates insulin resistance in obese mice models[11,12]. In addition, mature human adipocytes, despite generally considered to be postmitotic, are able to activate a cell cycle program in association with obesity and hyperinsulinemia which results in a pro-inflammatory senescent phenotype[13]. Li et al.[13] further demonstrated that targeting the adipocytic cell cycle program with metformin modulates the extent of senescence.

Here, we use data from the Bariatric Surgery in Amsterdam (BARIA) cohort study[14] to evaluate the association between hyperinsulinemia and adipose tissue senescence in humans. Insulin levels after a mixed meal test (MMT) positively correlated with p16 levels in mesenteric adipose tissue (MAT) but not in subcutaneous adipose tissue (SAT). We further found that metformin-treated individuals exhibit lower p16 protein levels in their MAT.

METHODS

Participant's recruitment and sample collection

The study was performed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC. All participants provided written informed consent. Individuals were recruited from our bariatric surgery cohort (the BARIA study) as previously described[14]. In brief, 103 individuals underwent a complete metabolic work-up prior to their bariatric surgery procedure. Anthropometric

measurements including height, weight, waist and hip circumference were taken. In addition, body fat percentage using bioelectrical impedance and blood pressure were measured. Fasting blood samples were used for the determination of hemoglobin, HbA1c, glucose and lipid profile. Within two months before surgery, a two-hour mixed meal test (MMT) was performed to assess insulin resistance. The MMT consisted of two Nutridrink compact 125ml (Nutricia®), in total containing 23.3 grams fat, 74.3 grams carbohydrates (of which 38.5 grams sugar) and 24.0 grams protein. The individuals received this meal after fasting for a minimum of nine hours. Time point zero refers to the moment at which the participant had fully consumed the meal. Blood samples were drawn *via* an intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes. Insulin and glucose were measured at these seven time points. All samples were stored at -80°C until further processing. The study was performed in accordance with the Declaration of Helsinki and was approved by the Academic Medical Center Ethics Committee of the Amsterdam UMC (approval code: NL55755.018.15). All participants provided written informed consent.

Diabetes and prediabetes definitions

We stratified individuals by classifications of glycaemic control according to plasma glucose and glycated haemoglobin (HbA1c) levels formulated in the American Diabetes Association (ADA) criteria: normoglycemia = fasting plasma glucose (FPG) <100mg/dl (<5.6mmol/l), hyperglycaemia = FPG >100 mg/dl (>5.6 mmol/l) and / or HbA1c >5.7% (>39mmol/mol) or metformin use. All subjects using metformin were also self-reported diagnosed with type 2 diabetes mellitus.

Adipose tissue biopsies and histology

Tissue biopsies were obtained during surgery. Two adipose tissue compartments were analysed in the present study and biopsies were obtained as follows: subcutaneous adipose tissue from one of the laparoscopic incisions in the upper abdomen and mesenteric adipose tissue from the appendices of the transverse colon. All biopsies were snap frozen in liquid nitrogen and stored at -80°C until further processing and paraffin-embedded for histology. For mesenteric adipose tissue 48 biopsies and subcutaneous 47, met the quality criteria for histological analysis. The p16 immunohistochemistry was performed on formalin-fixed sections using Anti-CDKN2A Mouse Monoclonal Antibody, (Immunologic, ILM 0632-C1). Six blinded consecutive non-overlapping fields were acquired per biopsy at × 400 magnification. Area of antibody staining was measured using ImageJ. The mean of the six images was used to determine the level of expression.

Transcriptomics

RNA from the mesenteric and subcutaneous adipose tissue biopsies, that were collected during the surgery were extracted using the TriPure Isolation Reagent (Roche) and homogenization in a FastPrep-24 instrument using Lysing Matrix D tubes (MP Biomedicals). The extracted RNA was purified using RNeasy MinElute spin columns and libraries for RNA sequencing were prepared using a r-RNA depletion method and sequenced at NOVOGENE, using a dUTP-based strand-specific protocol. After ribosomal RNA depletion using Ribo-Zero kits, the RNA was fragmented using fragmentation buffer. Double-stranded cDNA was then synthesized using mRNA template and random hexamer primers (for the first strand), followed by second-strand synthesis buffer, dNTPs, RNAse H, and DNA polymerase I for the second strand. After a series of terminal repair, A ligation and sequencing adaptor ligation, the cDNA library was size-selected and PCR-enriched. Strand-specificity was achieved by the incorporation of dUTPs instead of dTTPs in the second-strand synthesis. Digestion of dUTPs by uracil-DNA glycosylase (UGDase) prevents the second strand from being amplified. Library quality control was performed using Qubit 2.0 (library concentration), Agilent 2100 (insert size) and qPCR (for precised effective library concentration). Raw RNA sequencing reads were quality trimmed with Trimmomatic, removing adaptors, the first 6 bp from the 5' end, applying a sliding window quality trimming (window width 4, threshold Q-score 15), and removing reads that were shorter than 50 bp post-processing. Remaining high-quality reads were pseudoaligned to the human transcriptome (GRCh38) using kallisto (v.0.46.0) with 100 bootstraps and using the "-rf-stranded" and GC-bias correction options[15].

Statistical analyses

Data are expressed as median (interquartile range). The normal distribution of continuous variables was assessed with the Kolmogorov-Smirnov method. To gain normality, variables with a skewed distribution were log transformed. Correlation tests were performed using Pearson correlation coefficient and p-values were calculated using a t-test. Two-tailed significance was set at 0.05. Transcriptomic data was analysed using the following methods: abundances of transcripts were quantified using Kallisto[15]. Counts were normalized and differential expressions calculated using DESeq2[16]. For analyses of senescence genes, non-metformin using subjects were stratified in two groups of quintiles. One, for peripheral insulin resistance based on area under the curve of insulin during the mixed meal test. Second, for hepatic insulin resistance based on baseline insulin values. Differential expressions were calculated between the highest and lowest quintile. Analyses of genes was limited to genes in the KEGG senescence pathway, total of 149 genes. P-values of differentially expressed genes were adjusted using the Benjamini and Hochberg method for false discovery rate, significance for false discovery was set at $p<0.1$[17]. Data analyses were performed in R version

4.0.2 with RStudio version 1.3.1073 using packages, Tidyverse (version 1.3.0), Tableone (version 0.10.0), cowplot (version 1.0.0), ggpubr 0.4.0.

RESULTS

A total of 103 individuals underwent a complete metabolic work-up prior to their bariatric surgery procedure. Within two months before surgery, MMT was performed to assess insulin resistance (IR)[14]. Blood samples were drawn via an intravenous line at baseline, 10, 20, 30, 60, 90 and 120 minutes after ingestion of the meal. Insulin and glucose were measured at these 7 timepoints and areas under the curves (AUC) were calculated using trapezoidal integration. We stratified individuals in groups based on normoglycaemia, hyperglycaemia as formulated in the ADA criteria or metformin use. The baseline characteristics of the participants can be viewed in **Table 1**.

	Normoglycaemia	Hyperglycaemia	metformin use
n	35	53	15
age (years)	42.0 [37.5, 50.0]	47.0 [41.0, 54.0]	47.0 [43.0, 52.0]
female (%)	27 (77.1)	43 (81.1)	10 (66.7)
BMI	39.0 [37.4, 40.9]	39.8 [37.5, 41.3]	38.9 [36.3, 42.1]
FPG (mmol/l)	5.1 [5.0, 5.3]	5.8 [5.4, 6.3]	7.1 [6.3, 8.5]
fasting insulin (pmol/l)	69.4 [46.2, 111.7]	83.4 [69.0, 116.0]	96.6 [51.0, 153.5]
glucose AUC (mmol/l*120 min)	723 [657, 804]	881 [817, 953]	1189 [1108, 1433]
insulin AUC (nmol/l*120 min)	35 [27, 64]	57 [46, 71]	40 [22, 67]
HbA1c NGSP (%)	5.4 [5.3, 5.5]	5.8 [5.6, 5.9]	6.8 [6.3, 7.4]
HOMA2 IR	1.3 [0.9, 2.1]	1.6 [1.3, 2.2]	1.9 [1.0, 3.4]
Total cholesterol (mmol/l)	4.6 [4.1, 5.5]	5.2 [4.6, 5.8]	4.0 [3.5, 4.6]
HDL cholesterol (mmol/l)	1.1 [1.1, 1.4]	1.2 [0.9, 1.4]	1.0 [1.0, 1.2]
LDL cholesterol(mmol/l)	2.9 [2.4, 3.5]	3.3 [2.7, 4.0]	2.5 [2.0, 2.9]
triglycerides (mmol/l)	1.0 [0.9, 1.4]	1.3 [1.0, 1.8]	0.9 [0.8, 1.4]

Table 1. Baseline characteristics of all participants. All values are displayed as median with inter quartile range. FPG = fasting plasma glucose. AUC = area under the curve

Figures

During surgery, biopsies were taken from mesenteric (MAT) and subcutaneous adipose tissue (SAT) compartments and were snap frozen and paraffin-embedded for histological examination. A total of 48 and 47 biopsies for MAT and SAT, respectively, met the quality criteria for histological analysis and p16 protein immunohistochemistry was performed using Anti-CDKN2A Mouse Monoclonal Antibody (Immunologic, ILM 0632-C1). Six blinded consecutive images at x400 magnification were taken and area of staining was quantified using ImageJ. Individuals with hyperglycaemia trended towards increased staining of p16

protein in MAT compared with normoglycaemia (t(25)=-1.78, p-value=0.085). Interestingly, metformin usage was associated with reduced levels of p16 protein staining (t(25)=2.66, p-value=0.014), comparable to normoglycemic controls. P16 protein was differentially expressed in SAT (Figure 1A).

Hyperglycaemia can be classified into different categories such as increased fasting plasma glucose or increased postprandial glycaemia, represented by glucose tolerance or HbA1c. We used a MMT to investigate which facet of glycaemic disturbance represents the most likely driver of hyperglycaemia in our cohort[18]. We used four different parameters of the MMT, i.e., fasting insulin, fasting glucose, insulin AUC and glucose AUC, to evaluate the consequences of the different facets of glycaemic disturbance[14]. Fasting insulin represents the amount of insulin needed in the fasting state to maintain normal glucose levels: an increase in fasting insulin is associated with hepatic IR and inability to supress gluconeogenesis in the liver[19]. Elevated fasting glucose shows that the body is no longer able to maintain proper glucose levels in the fasting state. Insulin AUC represents the insulin response to ingestion of a meal, where peripheral IR will result in higher postprandial levels of insulin to maintain glucose levels. A high glucose AUC can be caused by peripheral IR when the increased amounts of insulin are not sufficient to maintain normal glucose levels or in situations with inadequate insulin sensing or production.

Of these different parameters measured we found no association levels for glucose, either fasting or AUC, and for fasting insulin regardless of metformin use (Figure 1B-D). In contrast, we did find a significant correlation between p16 protein expression and insulin AUC (r=0.39, p=0.015) in individuals not using metformin (Figure 1E). No such correlation was observed in individuals using metformin (Figure 1E). In SAT no differences in protein expression of p16 were found (data not shown). To obtain additional proof for differences in cellular senescence between adipose tissue depots, transcriptomics analysis were performed on RNA isolated from snap frozen biopsies of all participants. Abundances of transcripts were quantified using Kallisto[15]. Counts were normalized and differential expression was calculated using DESeq2[16]. For analyses of senescence genes, subjects were stratified in groups of quintiles based on insulin AUC during the MMT. Differential expressions were calculated between the highest and lowest quintile and adjusted for age. Analysis of genes was limited to genes in the KEGG senescence pathway. We found 7 of the genes in the KEGG senescence pathway to be significantly upregulated in the highest quintile of insulin AUC compared with the lowest quintile in MAT. These genes included the cell cycle inhibitor *CDKN2A* (encoding p16) and *TGFB1*. In addition, significant upregulation of the *HLA-A* was observed (Figure 1F). In agreement with the observation that p16 only was associated with hyperglycaemia in MAT, we did not observe regulation of senescence genes in the SAT transcriptome (Figure 1F).

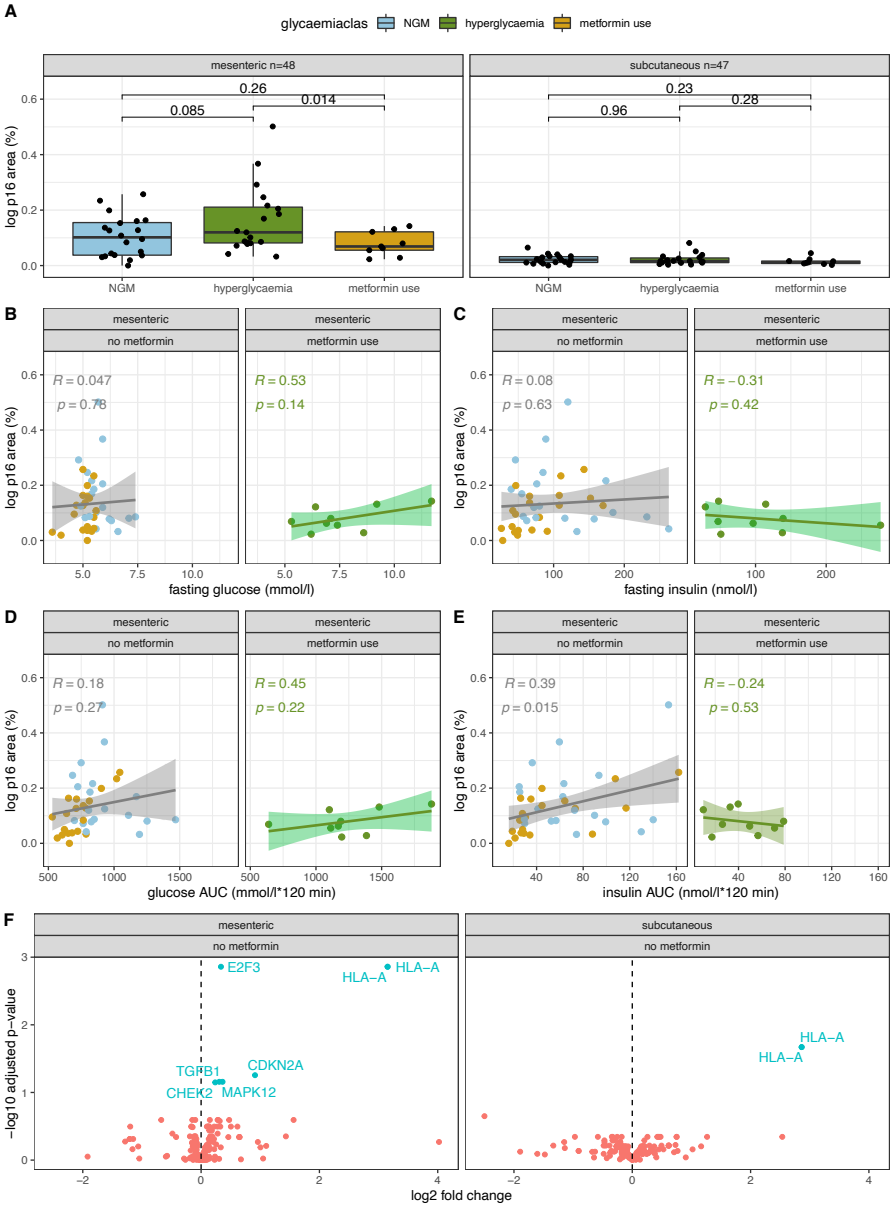


Figure 1. A: Boxplot of log transformed p16 area percent on the y-axis in mesenteric and subcutaneous adipose tissue from individuals divided in normal normoglycemic subjects, subjects with hyperglycaemia, and subjects using metformin displayed on the x-axis. In both boxplots, numbers are p-values calculated using Welch two sample t-test. **B-E:** Scatterplot of log transformed p16 area percent and fasting glucose, fasting insulin, glucose area under the curve (AUC) and insulin AUC during a 2-hour mixed meal test. Facets are divided in metformin use and not metformin use. Blue line is linear regression prediction line, grey area is the 95% confidence interval of the regression, rho is the Pearson correlation coefficient, and p-value is the significance level computed using a t-test. **F**. Volcano plot showing genes in the KEGG senescence pathway identified by RNA sequencing of mesenteric and subcutaneous adipose tissue biopsies, respectively in subjects not using metformin; blue genes are significantly upregulated in patients with highest quintile insulin AUC. RNA sequencing data were adjusted for age and individuals with T2DM were excluded from this analysis.

DISCUSSION

The marjor finding of this study is the strong association between plasma insulin levels after MMT and levels of p16 protein. In addition, P16 protein levels were increased irrespective of glucose concentrations. According to transcriptomics analyses of MAT, several genes associated with senescence: *CHKN2A, E2F3, TGFB1, MAPK12* and *CHEK2*, were upregulated in MAT of individuals with high insulin levels. Furthermore, we observed that p16 protein levels were significantly lower in individuals using metformin. These results are in line with a recent article showing an increase in adipocytic senescence exhibiting increased p16 protein contents in subjects with high insulin levels[13]. Also, our results indicate an effect of metformin treatment on senescence in humans *in vivo*. This is in line with reported *in vitro* effects of metformin on senescence markers in adipocytes. Metformin was shown to block mTOR-mediated mitogenic signalling[20], to supress the expression of senescent transcriptional programs and to induce expression of genes associated with quiescence. Our study design does not enable us to identify the causality of the lower senescence burden in metformin-treated subjects. It should be noted that the insulin AUC of those treated with metformin was in the mid to lower range, which may be a consequence of metformin treatment. This insulin lowering could have subsequently led to less senescence. It will be of interest to compare metformin *vs*. senolytics on IR and senescence in humans[11,12,21].

Interestingly, Li et al.[13] showed changes in adipocytic senescence in SAT whilst our study showed only differences in MAT. The methods used by Li et al. for processing of adipose tissues differed greatly from ours, i.e., use of isolated cell fractions or whole tissue biopsies, respectively. The use of isolated cells probably enabled Li et al. to pick up more subtle changes. Systemic insulin resistance is linked to changes in both subcutaneous and mesenteric adipose tissue depots with some studies suggesting a larger role for mesenteric adipose tissue[22–24]. Future studies should employ isolated cell analysis on MAT to assess the presence of cellular senescence more accurately.

Our observations are in agreement with a recent study showing that cellular senescence precedes hyperglycaemia[9]. Individuals with a genetic predisposition for T2DM (defined as first degree relative of an individual with T2DM) but without impaired glucose tolerance showed increased expression of senescence markers in SAT[9]. Unfortunately, insulin concentrations were not measured in this study. We can therefore only speculate on the presence of hyperinsulinemia, as is often the case in individuals with a genetic predisposition for T2DM even in the absence of obesity[25].

In summary, by using clinical, immunohistochemistry, and transcriptomics data derived from non-diabetic individuals with obesity, we have shown that peripheral IR is highly correlated with markers of cellular senescence in MAT prior to the presence of glycaemic

dysregulation. Our results confirm previous studies indicating that adipocytic cellular senescence may play an important role already in the earliest stages of IR, i.e., before the onset of T2DM. Furthermore, our findings confirm a potential role for metformin in decreasing senescence in adipose tissue. It is worth investigating the triggers of this early-stage senescence and whether there is a role for senolytics in its prevention and subsequent metabolic dysregulation. These findings emphasize the importance of preventive strategies and early intervention to halt progression of IR in obese individuals.

Acknowledgments

MN is supported by a personal ZONM-VICI grant 2020 [09150182010020] CO was appointed on a NNF GUTMMM grant 2016 (to MN, TWS and FB). The study reported here was additionally supported by le Ducq consortium grant 17CVD01 to FB and MN. MN HH and AKG are supported by a Dutch Heart Foundation CVON in CONTROL-2 consortium grant.

Author Contributions

Casper C. van Olden, Abraham S. Meijnikman: Have been instrumental in the design of the study, data acquisition, analysis and interpretation and drafting the manuscript. Omrum Aydin, Hilde Herrema, Maurits de Brauw, Arnold van de Laar, Victor E. A. Gerdes: Have made substantial contributions in data acquisition and critically revising the manuscript. Thue W. Schwartz, Jens Nielsen, Fredrik Bäckhed, Folkert Kuipers, Max Nieuwdorp, Albert K. Groen: Have made contributions to the design of the study. They have made substantial contributions in interpretation of the data and critically revising the content. All authors have given final approval of the version to be published and take public responsibility for the content. They agreed to be accountable for all aspects of the work on ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

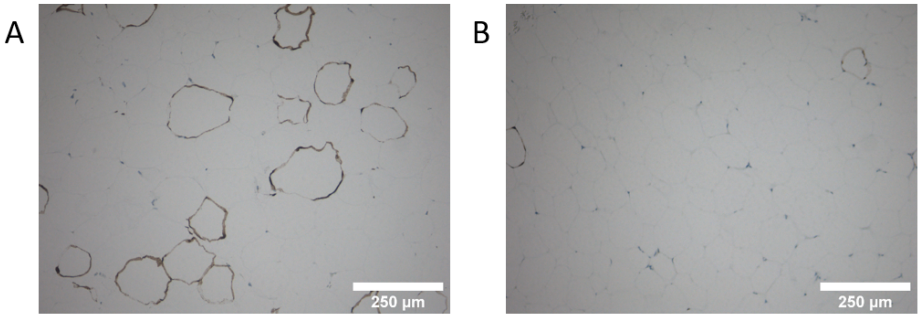
1. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. *Cell* **153**, 1194 (2013).
2. Ogrodnik, M., Salmonowicz, H. & Gladyshev, V. N. Integrating cellular senescence with the concept of damage accumulation in aging: Relevance for clearance of senescent cells. *Aging Cell* Preprint at https://doi.org/10.1111/acel.12841 (2019).
3. Gorgoulis, V. *et al.* Cellular Senescence: Defining a Path Forward. *Cell* Preprint at https://doi.org/10.1016/j.cell.2019.10.005 (2019).
4. Barzilai, N., Huffman, D. M., Muzumdar, R. H. & Bartke, A. The Critical Role of Metabolic Pathways in Aging. *Diabetes* **61**, 1315–1322 (2012).
5. Fontana, L., Partridge, L. & Longo, V. D. Extending healthy life span-from yeast to humans. *Science* Preprint at https://doi.org/10.1126/science.1172539 (2010).
6. Kenyon, C. J. The genetics of ageing. *Nature* Preprint at https://doi.org/10.1038/nature08980 (2010).
7. Holzenberger, M. *et al.* IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* (2003) doi:10.1038/nature01298.
8. Templeman, N. M. *et al.* Reduced Circulating Insulin Enhances Insulin Sensitivity in Old Mice and Extends Lifespan. *Cell Reports* (2017) doi:10.1016/j.celrep.2017.06.048.
9. Gustafson, B., Nerstedt, A. & Smith, U. Reduced subcutaneous adipogenesis in human hypertrophic obesity is linked to senescent precursor cells. *Nature Communications* (2019) doi:10.1038/s41467-019-10688-x.
10. Papatheodoridi, A.-M., Chrysavgis, L., Koutsilieris, M. & Chatzigeorgiou, A. The role of senescence in the development of non-alcoholic fatty liver disease and progression to non-alcoholic steatohepatitis. *Hepatology (Baltimore, Md.)* (2019) doi:10.1002/hep.30834.
11. Palmer, A. K. *et al.* Targeting senescent cells alleviates obesity-induced metabolic dysfunction. *Aging Cell* **18**, e12950 (2019).
12. Wang, L. *et al.* Targeting p21Cip1 highly expressing cells in adipose tissue alleviates insulin resistance in obesity. *Cell Metabolism* **34**, 75-89.e8 (2022).
13. Li, Q. *et al.* Obesity and hyperinsulinemia drive adipocytes to activate a cell cycle program and senesce. *Nature medicine* **27**, 1941–1953 (2021).
14. Van Olden, C. *et al.* A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *Journal of Internal Medicine* joim.13157 (2020) doi:10.1111/joim.13157.
15. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**, 525–527 (2016).
16. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
17. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (1995).
18. Dalla Man, C. *et al.* Two-Hour Seven-Sample Oral Glucose Tolerance Test and Meal Protocol: Minimal Model Assessment of -Cell Responsivity and Insulin Sensitivity in Nondiabetic Individuals. *Diabetes* **54**, 3265–3273 (2005).
19. Samuel, V. T. & Shulman, G. I. The pathogenesis of insulin resistance: Integrating signaling pathways and substrate flux. *Journal of Clinical Investigation* vol. 126 12–22 Preprint at https://doi.org/10.1172/JCI77812 (2016).
20. Sahra, I. Ben *et al.* Metformin, independent of AMPK, induces mTOR inhibition and cell-cycle arrest through REDD1. *Cancer research* **71**, 4366–4372 (2011).
21. Salaami, O. *et al.* Antidiabetic Effects of the Senolytic Agent Dasatinib. *Mayo Clinic Proceedings* **96**, 3021–3029 (2021).
22. Yang, Y.-K. *et al.* Human Mesenteric Adipose Tissue Plays Unique Role Versus Subcutaneous and Omental Fat in Obesity Related Diabetes. *Cellular Physiology and Biochemistry* **22**, 531–538 (2008).

23. Bigornia, S. J. *et al.* Relation of depot-specific adipose inflammation to insulin resistance in human obesity. *undefined* **2**, (2012).

24. Chan, D. C. *et al.* Adipose tissue compartments and insulin resistance in overweight-obese Caucasian men. *Diabetes Research and Clinical Practice* **63**, 77–85 (2004).

25. Ishikawa, M., Lourdes Pruneda, M., Adams-Huet, B. & Raskin, P. Obesity-independent hyperinsulinemia in nondiabetic first-degree relatives of individuals with type 2 diabetes. *Diabetes* (1998) doi:10.2337/diabetes.47.5.788.

SUPPLEMENTAL FIGURES



Supplemental figure 1. Representative images of adipose tissue immunohistochemistry. A shows high levels of P16 and B shows low levels of P16. Scale bar is shown in the lower left corner.

SUPPLEMENTAL TABLES

	NGT	Hyperglycaemia	metformin use
n	20	19	9
age (years)	40.0 [33.8, 49.2]	45.0 [38.0, 55.5]	47.0 [44.0, 51.0]
female (%)	17 (85.0)	15 (78.9)	6 (66.7)
BMI	39.5 [38.3, 40.8]	40.0 [37.3, 42.2]	38.5 [35.0, 40.2]
FPG (mmol/l)	5.2 [5.0, 5.3]	5.6 [5.3, 6.1]	7.1 [6.4, 8.6]
fasting insulin (pmol/l)	66.0 [45.8, 108.0]	85.0 [67.0, 145.2]	96.6 [47.7, 138.0]
glucose AUC (mmol/l*120 min)	713 [648, 785]	837 [776, 1019]	1189 [1113, 1384]
insulin AUC (nmol/l*120 min)	32 [27, 67]	63 [39, 92]	40 [26, 57]
HbA1c NGSP (%)	5.3 [5.3, 5.5]	5.8 [5.7, 6.0]	7.0 [6.6, 7.6]
HOMA2 IR	1.3 [0.8, 2.0]	1.6 [1.3, 2.8]	1.9 [0.9, 2.6]
Total cholesterol (mmol/l)	4.6 [4.1, 5.0]	5.0 [4.6, 5.8]	4.0 [3.5, 4.4]
HDL cholesterol (mmol/l)	1.1 [1.1, 1.4]	1.1 [0.9, 1.4]	1.2 [1.0, 1.2]
LDL cholesterol(mmol/l)	2.9 [2.5, 3.4]	3.3 [2.5, 3.9]	2.7 [2.2, 2.9]
triglycerides (mmol/l)	0.9 [0.7, 1.2]	1.4 [1.1, 1.9]	0.9 [0.9, 1.2]

Supplemental Table 1: Baseline characteristics of participants with biopsies taken. All values are displayed as median with inter quartile range. Weight loss = weight loss in the period between mixed meal test and surgery FPG = fasting plasma glucose. AUC = area under the curve

13

Summary

SUMMARY

Part 1. Systems biology in metabolic disease

Chapter 1 is an introduction to this thesis. In this thesis we used a systems biology approach to integrate panomics data to detect which pathway are dominant and which one are proximal in metabolic disease development. To move from association to causation, we performed intervention studies in humans with to manipulate the gut microbiota composition and functionality with the overarching aim to find new treatment targets for nonalcoholic fatty liver disease (NAFLD). Finally, we aimed to identify potential new drivers of cellular senenscence in metabolic disease.

Chapter 2 is a review, which summarizes the quality of the evidence of a causal role of the gut microbiome in the development of obesity and type 2 diabetes (T2D) at that time. Since the introduction of affordable next generation sequencing techniques, a plethora of studies have shown striking associations between the composition of the gut microbiome and obesity and T2D. Studies in humans using antibiotic therapy or fecal microbiota transplantation are suggestive of a causal relationship between the gut microbiome and metabolic disease development. However, effects size and evidence for causality are still marginal. Furthermore, these studies do not provide mechanistic insight into the interplay between the gut microbiome and host metabolism. Large prospective studies will be of critical importance to answer whether gut microbial composition is a reflection of the disease itself or the microbial composition was affected prior to disease development and hence was a driving factor. Thus, both prospective and intervention studies in large human cohorts combined with dedicated mechanistic studies in model systems are required to understand if and how gut the microbiome affects metabolic disease development.

In **Chapter 3** we aimed to reveal a microbial metagenomic signatures of obesity by using shotgun whole genome sequencing of individuals with a wide variety in BMI (i.*e.,* from lean to morbid obesity). A total number of 52 fecal bacterial species differed significantly in individuals with and without obesity. Moreover, we found that gut microbiome composition could explain more than 50% of the variance in clinical features of obesity such as BMI, waist, HDL and triglycerides. In addition, we observed differences in fecal microbial metabolic pathways in individuals with and without obesity, and particularly strong associations between amino acid metabolism, microbial species and obesity. Pathways involving biosynthesis of several amino acids, including histidine, lysine, and tryptophan were highly enriched in individuals with obesity, whereas pathways involved in the degradation of these amino acids, in particular histidine, were depleted. This suggests that the microbiome of individuals with obesity has a higher potential to produce several amino acids compared to subjects without obesity but have reduced capacity of catabolizing specific

amino acids. We thus, identified new bacterial species that are altered in individuals with obesity and also contribute to a possible robust consensus signature for the obese and lean gut microbiome.

Research into the role of the gut microbiome in the development of metabolic diseases is hampered by difficulties in obtaining biopsies from human tissues from the affected organs. Bariatric surgery provides the unique opportunity to obtain biopsies from several adipose tissue depots, small intestine, liver and even portal vein blood samples. In **Chapter 4** we describe the design and the aims of the BARIA study. Using this longitudinal bariatric surgery cohort, we aim to perform a systems biology approach identifying novel pathways in the pathogenesis of obesity, T2D and NAFLD which may be targets for future drug development. The aim of the BARIA study is to include 1500 individuals undergoing primary laparoscopic bariatric surgery. Before surgery, they are subjected to a 2-h mixed meal test (MMT) to assess insulin resistance and investigate dynamic alterations in circulating metabolites, blood and fecal sampling, and questionnaires, including psychology lists taken at the start of the MMT in all individuals at all time-points to minimize variation. During surgery, biopsies are obtained from three fat depots, jejunum, liver and samples from portal and peripheral venous blood. Thereafter, further sampling (MMT, blood and faecal samples) is performed. In the event of another surgery (revisional surgery, cholecystectomy) further biopsies can be obtained, which is included in the ethical protocol. We process tissues for RNA-sequencing, analyse gut microbiota and perform untargeted (postprandial) plasma metabolomics on both fasting and postprandial MMT plasma samples. By using a systems biology, we aim to integrate all the omics data to gain insight into the hierarchy of mechanisms underlying the development of metabolic diseases. In the next three chapters we used both clinical and omics data of the first 106 participants from the BARIA study of whom a complete omics data set was available.

In **Chapter 5**, we developed a novel concept for stratification of individuals with obesity based on fasting plasma metabolome data. This framework enables a non-biased stratification of individuals with obesity rather than purely clinical parameters that may fail to accurately encompass the multitude of nuances in human population-based studies, particularly when studying complex and multigenic diseases such as obesity and T2D. For example, classifiers for obesity-associated comorbidities such as hypertension, T2D, NAFLD and dyslipidemia may be treated as binary variables (present *vs.* absent), however the overall wellness of an individual with any of these disorders can vary significantly as a function of how well managed each of these conditions are, among many other factors. Metabolomics, in the context of obesity and cardiometabolic disease, is consistently being used as a means of evaluating metabolic health, measuring the effect of dietary intervention strategies and to identify predictive biomarkers characterizing a specific condition. In this light, use of

biologic parameters such as the plasma metabolome, as a direct readout of the overall status of an entire multiorgan system host and its microbiome, to determine grouping of individuals, rather than traditional clinical disease classifiers, offers a unique approach that may more accurately classify individuals into distinct disease physiological states. This method generated five distinct metabotypes for the BARIA bariatric surgery cohort. Grouping participants into five metabotypes reduced the dimensionality and enabled us to reveal associations and links between specific microbes, metabolites and transcriptome signatures of adipose and liver tissue. Our findings suggested that participants' stratification based on metabotyping could enhance our ability to get molecular insights into the causes of diseases from panomics integrative analysis

In **Chapter 6**, we performed plasma metabolomic profiling on both fasting and post-prandial samples and investigated global metabolic responses to an MMT. In contrast to chapter 5, we grouped the individuals based on their glucose tolerance in normal glucose tolerance (NGT), pre-diabetes and T2DM groups. The plasma metabolome revealed that the number of metabolites that differ significantly between these groups was most pronounced between the NGT and T2DM groups in the post prandial plasma samples compared to fasting condition, thus enabling us to discover abnormal metabolism related to (pre)diabetes that did not appear at fasting condition. Furthermore, we traced the differences in metabolic responses back to other omics sets including fecal metagenomics and transcriptomics data of liver, adipose tissue and jejunum. Finally, to further investigate possible driving factors for postprandial glucose regulation, we predicted glucose area under the curve (AUC) based on panomics data using ridge regression models with 5-fold cross-validation. Both fasting and post-prandial metabolomics data performed best to predict the actual glucose AUC. Amongst the most important metabolites for the prediction of glucose AUC were phenylalanine and 1–carboxyethylphenylalanine. The performance of the gut microbiome to predict glucose AUC was modest. We thus, identified possible new biomarkers for glycemic control including phenylalanine derived metabolites and show that that there is a metabolic imbalance in individuals with (pre)diabetes.

Chapter 7 focusses on the most prevalent chronic liver disease worldwide, NAFLD. In the general population, one in four individuals is affected by NAFLD, but the prevalence increases to over 80% in individuals with obesity. The rapidly growing prevalence of NAFLD and lack of effective treatment options to tackle this potentially debilitating disease, will further increase obesity-related burden on public health and economies. In order to develop appropriate, non-invasive diagnostic methods and treatment options, it is critical to deeply investigate the complex pathophysiology of NAFLD. Here, we used a systems biology approach to investigate the contribution of different organs to this disease. We analyzed transcriptomics profiles of liver and adipose tissues, fecal metagenomes and

plasma metabolomes of 55 women with and without early stage NAFLD. We observed significant differences in metabolites, expression of human genes and gut microbial features in individuals with and without NAFLD. In addition, by developing a multivariate model, we revealed that there is substantial crosstalk between these different omics sets. Our study provides a comprehensive panomics analysis of individuals with early stage NAFLD, providing a novel strategy to study the pathophysiology of NAFLD in humans.

Part 2. From association to causation

In this part, we used the steppingstones provided in part 1 to move from association to causation by performing interventions studies that manipulate the gut microbiome in individuals with NAFLD. In **Chapter 8** we showed that the human gut microbiota produces large amounts of ethanol that might be clinically relevant for the pathogenesis of NAFLD by performing four distinct different experiments. We obtained portal vein blood, which is the most relevant blood vessel to study microbial produced metabolites because blood from the gastrointestinal tract drains directly into this vein. High concentrations of ethanol were found in the portal vein and significantly correlated with NAFLD parameters. In addition, we also observed that ethanol in the peripheral circulation increased during a mixed meal. Therefore, ethanol production during a mixed meal test should be considered as a non-invasive diagnostic approach for the detection of high ethanol producing gut microbiomes and NAFLD risk. High post prandial plasma ethanol concentrations correlated particularly with high relative fecal abundance of lactic acid bacteria. By circumventing the first pass-effect by inhibiting alcohol dehydrogenase an almost 15 times increase in peripheral ethanol was observed in patients with NAFLD. This effect was abolished after treatment with broad spectrum antibiotics. Hence, we obtained causal evidence that the human gut microbiota can produce large amounts of ethanol. To what extent persistent endogenous ethanol production is causally involved in the highly complex pathogenesis of NAFLD where a combination of environmental factors, genetic variants, obesity and disturbed lipid homeostasis interact, remain to be elucidated.

In **Chapter 9** we describe the results of a double-blind randomized controlled proof-of-principle study in which individuals with NAFLD on ultrasound were randomized to two study arms; lean vegan donor (allogenic n=10) or own (autologous n=11) fecal microbiota transplantation (FMT), which were performed three times at eight-week intervals. A liver biopsy was performed at baseline and after 24 weeks in every subject to determine histopathology (NASH-CRN) classification and changes in hepatic gene expression based on RNA sequencing. Although the study was underpowered, allogenic FMT from vegan donors on a plant-based, low animal protein diet decreased the necro-inflammation score, also known as disease activity (i.e. hepatocyte inflammation and ballooning) independent

of steatosis is clinically the most relevant parameter of NAFLD in paired liver biopsies. In addition, allogenic FMT showed an effect on intestinal microbiota composition, which was associated with both beneficial changes in plasma metabolites and the expression of liver genes involved in inflammation and lipid metabolism after donor FMT.

Part 3. Cellular senescence, an old but new player in metabolic disease

Here, we report on an old but new player in metabolic disease; cellular senescence. Cellular senescence is a state of irreversible cell cycle arrest with important physiological functions. However, cellular senescence is also a hallmark of ageing and has been associated with several pathological conditions. **Chapter 10** is a review which summarizes the quality of the evidence of a causal role of cellular senescence in the development of NAFLD. In this review, we provide insight in the quality of the evidence that is supportive of a causal role of cellular senescence in the development of NAFLD in rodents and humans. We also elaborate on defining and discussing key cellular and molecular features of senescence. Finally, we discussed the efficacy and safety of novel senolytic drugs to treat or even prevent this disease.

In **Chapter 11**, we used clinical data, portal vein plasma, immunohistochemistry, and transcriptomics data derived from individuals with obesity to establish a link between insulin and senescence in the liver. We found strong correlations between markers of cellular senescence in the liver with circulating levels of insulin. Of specific interest, a very strong correlation between portal vein insulin and hepatocytic senescence was found. In addition, these correlations were independent of NAFLD. This observation suggests that hepatocytic senescence, driven by prevailing insulin concentrations, precedes NAFLD. By using a validation cohort consisting of 180 individuals, we were able to replicate the findings that markers of cellular senescence in the liver are strongly associated with fasting plasma insulin concentrations. In **Chapter 12,** we used the same approach as in the previous chapter by using clinical, immunohistochemistry, and transcriptomic data derived from non-diabetic individuals with obesity, we have shown that peripheral insulin resistance is highly correlated with markers of cellular senescence in mesenteric adipose tissue prior to the presence of glycaemic dysregulation. These results confirm previous studies indicating that adipocytic cellular senescence may play an important role already in the earliest stages of insulin resistance before the onset of T2DM Collectively, these chapters suggest that insulin might be inducing cellular senescence in different organs via distinct pathways, an observation that underscores the importance of lowering insulin secretion in obese individuals with insulin resistance.

14

General discussion and perspectives

GENERAL DISCUSSION AND PERSPECTIVES

The research in this thesis has tried to gain new insight into the pathogenesis of obesity, insulin resistance and nonalcoholic fatty liver disease (NAFLD). As appetizer, we used a systems biology approach with emphasis on the putative role of gut microbiota and observed remarkable associations between gut microbial composition, plasma metabolites and human genes and metabolic diseases. For the main course, we used the steppingstones that were provided by the comprehensive panomics analyses to move from association to causation by performing interventions studies that manipulate the gut microbiome in humans. For dessert, we investigated a new but old player in metabolic disease, cellular senescence. In contrast to the current paradigm, we revealed that insulin rather than glucose is associated with cellular senescence in metabolic disease.

The human gut microbiome, more than just a filter between the gut and host

Over the past decades, numerous studies reported associations between the gut microbiome and numerous diseases ranging from metabolic to psychiatric disorders[1,2]. High-throughput sequencing combined with machine learning methods such as mendelian randomization and mediation analyses, suggest causal relations between the gut microbiome and (metabolic) diseases, without performing an intervention trial. Yet, is it real? Or are the expectations distilled from the thousands of publications per year inflated, and is the truth still hidden in the shadow?

The gut microbiome is an extraordinarily complex ensemble, containing organisms that span several kingdoms[3]. Tremendous amount of research has been performed since Antony van Leeuwenhoek reported "small animalcules very swiftly moving" in his stools and dental plaque when looking through his microscope more than three centuries ago[4]. However, simple, and obvious questions frequently asked by patients (and researchers themselves) such as which microbes are good and which are bad, cannot be answered as straightforward as one would expect. Many researchers have struggled to define a "healthy gut microbiome"[5,6]. What is healthy? There is no universally accepted definition of health, and health can thus be considered as a relative condition. Moreover, according to studies in large human cohorts progressively including "healthy" participants, it became clear that variation in the gut microbiome is extensive[7]. However, most of the variance of the human gut microbiome is still unaccounted for[8–10]. To define a healthy but also an unhealthy gut microbiome, this challenge must be overcome. The term "dysbiosis", frequently used when the gut microbiome composition is associated with metabolic disease, implies that there is an understanding of a normal or healthy gut microbiome, and should therefore not be used[11]. As of to date, novel high-throughput sequencing methods can detect approximately 100-200 bacterial species[12],

whereas the human gut may represent up to 2000 bacterial species[13]. Part of the variance in the gut microbiome may be stochastic, yet several intrinsic and extrinsic factors such as host genetics, disease state, immune health, diet, socio-economic status, location, and medication are known to determine individual gut microbiomes[3]. Medication history should therefore be accounted for in any reference norm for the gut microbiome[11]. In part 1 of this thesis, we found alterations in the gut microbiome composition between individuals affected by either one or a combination of obesity, insulin resistance and NAFLD. So far, although of value, we only demonstrated differences. As customary for microbiome papers, reviews but also science in general, we concluded in these papers that we should move from correlation to causation, from observation to mechanism and from cross-sectional to longitudinal studies.

It is noteworthy, that most human gut microbiome studies are limited to use fecal rather than intestinal bacteria. The gastrointestinal tract is very heterogenous, and although limited studies are performed with upper gastrointestinal tract samples (i.e., small intestine), it is known that gut microbial diversity increases towards the colon[14]. Even in the colon, the diversity varies depending on the studied segment[14]. The fecal microbiome is an end-product that is a result of a dynamic process along the gastrointestinal tract[3]. Certainly, it provides insights into the general shifts within the gut microbiome, but species that are dominant throughout the gastrointestinal tract are not always detected in the feces[15], underscoring the need to sample along the gastrointestinal tract. Bearing in mind that laxatives have profound influence on gut microbiome composition, should gastroenterologists go for the dark (or brown) matter while we wait for ingestible devices that facilitate sampling from sites along the gastrointestinal tract?

Another customary ending of papers and reviews in the microbiome field is that we should focus on functionality instead of composition. Metabolic actions on distal tissues and organs by the gut microbiome are exerted amongst others via microbial metabolites[16]. Human plasma serves as a liquid conveyor for molecules inside the body. The thousands of circulating small molecules, collectively called the plasma metabolome provides a unique insight into the interactions of genetics, lifestyle, environment, medication use and microbial activity[17]. Notably, according to recent reports using rather sophisticated statistical analyses approximately 60% of the variance in the plasma metabolome can be explained by the gut microbiome[9]. The plasma metabolome can thus be used as a read-out of the functionality of the gut microbiome. Upon ingestion of nutrients, the gut microbiome determines which metabolites are formed and absorbed[18]. Microbial metabolites modulate many key features of metabolic diseases such as insulin resistance[19], platelet hyperreactivity[20], thrombosis potential[21] and atherogenic lipid profile[22], suggesting that the gut microbiome contributes to different metabolic diseases. The significance in disease development, progression or outcome nevertheless is still subject of debate. A key example is our ethanol paper described

in this thesis[23]. Although we obtained casual evidence that the gut microbiome can produce large amounts of ethanol, the direct impact of microbial ethanol on NAFLD was not assessed. According to our calculations, the daily amount of ethanol produced in the gut should be sufficient to induce profound alterations in hepatic metabolism including fatty acid oxidation, gluconeogenesis, and inflammation[24]. However, the paper has not identified the specific bacterial strains responsible for the ethanol production, nor has it unraveled under which conditions these strains switch to ethanol production. Nevertheless, it was already shown almost a decade ago that several bacterial strains can switch from ethanol to other microbial metabolites depending on the environments' redox balance and available substrate[25]. This suggests that ethanol production by bacterial strains can be altered and shifted towards less harmful or even beneficial metabolites. These exciting observations underline the opportunities of identifying therapeutic targets within the gut.

Targeting the gut microbiome

Considerable advances in gut microbial science have been made over the past decades including relationships between Helicobacter pylori and peptic ulceration[26], treating recurrent Clostridioides difficile infection via fecal microbiota transplantation (FMT)[27], and more recently the relationship between responsiveness to checkpoint inhibitors and gut microbiome composition in cancer[28]. For several other diseases, the microbiome seems an attractive and feasible target as the effects are restricted to the gut lumen, with beneficial systemic effects but with minimal risk of systemic off target effects.

Roughly, interventions can be divided into targeted and untargeted therapies. Untargeted therapies include FMT, fecal filtrate transplant (FFT) or oral administration of probiotics. The aim of FMT is to change a recipient's microbiome for therapeutic purposes[29]. In Part 2 of this thesis, we performed a proof of principle trial in patients with NAFLD whom either received an autologous or an allogenic FMT[30]. A trend towards improvement of lobular inflammation and hepatocyte ballooning in the liver biopsy was observed after treatment with allogenic FMT. Considering that from a clinical perspective, an FMT trial is successful if improvement or remission of a disease is induced, this trial was negative. From an ecological perspective, however, the extent to which the donor's microbiota can colonize the recipient microbiome is more important[31]. A recent study, reanalyzing 316 FMTs derived from a wide range of indications, suggest, however, that clinical success is not dependent on colonization of donor strains, displacement of recipient species or the reinstatement of specific bacterial functions[31]. This is line with the success of the autologous FMT, in both type 1 diabetes and inflammatory bowel disease[32,33]. Predicting the outcome of FMT, from a clinical and ecological view remains difficult. Recent advances suggest that recipient factors are more important than donor factors, which is in contrast to the concept of super-

donor's, representing individuals with a highly diverse microbiome that were considered most effective FMT donors[34]. Complementarity of donor and recipient microbiomes on community level and to specific strain population similarity are crucial to colonization and by proxy, clinical success. Matching donor-recipient microbiomes on community, species and strain levels could increase the success of colonization and hence clinical success and is therefore warranted[31]. In addition, there is little focus on the mycobiome and virome, which represent non-microbial inhabitants of the gut with significant impact on bacteria, which complicated the understanding of interactions between the different ecosystems in the gut and host[35,36]. To further increase progress in facilitating the implementation of microbiome-based interventions, focus should be on unraveling these unknown factors.

Probiotics are defined as "live microorganisms that, when administered in adequate amounts, confer a health benefit on the host"[37]. Probiotics come in a range of complexities but traditionally, probiotics consisted of single strains[38]. Advances in anaerobic culturing, in combination with improved sequencing technologies, enabled the production of strains for specific conditions and have resulted in a wider availability of strains and have been termed 'next generation probiotics'[39]. Strains derived from *Faecalibacterium prausnitzii*[40], *Bacteroides fragilis*[41], *Anaerobutyricum soehngenii*[42] and *Akkermansia muciniphilia*[43] are abundantly studied using next generation probiotics, the latter being the most dominant player in the field. Nevertheless, so far, human studies have failed to show a clinically relevant benefit of next generation probiotics use. Considering that most of these strains are strictly anaerobic, the lack of clinical success is often explained by failure of engraftment because the viability is reduced in the small intestine[44]. To overcome these challenges, the production of multiple bacterial strains into a single probiotic, whereby interactions can be directed to increase the success of engraftment or production of the desired metabolite[44,45]. These so-called multi-strain consortia or engineered microbes thereof increases the likelihood of achieving a specific clinical target.

Targeted therapies include bacteriophages (or phages) and post-biotics, which can influence microbiota-host, trans-kingdom, and inter-bacterial interactions[35,38]. Phages are viruses that target bacteria and might aid in depletion of a certain species within the ecosystem[35]. This strategy has shown beneficial effects in patients with Alcoholic Fatty liver disease by reducing the circulation of the detrimental protein cytolysin, via targeting the cytolysin producing bacteria *Enterococcusus feacalis*[46]. Most phages have a narrow host range, meaning they infect closely related strains within (related) species, limiting collateral damage to the microbiome of the recipient, via infecting other species[35]. Nevertheless, this narrow range might also be a reason why the ability of phages to modulate the gut microbiome is limited[35].

Postbiotics are bioactive molecules produced by bacteria and are now formally defined as a "preparation of inanimate micro-organisms and/or their components that confer a health benefit on the host"[47]. The preparation, however, is not alive or viable, preventing the chance of colonization. Hence, the possible health benefits conferred by the postbiotics depend on the regular intake to maintain the presence of the functionally bioactive molecules. Examples of postbiotics are short-chain fatty acids, secondary bile acids, but also pasteurized *A. muciniphila*[44]. All have shown their benefit on human metabolism in varying degrees, but none of them is the holy grail to cure metabolic diseases. Of interest, especially for individuals affected by or prone to develop NAFLD, the end products of the mixed acid fermentation are examples of postbiotics. The mixed acid fermentation pathway is the biological process in which under anaerobic conditions sugars are converted into a complex and variable mixture of acids including lactate, acetate, succinate, formate and ethanol[25,48]. This metabolic pathway is common in bacteria including Gram-negative (members of Enterobacteriaceae) and Gram-positive including members belonging to lactic acid bacteria (LAB)[25,48]. The formation of these gut microbial metabolites depends on the presence of certain key enzymes in the gut microbiota and the amount of oxidized Nicotinamide adenine dinucleotide (NAD). The first step is a glycolysis reaction where glucose is converted into pyruvate and where NAD is reduced to NADH. Pyruvate is then converted into acetyl-CoA and subsequently via the enzyme Alcohol Dehydrogenase (ADH) and oxidation of NADH to NAD, ethanol is produced. The variety in end products, dependence on NAD and that the balance between end products is not "fixed", suggests that the process can be altered and thus shifted towards different end products when the environment or redox potential is changed. Finding the right postbiotic to alter the redox balance and thereby reducing ethanol production is an interesting approach and warrants further research.

I predict that in the coming decades an explosion of so-called personalized microbiome companies will arise that promises to beneficially alter and individuals' microbiome. This is rather amusing, because at first, we still don't know what a healthy microbiome is, as explained previously but secondly, diet is the most successful intervention to change the gut microbiome composition[49,50]. Certainly, the correct dietary advice should be given preferably via a personalized microbiome medicine approach[49,51]. However, patients don't like to diet, and diet is notorious for not persevering. I wonder, will patients use any of the above-mentioned therapies just because scientist claim it is good for their microbiome and proxy, their health, just like diet? Nevertheless, in the already affected individuals or as secondary prevention, I expect that it will be difficult to find a one-size fits all treatment strategy because metabolic diseases are heterogenic, and the natural history and clinical features are different between patients[18]. Focus should therefore be first on identifying and secondly on validating mechanisms that contribute to the pathophysiology of the disease in humans and

adopt treatments in a personalized fashion. Therefore, in-depth phenotyping using omics data of patients before and after intervention is needed, that can be used to predict the response to a specific intervention[18]. For example, individuals with a deficiency in butyrate producing bacteria who are insulin resistant, could benefit from supplementation with the missing microbrobes in the form of probiotics or the metabolite itself[18]. Whereas individuals with elevated levels of harmful metabolites such as ethanol or phenylacetylglutamine, might respond better from supplementation with inhibitors, specifically designed to target microbial enzymes involved in the production of these metabolites[18].

The therapeutic landscape of nonalcoholic fatty liver disease

The prevalence of NAFLD have reached pandemic proportions, underscoring the need to unravel its pathophysiology and the risks associated with the condition. The presence of steatosis in the liver in the absence of significant fibrosis has long been considered a relative benign condition. It is now widely accepted that liver fibrosis as a result of liver injury secondary to NAFLD, is a major prognostic predictor for liver-related and overall morbidity and mortality[52,53]. Our understanding of the factors that determine disease progression has evolved, but we are still not able to identify those patients who will progress to a more advanced stage in the disease and those who will not[54]. In line with other metabolic diseases, it is becoming increasingly apparent that the molecular and cellular processes driving NAFLD are highly heterogenous from one patient to the next[55]. I expect therefore that it will be difficult to find a one-size fits all treatment strategy for NAFLD. This understanding raises the question; why are we searching for the one and only predictive core signature, as we also tried in this thesis? We should put aside the aim to find the holy grail among signatures and biomarkers for NAFLD and embrace the heterogeneity of the disease. If there is one metabolic disease that is a suitable target for a precision medicine approach, identifying subpopulations and eventually raising the possibilities to match the therapeutic strategy to the disease drivers specific for the affected individual, it is NAFLD. The development of such precision medicine approaches will require a large amount of panomic data to create a 'Liver Atlas', that is likely to have a transformative effect on the NAFLD field like the Cancer Genome Atlas project[56]. To create such an atlas, we should regularly perform biopsies, not because gastroenterologists like to perform biopsies, but because it is necessary to better classify affected individuals. Obtained liver biopsies can be used to deeply phenotype patients on single cell or spatial resolved omics level, before and after an intervention, which can be used to predict the response to a specific intervention. Therefore, I believe that the liver biopsy will not be replaced in the coming decades for a certain biomarker for the diagnoses and thus molecular pattern, specific for the patient. It is however undeniable that we need biomarkers to track changes over time and to evaluate the impact of specific interventions[57].

Obviously, this is crucial when we are trying to target appropriate therapies to those who are most likely to benefit[57].

Although there is no registered treatment, the future therapeutic landscape is enriched with an impressive range of agents with mechanisms of action that target different factors of the pathogenesis of NAFLD[57]. In parallel with the current treatment strategies for hypertension and T2D, the future of NAFLD pharmacotherapy will certainly include combination therapies[54]. Unfortunately, this inevitably will lead to polypharmacy and thereby decreases treatment adherence and increases the risk of adverse events and interactions with other drugs. Finding a cure that targets different elements or the disease itself instead of controlling concomitant metabolic diseases is therefore warranted. Homeostasis and the microenvironment of the liver in NAFLD is remarkably different compared to healthy livers. In the whole spectrum of the disease (i.e., from steatosis to cirrhosis), the hepatocyte exhibits one or more of the following hallmarks: prolonged and generally irreversible cell cycle arrest, macromolecular damage, secretory features and deregulated metabolism[58]. Hence, part of the hepatocytes in individuals with NAFLD are considered to be senescent. Many of the human cells that have a role in metabolic disease are postmitotic, and the repercussions of postmitotic cellular senescence on tissue and health function are ill-defined[59]. The role of cellular senescence, however, is specifically relevant for liver disease because hepatocytes are considered to be reversed postmitotic cells that preserve their proliferating potential enabling that damaged hepatocytes can be replaced by healthy cells[58]. However senescent cells depend on senescence-associated pathways (SCAPs), which makes senescent cells resistant to apoptosis[60]. SCAPs were thus identified as the "Achilles heel" of senescent cells and since this discovery, intense research has focused on identifying molecules that selectively induce apoptosis in senescent cells[61]. This has been successful and, it is now possible to specifically target senescent cells. Remarkably, according to *in vitro* data, a brief disruption of pro-survival pathways is sufficient to clear senescent cells[61,62]. Because senescent cells take weeks to reaccumulate, treatment aimed at clearing senescence cells (senolytics) can be administrated intermittently, which reduces the risk of adverse effects compared to continuous treatment[61,62].

Multiple senolytic candidates have been discovered and are among currently available drugs to target senescent cells in a multitude of diseases. Dasatinib, which is an EMA-approved oral tyrosine kinase inhibitor and the antioxidant quercetin, which is a flavonoid; present in many fruits and vegetables, successfully induce apoptosis in senescent cells in rodent and human studies improving meaningful clinical outcomes in patients with idiopathic lung fibrosis and diabetic associated kidney disease[63,64]. The combination of dasatinib and quercetin acts on multiple SCAP network nodes and thereby increases the potency towards clearing senescent cells, including in the liver[62]. Despite these exciting translational

findings, we must wait for the results of studies assessing the safety and efficacy of this drug combination in humans. A proof-of-principle trial with the senolytic combination dasatinib and quercetin in patients with liver fibrosis is already on its way. While we wait, the key to significantly improve the care of patients with NAFLD lies in adopting a multidisciplinary approach combining skills and expertise of multiple (para) medical professions combined with dedicated research to find the right treatment for the right patient.

REFERENCES

1. Manor, O. *et al.* Health and disease markers correlate with gut microbiome composition across thousands of people. *Nature Communications 2020 11:1* **11**, 1–12 (2020).
2. Chakaroun, R. M., Olsson, L. M. & Bäckhed, F. The potential of tailoring the gut microbiome to prevent and treat cardiometabolic disease. *Nature Reviews Cardiology 2022* 1–19 (2022) doi:10.1038/s41569-022-00771-0.
3. Meijnikman, A. S., Gerdes, V. E., Nieuwdorp, M. & Herrema, H. Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. *Endocr Rev* **39**, 133–153 (2018).
4. Lane, N. The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140344 (2015).
5. Bäckhed, F. *et al.* Defining a healthy human gut microbiome: Current concepts, future directions, and clinical applications. *Cell Host Microbe* **12**, 611–622 (2012).
6. Eisenstein, M. The hunt for a healthy microbiome. *Nature* **577**, (2020).
7. Zhernakova, A. *et al.* Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science (1979)* **352**, 565–569 (2016).
8. Asnicar, F. *et al.* Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat. Med.* **27**, 321–332 (2021).
9. Chen, L. *et al.* Influence of the microbiome, diet and genetics on inter-individual variation in the human plasma metabolome. *Nature Medicine 2022 28:11* **28**, 2333–2343 (2022).
10. Hughes, D. A. *et al.* Genome-wide associations of human gut microbiome variation and implications for causal inference analyses. *Nature Microbiology 2020 5:9* **5**, 1079–1087 (2020).
11. Shanahan, F., Ghosh, T. S. & O'Toole, P. W. The Healthy Microbiome—What Is the Definition of a Healthy Gut Microbiome? *Gastroenterology* **160**, 483–494 (2021).
12. Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science (1979)* **341**, (2013).
13. Almeida, A. *et al.* A new genomic blueprint of the human gut microbiota. *Nature 2019 568:7753* **568**, 499–504 (2019).
14. Zmora, N. *et al.* Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* **174**, 1388-1405.e21 (2018).
15. Vaga, S. *et al.* Compositional and functional differences of the mucosal microbiota along the intestine of healthy individuals. *Scientific Reports 2020 10:1* **10**, 1–12 (2020).
16. Herrema, H. & Niess, J. H. Intestinal microbial metabolites in human metabolism and type 2 diabetes. *Diabetologia* 1–15 Preprint at https://doi.org/10.1007/s00125-020-05268-4 (2020).
17. Talmor-Barkan, Y. *et al.* Metabolomic and microbiome profiling reveals personalized risk factors for coronary artery disease. *Nat Med* **28**, (2022).
18. Olofsson, L. E. & Bäckhed, F. The Metabolic Role and Therapeutic Potential of the Microbiome. *Endocr Rev* **43**, 907–926 (2022).
19. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).
20. Nemet, I. *et al.* A Cardiovascular Disease-Linked Gut Microbial Metabolite Acts via Adrenergic Receptors. *Cell* (2020) doi:10.1016/j.cell.2020.02.016.
21. Tang, W. H. W. *et al.* Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. *New England Journal of Medicine* (2013) doi:10.1056/NEJMoa1109400.
22. Haghikia, A. *et al.* Propionate attenuates atherosclerosis by immune-dependent regulation of intestinal cholesterol metabolism. *Eur Heart J* **43**, 518–533 (2022).
23. Meijnikman, A. S. *et al.* Microbiome-derived ethanol in nonalcoholic fatty liver disease. *Nature Medicine 2022 28:10* **28**, 2100–2106 (2022).
24. Wilson D.F, M. F. M. Ethanol metabolism: The good, the bad, and the ugly. *Med Hypotheses* **140**, (2020).

25. Elshaghabee, F. *et al.* Ethanol Production by Selected Intestinal Microorganisms and Lactic Acid Bacteria Growing under Different Nutritional Conditions. *Front Microbiol* **7**, (2016).

26. Kodaman, N. *et al.* Human and Helicobacter pylori coevolution shapes the risk of gastric disease. *Proc Natl Acad Sci U S A* **111**, 1455–1460 (2014).

27. van Nood, E. *et al.* Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med* **368**, 407–415 (2013).

28. Morad, G., Helmink, B., Sharma, P. & al., et. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell* **184**, 5309–5337 (2021).

29. Hanssen, N. M. J., de Vos, W. M. & Nieuwdorp, M. Fecal microbiota transplantation in human metabolic diseases: from a murky past to a bright future? *Cell Metab.* **33**, 1098–1110 (2021).

30. Witjes, J. J. *et al.* Donor Fecal Microbiota Transplantation Alters Gut Microbiota and Metabolites in Obese Individuals With Steatohepatitis. *Hepatol Commun* hep4.1601 (2020) doi:10.1002/hep4.1601.

31. Schmidt, T. S. B. *et al.* Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nature Medicine 2022 28:9* **28**, 1902–1912 (2022).

32. de Groot, P. *et al.* Donor metabolic characteristics drive effects of faecal microbiota transplantation on recipient insulin sensitivity, energy expenditure and intestinal transit time. *Gut* (2019) doi:10.1136/gutjnl-2019-318320.

33. Basson, A. R., Zhou, Y., Seo, B., Rodriguez-Palacios, A. & Cominelli, F. Autologous fecal microbiota transplantation for the treatment of inflammatory bowel disease. *Transl Res* **226**, 1–11 (2020).

34. Wilson, B. C., Vatanen, T., Cutfield, W. S. & O'Sullivan, J. M. The Super-Donor Phenomenon in Fecal Microbiota Transplantation. *Front Cell Infect Microbiol* **9**, (2019).

35. Duan, Y., Young, R. & Schnabl, B. Bacteriophages and their potential for treatment of gastrointestinal diseases. *Nat Rev Gastroenterol Hepatol* **19**, 135–144 (2022).

36. Demir, M. *et al.* The fecal mycobiome in non-alcoholic fatty liver disease. *J Hepatol* **76**, 788–799 (2022).

37. Hill, C. *et al.* Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol* **11**, 506–514 (2014).

38. Bajaj, J. S., Ng, S. C. & Schnabl, B. Promises of microbiome-based therapies. *J Hepatol* **76**, 1379–1391 (2022).

39. O'Toole, P. W., Marchesi, J. R. & Hill, C. Next-generation probiotics: the spectrum from probiotics to live biotherapeutics. *Nature Microbiology 2017 2:5* **2**, 1–6 (2017).

40. Ueda, A. *et al.* Identification of Faecalibacterium prausnitzii strains for gut microbiome-based intervention in Alzheimer's-type dementia. *Cell Rep Med* **2**, (2021).

41. Wang, Y. *et al.* Safety evaluation of a novel strain of Bacteroides fragilis. *Front Microbiol* **8**, 435 (2017).

42. Koopen, A. *et al.* Duodenal Anaerobutyricum soehngenii infusion stimulates GLP-1 production, ameliorates glycaemic control and beneficially shapes the duodenal transcriptome in metabolic syndrome subjects: a randomised double-blind placebo-controlled cross-over study. *Gut* **71**, 1577–1587 (2022).

43. Depommier, C. *et al.* Supplementation with Akkermansia muciniphila in overweight and obese human volunteers: a proof-of-concept exploratory study. *Nat Med* **25**, 1096–1103 (2019).

44. Hitch, T. C. A. *et al.* Microbiome-based interventions to modulate gut ecology and the immune system. *Mucosal Immunology 2022 15:6* **15**, 1095–1113 (2022).

45. Moens, F. *et al.* A four-strain probiotic exerts positive immunomodulatory effects by enhancing colonic butyrate production in vitro. *Int J Pharm* **555**, 1–10 (2019).

46. Duan, Y. *et al.* Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease. *Nature* **575**, 505–511 (2019).

47. Salminen, S. *et al.* The International Scientific Association of Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of postbiotics. *Nat Rev Gastroenterol Hepatol* **18**, 649–667 (2021).

48. Förster, A. H. & Gescher, J. Metabolic engineering of Escherichia coli for production of mixed-acid fermentation end products. *Front Bioeng Biotechnol* **2**, 16 (2014).

49. Leshem, A., Segal, E. & Elinav, E. The Gut Microbiome and Individual-Specific Responses to Diet. *mSystems* **5**, (2020).

50. Schmidt, T. S. B., Raes, J. & Bork, P. The Human Gut Microbiome: From Association to Modulation. *Cell* **172**, 1198–1215 (2018).

51. Suez, J. *et al.* Personalized microbiome-driven effects of non-nutritive sweeteners on human glucose tolerance. *Cell* **185**, 3307-3328.e19 (2022).

52. Targher, G., Byrne, C. D. & Tilg, H. NAFLD and increased risk of cardiovascular disease: Clinical associations, pathophysiological mechanisms and pharmacological implications. *Gut* Preprint at https://doi.org/10.1136/gutjnl-2020-320622 (2020).

53. Sanyal, A. J. *et al.* Prospective Study of Outcomes in Adults with Nonalcoholic Fatty Liver Disease. *New England Journal of Medicine* **385**, 1559–1569 (2021).

54. Marjot, T., Moolla, A., Cobbold, J. F., Hodson, L. & Tomlinson, J. W. Nonalcoholic Fatty Liver Disease in Adults: Current Concepts in Etiology, Outcomes, and Management. *Endocr Rev* **41**, 66–117 (2020).

55. Ratziu, V., Francque, S. & Sanyal, A. Breakthroughs in therapies for NASH and remaining challenges. *J Hepatol* **76**, 1263–1278 (2022).

56. Sanyal, A. J. Past, present and future perspectives in nonalcoholic fatty liver disease. *Nature Reviews Gastroenterology & Hepatology 2019 16:6* **16**, 377–386 (2019).

57. Francque, S. M., van der Graaff, D. & Kwanten, W. J. Non-alcoholic fatty liver disease and cardiovascular risk: Pathophysiological mechanisms and implications. *Journal of Hepatology* Preprint at https://doi.org/10.1016/j.jhep.2016.04.005 (2016).

58. Meijnikman, A. S. *et al.* Evaluating causality of cellular senescence in Non-Alcoholic Fatty Liver Disease. *JHEP Reports* **0**, 100301 (2021).

59. Sapieha, P. & Mallette, F. A. Cellular Senescence in Postmitotic Cells: Beyond Growth Arrest. *Trends Cell Biol* **28**, 595–607 (2018).

60. Khosla, S., Farr, J. N., Tchkonia, T. & Kirkland, J. L. The role of cellular senescence in ageing and endocrine disease. *Nature Reviews Endocrinology* Preprint at https://doi.org/10.1038/s41574-020-0335-y (2020).

61. Zhu, Y. *et al.* The achilles' heel of senescent cells: From transcriptome to senolytic drugs. *Aging Cell* (2015) doi:10.1111/acel.12344.

62. Chaib, S., Tchkonia, T. & Kirkland, J. L. Cellular senescence and senolytics: the path to the clinic. *Nature Medicine 2022* 1–13 (2022) doi:10.1038/s41591-022-01923-y.

63. Justice, J. N. *et al.* Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study. *EBioMedicine* (2019) doi:10.1016/j.ebiom.2018.12.052.

64. Hickson, L. T. J. *et al.* Senolytics decrease senescent cells in humans: Preliminary report from a clinical trial of Dasatinib plus Quercetin in individuals with diabetic kidney disease. *EBioMedicine* (2019) doi:10.1016/j.ebiom.2019.08.069.

APPENDICES

NEDERLANDSE SAMENVATTING

Deel 1. Systeembiologie in metabole ziekten

Hoofdstuk 1 is een introductie van dit proefschrift. Dit proefschrift beschrijft eerst de rol van een systeem biologische benadering om grote hoeveelheden data (hierna omics genoemd) te integreren met als doel de hiërarchie in de metabole paden die betrokken zijn in de ontwikkeling van metabole ziekten te ontdekken. Gebruikmakend van de opgedane kennis in het eerste deel van dit proefschrift, veranderden wij in deel 2 de samenstelling en functionaliteit van de darmmicrobiota door middel van interventie studies in mensen met als overkoepelend doel om nieuwe behandelmogelijkheden te vinden voor niet alcoholische leververvetting (NAFLD). Tot slot hebben wij geprobeerd om nieuwe drijvers van cellulaire senescence te identificeren in het laatste deel.

Hoofdstuk 2 is een overzicht van de kwaliteit van het toenmalige bewijs voor een causale rol van de darmmicrobiota in de ontwikkeling van obesitas en type 2 diabetes (T2D). Sinds de opkomst van betaalbare "next generation sequencing-technieken" heeft een groot aantal studies intrigerende verbanden aangetoond tussen de samenstelling van de darmmicrobiota, obesitas en T2D. Studies uitgevoerd in mensen waarin de samenstelling van de darmmicrobiota werd veranderd, bijvoorbeeld door antibioticabehandeling of een fecestransplantatie (FMT), wijzen op een causaal verband tussen de darmmicrobiota en de ontwikkeling van metabole ziekten. Echter is het bewijs voor een causale rol nog erg klein. Bovendien bieden deze studies geen mechanistisch inzicht in de wisselwerking tussen de darmmicrobiota en het humane metabolisme. Grote prospectieve studies zullen van cruciaal belang zijn om bewijs te vinden voor de veronderstelling dat de samenstelling van de darmmicrobiota een weerspiegeling is van de ziekte, of dat anderzijds de samenstelling van de darmmicrobiota vóór de ontwikkeling van de ziekte al veranderd was en dus een drijvende factor is. Om dit bewijs te leveren zijn er dus zowel grote prospectieve- als interventie-studies in mensen nodig, in combinatie met mechanistische studies in proefdieren en modelsystemen.

In **hoofdstuk 3** valt te lezen dat wij hebben geprobeerd om een darmmicrobiota-signatuur te ontrafelen voor obesitas, door middel van 'whole genome shotgun sequencing' van feces van mensen met een grote variatie in de body mass index (BMI). Het BMI van de personen varieerde van slank tot morbide obesitas. Uit deze analyses bleek dat een totaal van 52 fecale bacteriesoorten significant verschillend waren tussen mensen met en zonder obesitas. Ook vonden wij dat de samenstelling van de darmmicrobiota meer dan 50% van de variantie in klinische kenmerken van obesitas kon verklaren, zoals BMI, middelomtrek, High-Density Lipoprotein en triglyceriden. Daarnaast observeerden wij verschillen in metabole paden, potentieel aanwezig in de darmbacteriën bij personen met en zonder

obesitas en vonden wij bijzonder sterke associaties tussen aminozuurmetabolisme, diverse darmbacteriesoorten en obesitas. Metabole paden die betrokken zijn bij de biosynthese van verschillende aminozuren (waaronder histidine, lysine en tryptofaan) waren sterk verhoogd bij personen met obesitas, terwijl paden die betrokken zijn bij de afbraak van deze aminozuren, met name histidine, waren uitgeput. Dit suggereert dat de darmmicrobiota van personen met obesitas meer potentie heeft om verschillende aminozuren te produceren in vergelijking met personen zonder obesitas, maar een verminderde capaciteit heeft om specifieke aminozuren af te breken.

Onderzoek naar de rol van de darmmicrobiota in de ontwikkeling van metabole ziekten is uitdagend omdat het moeilijk is om biopten te verzamelen van de aangetaste organen zoals lever, darm en vetweefsel. Bariatrische chirurgie biedt de unieke mogelijkheid om biopten te verkrijgen uit verschillende vetweefseldepots, de dunne darm, lever en zelfs bloed uit de poortader.

In **hoofdstuk 4** beschrijven wij de opzet en de doelstellingen van de BARIA-studie. Met behulp van dit longitudinale bariatrische chirurgie cohort, wilden we een systeembiologische benadering uitvoeren om nieuwe pathogene paden in de pathogenese van obesitas, T2D en NAFLD te identificeren. Deze nieuwe paden zouden een target kunnen zijn voor nieuwe toekomstige geneesmiddelenontwikkeling. Om genoeg power te hebben om deze nieuwe paden te identificeren hebben wij als doel om in de BARIA-studie 1500 personen te includeren die een primaire laparoscopische bariatrische operatie ondergaan. Voor de operatie krijgen de deelnemers een twee uur durende maaltijdtest (MMT) om de insulineresistentie te beoordelen en de dynamische veranderingen in circulerende metabolieten te onderzoeken. Ook worden er bloed- en fecale samples afgenomen en worden er vragenlijsten ingevuld, waaronder psychologische vragenlijsten. Tijdens de operatie worden biopten genomen van drie vetdepots, jejunum, lever en in een deel van de patiënten wordt ook portaal bloed afgenomen.

Als een patiënt opnieuw geopereerd moet worden (revisieoperatie, cholecystectomie) kunnen wederom biopten worden afgenomen. De verzamelde feces samples, bloed en biopten van de verschillende weefsels kunnen worden gebruikt voor darmmicrobiota analyses, RNA-sequencing, en plasma metabolomics (nuchter alsook in de postprandiale fase). Door gebruik te maken van systeembiologie beschikken we over de mogelijkheid om alle omics-data te integreren om inzicht te krijgen in de hiërarchie van de mechanismen die ten grondslag liggen aan de ontwikkeling van metabole ziekten. In de volgende drie hoofdstukken beschrijven we zowel klinische als panomics data van de eerste 106 deelnemers va de BARIA-studie.

In **hoofdstuk 5** staat beschreven dat wij een nieuw concept hebben ontwikkeld voor de stratificatie van personen met obesitas. Dit nieuwe concept maakt geen gebruik van

klinische parameters zoals leeftijd, geslacht, aanwezigheid van bepaalde aandoeningen, maar stratificeert alleen op de verhouding in plasma metabolieten. Het plasma metaboloom geeft een uniek inzicht van de algehele fysiologische conditie van de mens en de interactie tussen genetica, leefstijl, medicijngebruik en de activiteit van het darmmicrobioom. In de context van obesitas en cardiometabole ziekten wordt metabolomics gebruikt als middel om de metabole gezondheid te evalueren, het effect van dieetinterventiestrategieën te meten en voorspellende biomarkers te identificeren die een specifieke aandoening karakteriseren. Op deze manier hoopten wij kleine nuances tussen personen te kunnen onderscheiden omdat deze vaak niet accuraat worden weergegeven door het gebruik van louter klinische parameters in complexe aandoeningen zoals obesitas en T2D. Het is namelijk zo dat aandoeningen zoals obesitas en daarbij horende comorbiditeiten zoals hypertensie, T2D, NAFLD en dyslipidemie als binaire variabelen worden weergegeven (aanwezig of afwezig), maar dit geeft de complexiteit van de aandoening en het individu niet goed weer. Zo kan het zijn dat de aandoening van een persoon goed, of juist niet goed is gereguleerd. En zo zijn er nog tal van andere factoren die niet goed worden weergeven als een aandoening enkel als binair wordt geclassificeerd. Deze nieuwe methode leverde vijf verschillende clusters in het BARIA-cohort op die we “metabotypes” noemden. Het classificeren van patiënten door middel van deze metabotypes stelde ons in staat om associaties en verbanden bloot te leggen tussen verschillende darmbacteriën, metabolieten en het transcriptoom van vet en leverweefsel. Onze bevindingen suggereren dat stratificatie van patiënten op basis van metabotyping, ons nieuwe moleculaire inzichten kan opleveren.

In **hoofdstuk 6** beschrijven we de resultaten van plasma metabolomics analyses op zowel nuchtere als postprandiale bloedsamples. We hebben de globale metabole reacties onderzocht op een maaltijdtest. We stratificeerden de proefpersoenen op basis van hun glucosetolerantie in normale glucosetolerantie (NGT), pre-diabetes en T2D groepen. Deze analyses toonden aan dat het aantal metabolieten dat significant verschilden tussen deze groepen het grootst was tussen de NGT- en T2D-groepen in de post prandiale fase in vergelijking met de nuchtere fase. Hierdoor kregen we nieuwe inzichten in het verstoorde metabolisme bij patiënten met insuline resistentie wat niet naar voren zou zijn gekomen als we alleen de nuchtere fase plasma metabolomics hadden geanalyseerd. Verder konden we de verschillen in plasma metabolieten terug herleiden naar andere omics sets, waaronder fecale metagenomics en transcriptomics van lever, vetweefsel en jejunum. We vonden sterke associaties tussen metabolieten en genen in de lever, vet en darm. Door het voorspellen van de glucose oppervlakte onder de curve (AUC) met omics data, waren we in staat om nieuwe drijvers en mogelijke biomarkers voor insuline resistentie te identificeren. Plasma metabolomics, zowel in de nuchtere als post prandiale fase voorspelden de glucose AUC het beste. Belangrijkste metabolieten in deze machine learning analyses waren fenylalanine en

1-carboxyethylfenylalanine, welke mogelijk gebruikt kunnen worden als biomarkers voor glycemische controle. De darmmicrobiota samenstelling was minder goed in staat om de glucose AUC te voorspellen.

In **hoofdstuk 7** wordt ingegaan op de wereldwijd meest voorkomende chronische leverziekte, NAFLD. Naar schatting heeft één op de vier personen in de algemene bevolking NAFLD. Deze prevalentie neemt toe tot meer dan 80% bij personen met obesitas. De drastische toenemende prevalentie van NAFLD en het gebrek aan effectieve en geregistreerde behandelingsmogelijkheden om deze ziekte aan te pakken, zorgt voor een grote druk op de gezondheidszorg en economie. Om geschikte, niet-invasieve diagnostische methoden en behandelingsopties te ontwikkelen, is het van cruciaal belang om de complexe pathofysiologie van NAFLD grondig te onderzoeken. In dit hoofdstuk maakten we gebruik van een systeembiologische benadering om zo de bijdrage van verschillende organen aan deze ziekte te onderzoeken. We analyseerden transcriptomics data van lever- en vetweefsel, fecale metagenoom en plasmametaboloom van 55 vrouwen met en zonder NAFLD. Significante verschillen in metabolieten, expressie van genen in de lever en vetweefsel en darmmicrobiota tussen vrouwen met en zonder NAFLD werden geobserveerd. Door een multivariaat model te ontwikkelen toonden we bovendien aan dat er een aanzienlijke interactie is tussen deze verschillende omics-sets. Deze studie biedt een uitgebreide panomics-analyse van personen met NAFLD in een vroeg stadium en laat daarnaast een nieuwe strategie zien om de pathofysiologie van NAFLD in mensen te bestuderen.

Deel 2: Van associatie naar causaal bewijs

Deel 2 laat zien hoe de opgedane kennis beschreven in deel 1 de mogelijkheid bracht om van associatie tot causaliteit te komen door middel van het uitvoeren van interventiestudies waardoor de samenstelling en functionaliteit van de darmmicrobiota veranderde in patiënten met NAFLD.

In **hoofdstuk 8** staat dat wij door vier verschillende experimenten uit te voeren aan konden tonen dat de darmbacteriën van mensen grote hoeveelheden ethanol kunnen produceren die klinisch relevant kunnen zijn voor de pathogenese en progressie van NAFLD. Wij gebruikten hiervoor onder andere bloed uit de poortader. Dit is het meest relevante bloedvat om microbieel geproduceerde metabolieten te bestuderen, omdat bloed uit het maagdarmkanaal rechtstreeks in deze ader wordt afgevoerd naar de lever. In de poortader werden hoge ethanolconcentraties aangetroffen, die significant correleerden met NAFLD-parameters. Bovendien zagen wij dat het ethanol in de perifere circulatie toenam tijdens een maaltijd. Hoge postprandiale plasma ethanol concentraties correleerden met name met een hoge relatieve aanwezigheid van lactic acid bacteriën in de feces. Door het remmen van alcoholdehydrogenase, werd het “first-pass effect” van de lever omzeild en steeg

de ethanol concentratie met een factor 15 in het perifere bloed bij patiënten met NAFLD. Dit effect werd volledig tenietgedaan na behandeling met breedspectrumantibiotica. Wij verkregen hiermee causaal bewijs dat de darmbacteriën van mensen grote hoeveelheden ethanol kunnen produceren. In hoeverre de aanhoudende endogene ethanolproductie causaal betrokken is bij de zeer complexe pathogenese van NAFLD, waarbij een combinatie van omgevingsfactoren, genetische varianten, obesitas en een verstoorde lipidenhomeostase een rol spelen, moet echter nog verder worden onderzocht.

In **hoofdstuk 9** beschrijven wij de resultaten van een dubbelblind gerandomiseerd gecontroleerd onderzoek waarin personen met NAFLD, werden gerandomiseerd naar twee studiearmen; slanke veganistische donor (allogeen n=10) of eigen (autoloog n=11) FMT. Elke deelnemer kreeg na randomisatie driemaal een FMT met tussenpozen van acht weken. Bij elke proefpersoon werd twee keer een leverbiopsie verricht. Een keer voor de start en een keer 24 weken na de start van de behandeling om de histopathologie (NASH-CRN) classificatie en veranderingen in levergenexpressie op basis van RNA-sequencing te bepalen. Hoewel de studie niet genoeg power had, werd er een trend in verbetering van de necroinflammatie score (lobulaire inflammatie en hepatocyte ballooning) gevonden, ook wel bekend als ziekteactiviteit geobserveerd de allogene FMT-groep. Deze FMT bestond uit feces afkomstig van veganistische donoren welke een plantaardig dieet met weinig dierlijke eiwitten hadden. Bovendien had de allogene FMT een effect op de samenstelling van de darmmicrobiota, welke geassocieerd was met zowel gunstige veranderingen in plasmametabolieten als de expressie van levergenen die betrokken zijn bij ontsteking en lipidenmetabolisme.

Deel 3: cellulaire senescence, een oude maar nieuwe speler in metabole ziekten

Dit deel gaat over een zowel een oude als nieuwe speler bij metabole ziekten: cellulaire senescence. Cellulaire senescence is een phenomeen dat wordt gekenmerkt door een onomkeerbare stop van de celcyclus welke belangrijke fysiologische functies voor de cel heeft. Cellulaire senescence is echter ook een kenmerk van veroudering en wordt in verband gebracht met verschillende metabole ziekten.

Hoofdstuk 10 is een overzicht waarin de kwaliteit van het bewijs voor een causale rol van cellulaire senescence in het ontstaan van NAFLD wordt samengevat. We bespreken hier onder andere studies in proefdieren en mensen waaruit zou blijken dat cellulaire senescence een causale rol speelt in de ontwikkeling en progressie van NAFLD. Wij gaan ook dieper in op de definitie van cellulaire senescence en bespreken de belangrijkste cellulaire en moleculaire kenmerken van senescence in de lever. Tenslotte bespreken we de effectiviteit en veiligheid van nieuwe senolytische geneesmiddelen die mogelijk in de toekomst gebruikt kunnen worden om NAFLD te behandelen of te voorkomen.

In **hoofdstuk 11** beschrijven wij het verband tussen insuline en senescence in de lever. Dit verband konden wij leggen door gebruik te maken van klinische gegevens, bloed uit de poortader, immunohistochemische en transcriptomics data van patiënten met obesitas. Wij vonden sterke correlaties tussen markers van cellulaire senescence in de lever met insuline. Daarnaast observeerden wij een zeer sterke correlatie van insuline in de poortader en senescence in de lever, onafhankelijk van NAFLD. Dit suggereert dat senescence in de lever, gedreven door hoge concentraties van insuline, optreedt voordat NAFLD is ontwikkeld. Met behulp van een valideringscohort van 180 personen konden wij deze bevindingen valideren.

In **hoofdstuk 12** hebben wij dezelfde aanpak gebruikt als in het vorige hoofdstuk. Hier gebruikten we klinische gegevens, immunohistochemische en transcriptomics data van patiënten zonder T2D en toonden wij aan dat perifere insulineresistentie sterk gecorreleerd is met markers van cellulaire senescence in visceraal vetweefsel. Deze resultaten bevestigen eerdere studies die erop wijzen dat cellulaire senescence in vetweefsel een belangrijke rol kan spelen in een vroeg stadium van insulineresistentie.

LIST OF PUBLICATIONS

Included in this thesis

1. **Meijnikman, A. S.**, et al., Evaluating Causality of Gut Microbiota in Obesity and Diabetes in Humans. *Endocrine Reviews* **39**, 133–153 (2018).
2. **Meijnikman, A. S.** *et al.* Distinct differences in gut microbial composition and functional potential from lean to morbidly obese subjects. *Journal of Internal Medicine* (2020) doi:10.1111/joim.13137
3. Van Olden C.C. Van de Laar A, **Meijnkman A.S.,** et al.. A Systems Biology approach to understand gut microbiota and host metabolism in morbid obesity: design of the BARIA Longitudinal Cohort Study. *Journal of Internal Medicine* joim.13157 (2020) doi:10.1111/joim.13157.
4. Li P, Ji B, Lappa D, **Meijnikman AS**, Olsson LM, Aydin Ö, et al. Systems analysis of metabolic responses to a mixed meal test in an obese cohort reveals links between tissue metabolism and the gut microbiota. bioRxiv 2022
5. Lappa D, **Meijnikman AS**, et al. Self-organized metabotyping of obese individuals identifies clusters responding differently to bariatric surgery. *Plos one. 2022*. epub ahead of print
6. **Meijnikman, A.S.**, et al. A systems biology approach to study non-alcoholic fatty liver (NAFL) in women with obesity, *iScience,* 2022
7. **Meijnikman**, A.S., et al. Microbiome derived ethanol and non-alcoholic fatty liver disease. *Nature medicine 2022.*
8. Witjes J, Smits L, Pekmez C, Prodan A, **Meijnikman A.S.,** et al.*,* Donor Fecal Microbiota Transplantation Alters Gut Microbiota and Metabolites in Obese Individuals With Steatohepatitis. *Hepatology Communications*. 1601 (2020) doi:10.1002/hep4.1601.
9. **Meijnikman, A. S**. *et al.* Evaluating causality of cellular senescence in Non-Alcoholic Fatty Liver Disease. *JHEP Reports* **0**, 100301 (2021).
10. **Meijnikman, A.S.**, *et* al. Hyperinsulinemia is highly associated with markers of hepatocytic senescence in two independent cohorts. *Diabetes. 2022*
11. Van Olden CC, **Meijnikman AS**., et al. Adipose tissue senescence is associated with increased postprandial plasma insulin concentrations in bariatric surgery patients. *Submitted after revisions.*

Other

12. Mikkelsen RB, Arora T, Trošt K, Dmytriyeva O, Kjærsgaard Jensen S, **Meijnikman AS,** et al. Type 2 diabetes is associated with increased circulating levels of 3-hydroxydecanoate activating GPR84 and neutrophil migration. *iScience 2022*;

13. Sindhunata DP, **Meijnikman AS**, Gerdes VEA, Nieuwdorp M. Dietary fructose as a metabolic risk factor. *Am J Physiol Cell Physiol 2022*.
14. van Rijswijk A, van Olst N, **Meijnikman AS**, et al. The effects of laparoscopic Roux-en-Y gastric bypass and one-anastomosis gastric bypass on glycemic control and remission of type 2 diabetes mellitus: study protocol for a multi-center randomized controlled trial (the DIABAR-trial). *Trials 2022*.
15. Guman MSS, Hoozemans JB, Haal S, De Jonge PA, Aydin O, **Meijnikman AS**, et al. Adipose Tissue, Bile Acids, and Gut Microbiome Species Associated With Gallstones After Bariatric Surgery. *J Lipid Res 2022*.
16. **Meijnikman, A. S.,** et al., Increased expression of key SARS-CoV-2 entry points in multiple tissues in individuals with NAFLD. *Journal of Hepatology* (2021)
17. Scheithauer, T. P. M. Davids M, Winkelmeijer M, Aydin O, De Brauw M, Van De Laar A, **Meijnikman, A.S.** *et al.* Compensatory intestinal antibody response against pro-inflammatory microbiota after bariatric surgery. *Gut Microbes* **14**, (2022).
18. Bakker G.J., **Meijnikman A.S**., et al. Fecal microbiota transplantation does not alter bacterial translocation and visceral adipose tissue inflammation in individuals with obesity. Obes Sci Pract. 2021 Aug 13;8(1):56-65. doi: 10.1002/osp4.545.
19. Husted A.S, Ekberg J.H, Tripp E, Nissen T.A.D, **Meijnikman A.S.** et al.Autocrine negative feedback regulation of lipolysis through sensing of NEFAs by FFAR4/GPR120 in WAT. *Mol Metab* **42**, (2020).
20. **Meijnikman, A. S.,** de Block, C. E. M., Verrijken, A., Mertens, I. & van Gaal, L. F. Predicting type 2 diabetes mellitus: A comparison between the FINDRISC score and the metabolic syndrome. *Diabetology and Metabolic Syndrome* (2018) doi:10.1186/s13098-018-0310-0.
21. **Meijnikman, A. S.** *et al.* Not performing an OGTT results in significant underdiagnosis of (pre)diabetes in a high risk adult Caucasian population. in *International Journal of Obesity* (2017). doi:10.1038/ijo.2017.165.
22. **Meijnikman, A. S**. *et al.* Screening for type 2 diabetes mellitus in overweight and obese subjects made easy by the FINDRISC score. *Journal of Diabetes and its Complications* **30**, 1043–1049 (2016).

DANKWOORD

In dit dankwoord wil ik iedereen bedanken die op welke manier dan ook heeft bijgedragen aan de totstandkoming van dit proefschrift.

Allereerst gaat mijn dank uit naar alle deelnemers van de studies beschreven in dit proefschrift. Het is bewonderenswaardig dat zoveel mensen vrijwillig hebben meegedaan aan onderzoek variërend van poeptransplantaties tot aan puncties in de poortader. Dank voor het vertrouwen!

Beste Max, ik beschouw het als een eer dat ik bij jou mocht promoveren. Mijn dank is groot voor de vrijheid die je me hebt gegeven om mijn eigen pad te bewandelen en om er te zijn als er bijgestuurd moest worden. Je hebt een zeldzaam talent om overal mogelijkheden en oplossingen te zien waar anderen stagneren. Ik heb dan ook veel bewondering voor de manier waarop jij het hoogleraarschap, promovendi, hoofd van drie afdelingen en gezinsleven combineert. Je ongekende efficiëntie en doelmatigheid is een inspiratiebron, vooral omdat je in de drukte rekening blijft houden met de mensen waarmee je samenwerkt. Je probeert het beste uit mensen te halen en dat gaat verder dan enkel de wetenschap en carrière. Veel dank voor alle mooie momenten, dat er nog maar vele mogen volgen.

Beste Bert, wat ben ik dankbaar dat ik jou heb mogen ontmoeten. Ondanks dat je het zelf vaak ontkent is je kennis ongeëvenaard. Ontelbaar vaak dacht ik dat ik iets slims en innovatiefs had bedacht maar dan kwam jij toch weer met een nieuwe invalshoek, die dan vaak weer bleek te kloppen. Wetenschap op zijn mooist. Ooit, zal ik de metabole processen in de lever volledig begrijpen en hoef jij me niet meer te corrigeren. Je bent dan ook een geweldige supervisor en mentor. Je brengt zoveel rust en laat mij zweven wanneer het kan (ethanol na 4-methylpyrazol experiment) en zet me daarna weer met beide benen op de grond. Naast de supervisie op het wetenschappelijke gebied ben je er ook op andere vlakken voor mij geweest (en nog steeds). Door jou kon ik verder groeien. Mijn dankbaarheid en gevoel is moeilijk uit te drukken in woorden maar ik zie het als een zegen dat iemand zoals jij er altijd voor me wil zijn.

Beste Victor, dit hele proefschrift zou er niet geweest zijn als jij het niet in mij zag zitten om in jouw ziekenhuis de BARIA-studie op te zetten. De vrijheid die je me gaf om “lekker te klussen” greep ik van harte aan, wetende dat als ik er echt niet uitkwam ik terug kon vallen op jou. Soms trapte je wel eens op de rem, bijvoorbeeld toen we op een gegeven moment een hele afdeling voor het onderzoek hadden geclaimd. Zelfs jij vond dit een beetje te veel

van het goede. Ik heb bewondering voor je hoe je om bent gegaan met het faillissement van het Slotervaart Ziekenhuis en in alle tumult de rust bewaarde. Wetenschappelijk heb ik er bewondering voor dat je echt helemaal niks om impact factors geeft en enkel zuivere wetenschap wil bedrijven.

Beste Hilde, met stip op één heb jij de meeste whatappjes en mailtjes van mij ontvangen. Ooit aangeschoven bij een overleg over zonering in de lever en daarna was je niet meer weg te denken uit mijn promotieteam. Dank voor al je input, kritische blik en dat je bereid was om je fantastische schrijfkunst los te laten op mijn stukken. Ik waardeer je tomeloze inzet voor het laboratorium en als moeder van de groep waar iedereen zijn hart kan luchten, inclusief ikzelf. Bovenal ben je gewoon echt een lief mens. Aan je wetenschappelijke kennis en kunde twijfelt niemand. Toch blijven die bacteriofagen een beetje vaag in metabole aandoeningen. Ik wacht ongeduldig totdat je deze gaat linken aan levermetabolisme.

Dear colleagues from the University of Gothenburg and Copenhagen. Fredrik, it was a pleasure and honor working with you on the BARIA cohort in the NNF consortium. Thue, your scientific mind still amazes me, your twitter capacities are mind blowing. Jens, I still don't know how you can perform cutting edge science and be CEO of the BioInnovation institute, and this with drinking coffee only until 12:00. Valentina, Louise, Annika, Siv and Kimberley, thank you for the fruitful collaboration. Your microbiome expertise is unparalleled.

Beste Christophe en Sven, in 2016 verliet ik de stad waar ik mijn hart aan heb verpand, Antwerpen. Ik ben dankbaar dat we na al die jaren nog samenwerken. Veel dank daarvoor.

Geachte leden van de promotiecommissie, prof. dr. Kuipers, prof. dr. Rensen, prof. dr. Beuers, prof. dr. De Jonge, prof. dr. Bergman, prof. dr. Houtkooper. Dank voor het zitting nemen in mijn promotiecommissie en de kritische beoordeling van dit proefschrift.

Het BARIA-team. Casper, Hanneke, Ömrüm, Anne-Sophie, Jacqueline, Kadriye en Nienke. Wat hebben wij een hoop bloed afgenomen en poep verstouwd. Af en toe leek het net goede tijden slechte tijden en gelukkig bestaat er zoiets als relatietherapie, maar dan voor collega's. De goede tijden overheersten gelukkig en waren ook echt goed. We konden ook genieten van onze samenwerking. Casper, je hebt delegeren tot kunst verheven maar hard werken ga jij ook niet uit de weg. Ik heb veel respect voor je hoe jij je moeilijk analyses en kleuringen eigen maakt zodat je dan toch ook weer onafhankelijk kan zijn van zo'n beetje iedereen. Ik ben er dan ook van overtuigd dat je buiten het ziekenhuis een prachtige succesvolle

carrière tegemoet gaat. Hanneke, waar zou de BARIA-studie zijn gebleven zonder jou? Van elke patiënt ken jij de voorgeschiedenis en de grote levensgebeurtenissen. De persoonlijke benadering en oprechte interesse in de mens achter de patiënt is de reden dat we zo weinig loss to follow-up hebben. De BARIA zou jouw laatste relaxte baan zijn om je carrière af te sluiten. Het liep iets anders maar wat heb ik van jouw tomeloze inzet en passie genoten. Je bent een geweldig mens met een groot hart en ik ben blij dat ik je heb leren kennen! Nienke, onze chirurg in spe, wat kan jij lekker beuken, hopelijk mag je binnenkort officieel aan je snijkunsten werken. Jacqueline, karaoke met jou was fantastisch en wat ben je een sociaal en lief mens. Kadriye, welkome versterking voor het BARIA-team, je pragmatisme wordt op prijs gesteld, succes met de echo's. Anne-Sophie, wat hebben we heerlijke gelachen en ook heerlijk ruzie gemaakt. Ik ben blij voor je dat je zo je draai hebt gevonden in Brabant. Hoe dan ook komen onze gezamenlijke stukken ooit nog een keer af. Daniko en Max, jullie zijn waardige opvolgers.

Collega's uit het Slotervaart en Spaarne Gasthuis. Thomas, de vader van de groep. Vaak heb ik dubbel gelegen om je grappen en originele blik op de dagelijkse gang van zaken. Je bent degene die zijn zaakjes altijd keurig op orde heeft en daar was je promotie een voorbeeld van. Een boekje vol met retrospectief onderzoek van soms twijfelachtig allooi (citaat uit je eigen dankwoord) maar dan wel weer een ZonMw beurs binnenhalen waar uiterst klinisch relevante onderzoeksvragen mee beantwoord zijn. Wel beetje jammer dat microbioom onderzoek na bariatrische chirurgie totaal verwoest gaat worden dankzij de inname van UDCA. Godzijdank is Nederland zo bureaucratisch dat het nog jaren duurt alvorens UDCA opgenomen wordt in de richtlijn. Floris, jij was altijd wel in voor een praatje, een koffietje of een potje tafeltennis. Je bent de rust hemzelf en eigenlijk een veel geschiktere vader voor de groep. Dat nam je ook wel letterlijk en plots was je vader van twee kinderen. Het was nog gepland ook. Ik heb veel respect voor je hoe je alle ballen hooghoudt en gestaag verder werkt als voorzitter van onze stichting. Je integriteit en je zelfbeheersing om compromitterende foto's van bepaalde personen niet te delen waardeer ik (en Thomas) enorm. Laten we die wintersport nog een keer overdoen! Paula, toen ik jou leerde kennen was je net met je "green happiness" sapjes bezig. Eigenlijk best goed voor je darmflora. Je sapjes mochten dan wel groen en happy zijn, de discussies over een der welk onderwerp waren altijd vol passie. Op het handbalveld was je meer van de kleur geel en rood, natuurlijk altijd onterecht. Ik vind het bijzonder knap dat al die kasten vol vragenlijsten toch gewoon zijn geanalyseerd en je vorig jaar bent gepromoveerd. Daarnaast heb ik veel waardering voor je inzet en collegialiteit.

Karin, mijn allereerste kamergenoot. Een groter verschil in persoonlijkheden bestaat er eigenlijk niet en dat is zo'n beetje het enige waar wij het volledig met elkaar over eens zijn. Ik heb dan ook veel van je geleerd. Maimoena, mede cola drinker, dank dat ik af en toe op je polikamer op de onderzoeksbank kon neerploffen op vrijdagmiddag. Ik heb bewondering voor je collegialiteit en verantwoordelijkheidsgevoel. Sylke, de meest gestructureerde en doelgerichte collega van de groep met als levensmotto: "work hard, play hard". Als enige collega had je een kamer alleen en dat kwam goed van pas want dan kon je tenminste beetje doorwerken. Op borrels, wintersport, skiweekend of congressen was dat wel anders en was je altijd de gangmaker met als hoogtepunt de UEG in Barcelona. Het lot heeft bepaald dat je niet in het ziekenhuis gaat werken maar ik weet zeker dat je ook buiten het ziekenhuis een mooie toekomst tegemoet gaat samen met Koen en Gosse.

Mijn dank gaat ook uit naar de internisten van het Slotervaart. Marcel, Jan-Willem, Hans-Martin, Saskia, Fanny, Roeland en Eelco. Een groep zo uiteen zien vallen is niet niks maar gelukkig (en vanzelfsprekend) zijn jullie allemaal goed terecht gekomen. Dees, dank voor je steun en geloof in mijn projecten. Nico, dank voor alle papers die ook maar iets met ethanol te maken hadden. Verder gaat mijn dank uit naar alle doktersassistenten van de poli bariatrie Jantien, Crista, Jorinde en Simone.

Zonder chirurgen geen bariatrie dus ook geen proefschrift. Maurtis, Sjoerd, Arnold, Yair en Rutger, veel dank voor het meehelpen om een half mens in de vriezer te stoppen. Maurits, levensgenieter, ik benijd je levenswijze. Wanneer gaan we nu eens samen kiten? Sjoerd, toch mooi hoe je van "hoezo zoveel biopten" naar "wil je nog wat meer?" bent gegaan. De reden hierachter is dat je je zorgen maakte over de gezondheid van jouw patiënten en dat kan ik alleen maar waarderen. Patiënten zijn in meer dan goede handen bij BARIA Nederland met jou als hoofd. Arnold, wat kun je heerlijk kritisch zijn en de vinger op de zere plek leggen, maar wat verwacht je anders van iemand die is opgeleid in Leuven, de twee na beste universiteit van België. Yair, de onvoorwaardelijke steun die jij voor de groep hebt is noemenswaardig en is een voorbeeld voor mij. Rutger, nieuwe aanwinst, heerlijk pragmatisch en je gaat nog promoveren ook. Als dat geen bier drinken wordt.

G1 boyz: Ilias, Koen, Torsten, wat een wilde rit was dit mannen. De avondjes weg met jullie waren nooit saai en liepen eigenlijk altijd uit de hand. Details zal ik de lezers besparen, maar het was goed. Ilias, ik snap nog steeds niet helemaal waarom we elkaar blindelings vertrouwden om infusen bij elkaar te prikken na een avondje stappen. Dank voor de vriendschap mannen.

Collega's van het vasculaire lab op G1. Wat is het toch fijn om de dag te starten met een bakje koffie in de koffiekamer rond 08:00 uur. Ondanks dat de werkdag pas echt begon rond 09:00 uur had je toch altijd het idee dat je al lekker bezig was geweest. Maar leg me nu toch nog een keer uit waarom je niet met een witte jas in de koffiekamer mag komen, en dat je bij overtreding koekjes moet trakteren. Ik denk dat ik er persoonlijk verantwoordelijk voor ben dat iedereen 1k g per jaar gewicht is aangekomen. Dit alles onder streng toeziend oog van Wil. Zo streng als op de koekjes was je ook in het netjes pipeteren. Nogmaals en alvast mijn excuses voor het laten aanstaan van de computers. Stefan, dank voor de hulp met de kleuringen. Is het nu Texas Red, Fitch of toch Cy-3, irritant hè die TAF kleuringen, maar wel leuk. Hans, ongelooflijke held, veel dank voor de magic mix en de eerste bepalingen van ethanol. Maaike, je bent een vrouw om van te houden en je vriezerindeling is beter dan die van mij maar toch houden we het lekker zo. Veel dank voor al het werk wat je voor de BARIA doet ondanks alle andere werkzaamheden die je hebt. Han, onze HPLC-, tennis-, en niet te vergeten, op de skipiste "carvende" held. Dank voor al je werk. Miranda, Jorge en Alinda, het lab is niet compleet zonder jullie. Geesje, eindelijk hebben we een metaboliet volledig uitgewerkt, en kijk eens naar het resultaat. Misschien hadden we wat eerder naar je moeten luisteren. Dank voor de fijne introductie die je me hebt gegeven op het lab. Aldo, stille kracht met een bak aan kennis en ervaring, je innovatieve manier van werken vind ik mooi. Jef, je onderzoek is minstens net zo goed als je bierbrouw kunsten. Dank voor de vele koffie momentjes. Nu ik erover nadenk hebben we te weinig bier gedronken, hint. Elena, dank voor al je support. Agnes, Tanja en Silvia, onze afdeling zou gewoon niet kunnen draaien zonder jullie, dank voor al jullie support. Ook veel dank aan Jorn en Xanthe van het Microbiota Center Amsterdam voor jullie tomeloze inzet en de strakke regie op de klinische data.

Alle onderzoekers van het lab: Kim, alleskunner en doener, dank voor alle gezelligheid tijdens de koffie momentjes. Ik weet zeker dat je het gaat maken in de States. Moritz, wat ben je een fijne collega die bovenal gewoon hard werkt. Ik heb bewondering voor je inzet en je collegialiteit. Anne-Linde, als dokter muizenproeven doen is geen sinecure maar toch doe je het met ogenschijnlijk weinig moeite. De meest indrukwekkende congressen heb ik samen met jou meegemaakt en wat hebben we een plezier gehad. Dank voor de fijne tijd. Coco, de definitie van hard werken. In twee jaar tijd heb je een cohort van 500 patiënten opgebouwd en rond je succesvol twee interventie studies af. Ongelooflijk knap en wat heb ik een respect voor je. Nog zo'n harde weker, Madelief, grootgebruiker van de MRI in het weekend. Naast het harde werken kun je ook heel hard feesten en dat is precies waarom je zo'n leuk mens bent. Barbara, je bent goed in pipeteren, werken met grote datasets en interventiestudies, de combinatie om het te maken in de onderzoekswereld. Veera, toch knap dat je een

Nederlandse PIF geschreven hebt in het begin van je promotietraject zonder een woord Nederlands te spreken. Om dan ook nog een studie succesvol af te ronden vind ik erg knap. En dan de nieuwelingen: Merel, Melany, Mia, Katie, Q en Bas, er wordt ogenschijnlijk heel erg hard gewerkt en zelfs mijn tv-scherm en pocket super Nintendo zijn weg gedaan, ga zo door. Charlotte T, wat fijn dat we weer samenwerken maar dan in het AMC.

De oude Max-minions, Guido, Annefleur, Nicolien en Julia een voor een zijn jullie bijzondere mensen en ik dank jullie voor de fijne samenwerking en warme ontvangst in de groep. Guido, ik genoot van de momenten dat we probeerden samen patiënten in de Ralstonia studie te includeren, wat een opgave was dat zeg. Blijf het bijzonder vinden dat jij zo positief bleef en ook bleef geloven in de studie. Anne-Fleur ik heb bewondering voor hoe jij met alle tegenslagen in korte tijd bent omgegaan en nog altijd positief en vrolijk bent. Nicolien, dank voor alle tips variërend van wetenschap tot het vaderschap, ik heb er veel van opgestoken. Andere onderzoekers van het oude F4/M0/1 en VU, Floris, Frits, Eva, Kristien, Nick, Shirin, Tycho, Anne-Marieke, Valerie, Hanke, Arjen, Silvia, Rens en Lotte dank voor alle leuke momenten. In het bijzonder nog Renate voor alle leuke uitstapjes. Was toch een ander tijdperk he.

Zonder data scientists en machine learning experts geen boekje. Mark, wat was het mooi om samen helemaal op te gaan in het ethanol stuk. Zonder jouw advies om voor clindamycine te gaan weet ik niet of de studie zo succesvol zou zijn geweest. Je inzet daarnaast is ongeëvenaard en daar ben ik je enorm dankbaar voor. Andrei, veel dank voor de analyses en je geduld. Vaker wel dan niet moest de onderzoeksvraag toch net weer even anders. Ulrika, ongekend talent, wat jammer dat je net nu het leuk begint te worden weer teruggaat naar Scandinavië. De man/vrouw paper komt af, ik beloof het. Patrick, als faagspecialist breng je ongekend veel expertise mee en dat is hard nodig om de dark-matter puzzel een beetje lichter van stof te maken. Manon en Eduard, ik heb er geen actieve herinnering aan dat ik jullie heb lastiggevallen om informatie te krijgen over analyses. Ik heb wel een actieve herinnering dat het altijd gezellig was als ik jullie even kwam storen.
Toen het Slotervaart failliet ging konden we ons onderzoek voortzetten mede omdat we zo hartelijke zijn ontvangen op de clinical trial unit (CTU) van de afdeling vasculaire geneeskunde van het AMC. Veel dank gaat dan ook uit naar Diona, Hans, Liesbeth, Linda, Marianne, Nanet, Petra, Sandra en Tanja. In het bijzonder wil ik Daniela bedanken. Je bent nagenoeg altijd positief en geeft op een hele fijne manier leiding aan het CTU. Zelfs toen je me met 130 km/uur over de vluchtstrook naar het ziekenhuis in Leiden bracht omdat Charlotte daar opgenomen was bleef je positief. Ik hoop dat we nog lang kunnen samenwerken. We gaan iets ontzettends moois maken van alle epigenetica projecten! Ook wil ik de Spaarne academie, onder bezielende leiding van Greetje bedanken voor de hartelijke ontvangst.

Beste bazen van de vasculaire, Harry, Erik, Bert-Jan, Danny, Kees, Nordin en Onno, wat legden jullie de lat altijd hoog, zowel tijdens de journal club meetings als tijdens het onderzoeksproces zelf. Vanzelfsprekend sijpelde dit door, ook naar mijn onderzoek. Veel dank daarvoor. Nordin, de Amsterdamse branie had je als geboren Limburger altijd al. Geen wonder dat je helemaal op je plek zit in het AMC. Ik kijk uit naar onze verdere samenwerking.

Daniël, veel dank voor je scherpe en kritische blik en je bereidheid om mee te denken, zelfs als je daar eigenlijk helemaal geen tijd voor hebt. Je vergelijking met Narcos zal me altijd bijblijven.

Joanne, die ballooning van die hepatocyten blijft lastig. Dank dat je me zo kritisch hebt leren kijken naar leverweefsel. Gelukkig hebben we nog zo'n kleine duizend biopten te gaan dus als het aan mij ligt, werken we nog lang samen.

Collega's uit het Groene Hart Ziekenhuis. Ik kan me geen fijnere plek bedenken waar ik mijn vooropleiding had willen doen. Na een lang promotietraject was het spannend om weer te beginnen in de kliniek maar het GHZ voelde als thuiskomen. Annewieke en Ted wat zijn jullie fijne opleiders die echt staan voor "hun" assistenten. Ted, je bent er altijd voor mij geweest wanneer ik je nodig had en dat waardeer ik ontzettend. De fijne en veilige sfeer wordt mede gecreëerd door alle andere internisten, longartsen, cardiologen en MDL-artsen. Het warme welkom dat ik voelde was ook te danken aan de arts-assistenten. Marvyn, Suzanne, Nicky, Roos, Friso, Wouter, Cynthia, Jarieke, Michiel, Robert, Marjella, Lars, Fleur, Sharisa, Charlotte, Maria, Marlot, Jesse, Nina, Marloes, Kevin, Renske, Laurien en alle andere assistenten. Bedankt voor de collegialiteit en goede sfeer.

Maarten, dank voor het vertrouwen om bij jullie te mogen starten met de opleiding en veel dank dat je altijd meedenkt met onderzoek en meehelpt om patiënten te werven.

Ruud, je flexibiliteit en omgang met patiënten is voor mij een groot voorbeeld. Veel dank dat ik vaak tussendoor even naar het Alrijne kon komen voor een oogcheck.

Lieve paranimfen, lieve Ömrüm, duizendpoot. Wat ben ik blij dat jij bij mijn verdediging naast mij staat. Als er iemand mij uit de brand kan helpen dan ben jij het wel met je fabelachtige manier van presenteren en overtuigingskracht. Je verzet ontzettend veel werk en bent erg succesvol in wat je doet en blijft daarbij bescheiden. Het mag best wat Hollandser zo af en toe. Veel dank dat we vrienden zijn geworden. My dearest Torsten, what a crazy ride. We had so much fun and I'm grateful to have shared this journey with you. You were there

for me when I needed you. Let's drive to Heidelberg together, one more time. Thank you for your support throughout the years and for being a friend.

Kevin, Rogier, Onno, ik beschouw het als een groot voorrecht om zo'n hechte vriendengroep te hebben. Vanzelfsprekend zien we elkaar wat minder dan voorheen maar als we elkaar zien is het gevoel nagenoeg hetzelfde als tijdens onze studententijd. Dat het nog lang zo moge blijven.

Lieve Rob en Gerda, ik prijs mij gelukkig met jullie als schoonouders. Veel dank voor alle etentjes en uitjes. Vanaf het begin was het altijd een feestje om bij jullie te komen eten. Het op de juiste waarde weten te schatten van wijn in combinatie met goed eten hebben jullie mij dan ook bijgebracht. Bovenal ben ik dankbaar dat jullie er altijd zijn wanneer we jullie nodig hebben.

Lieve Joppe, mijn grote broer. Je bent onbaatzuchtig en je probeert altijd het goede te doen voor je medemens. Ik ken weinig mensen die als doel in hun werk hebben om zichzelf overbodig te maken in plaats van onmisbaar. Hierdoor ben je dan juist van onschatbare waarde. Samen met Marlies en je dochters bouw je verder aan de toekomst op een plek waar al zo'n lange geschiedenis ligt van onze familie. Ik hou van je.

Lieve mama, ik heb altijd jouw onvoorwaardelijke liefde en steun ervaren en mijn kinderen ervaren die nu ook. Je bent altijd positief en probeert altijd iets moois van het leven te maken. Vanaf mijn derde jaar gingen we elk half jaar met de trein naar het AMC voor mijn diabetescontrole. Je maakte er een leuk uitje van en leerde mij vanaf jongs af aan te kijken naar wat wel kan in plaats van te denken in belemmeringen. Deze positieve en creatieve insteek helpt mij nog elke dag. Lieve papa, ik mis je nog elke dag. Kon je me nu maar zien. Wat had ik graag een colaatje met je willen drinken op het eindresultaat en een knuffel van je gehad. Ik hoop dat je trots op me bent. Weet dat ik van je hou.

Lieve Charlotte, voor jou heb ik nog de meest bewondering. Ik zou niet weten welke andere vrouw ter wereld het uit zou houden met een man zoals ik waar "ik kom over 5 minuten naar bed" eigenlijk betekent, tot over 3,5 uur. Ik zal in ieder geval bij deze beloven dat ik nooit meer zal zeggen dat het hierna rustiger gaat worden. Wat ik ook beloof is dat ik er altijd voor jou en de kinderen zal zijn. Het is moeilijk in woorden uit te leggen hoeveel ik van je hou. Je bent de moeder van onze kinderen, het allermooiste geschenk.
Lieve Hugo en Louise, jullie vervullen mijn dag met vrolijkheid, liefde en plezier (en snot). Wat is het genieten om jullie op te zien groeien, elke dag weer.

PHD PORTFOLIO

Name PhD student:	A.S. Meijnikman
PhD period:	June 2016 – March 2021
Names of PhD supervisor(s) & co-supervisor(s):	Prof. dr. M. Nieuwdorp, Prof. dr. A.K. Groen, dr. V.E.A. Gerdes, dr. H.J. Herrema

1. PhD training		
	Year	**ECTS**
General courses		
BROK (basiscursus regelgeving klinisch onderzoek)	2016	1.0
Seminars, workshops and master classes		
- Weekly journal club, department of vascular medicine	2016-2021	5.0
- Weekly clinical education, department of vascular medicine	2016-2021	5.0
- Monthly research meeting, Spaarne Gasthuis	2016-2021	1.0
- Two weekly Diabetes AMC/VUMC meeting	2016-2021	2.0
- Monthly Microbiota journal club	2016-2021	2.0
		1.0
Presentations (oral)		
- EASL ILC 2022: microbiome produced ethanol: an underestimated burden on the liver.	2022	0.5
- Digestive Disease Days: microbiome derived ethanol and non-alcholic fatty liver disease	2022	0.5
- AG&M: Treating NASH	2022	0.5
- Student en leefstijl symposium: Non-Alcoholic Fatty Liver disease	2022	0.5
- NNF challenge symposium 2021: Eat before you drink	2021	0.5
- Keystone symposia: hyperinsulinemia, a key player in hepatic and adipose tissue cellular senescence	2020	0.5
- NNF challenge symposium: looking beyond the fecal bias	2019	0.5
- UEG week: Geographic and ethnicity effects on microbiome	2019	0.5
- DSMBS Does Mini Gastric Bypass beat Roux-en-Y Gastric Bypass in glycemic control?	2017	0.5
(Inter)national conferences		
- EASL ILC, London, England	2022	1.0
- Digestive Disease Days	2022	0.5
- Keystone symposia: obesity and NAFLD mechanisms and therapeutics, Banff, Canada	2020	1.0
- EASL ILC, Vienna, Austria	2019	1.0
- UEG week, Barcelona, Spain	2019	1.0
- EMBO | EMBL symposium: The human microbiome, Heidelberg, Germany	2018	1.0
- Amsterdam symposium on vascular & metabolic disease, Amsterdam, The Netherlands	2018	0.5
- NNF challenge symposium, Copenhagen, Denmark	2019, 2021	1.0
- World obesity week, Washington DC, USA	2017	1.0
- DSMBS, Netherlands	2017	0.5

2. Teaching		
	Year	**ECTS**
Tutoring, Mentoring		
Keuze onderwijs Cardiovasculaire Research and Care, Bachelor year 2 medicine	2018, 2019	1.0
Supervising		
- Master student Eline Bakker	2017	2.0
- Master student Natasha Bosma	2018	2.0
Other		
Consultation at the outpatient department for bariatric surgery Spaarne Gasthuis, 0.4 fte per week	2016-2020	96

3. Parameters of Esteem	
	Year
Grants	
- ZonMw Rubicon	2023
- Niels Stensen Fellowship	2022
- ZonMw leren van elkaar	2022
- ZonMw Goed Gebruik Geneesmiddelen Rediscovery	2022
- Phoenix Foundation	2021
- EFSD young investigator grant	2021
- Subsidie Commissie Stimulering Wetenschap Spaarne Gasthuis (REFER trial)	2020
Awards and Prizes	
- EASL ILC young investigator bursary	2022
- EFSD young investigator award	2021
- EASL NAFLD summit: young investigator bursary	2019
- Krom Fund exchange program with Mayo Clinic Rochester	2018

CURRICULUM VITAE

Abraham Stijn Meijnikman is op 7 mei 1988 geboren in Alkmaar. Hij groeide op in het kleine dorp Sint-Maarten in Noord-Holland.

Na het behalen van zijn vwo-diploma verhuisde Stijn naar Antwerpen, waar hij uiteindelijk met de opleiding geneeskunde startte in 2009. Tijdens zijn studie geneeskunde begon Stijn met wetenschappelijk onderzoek uit te voeren onder begeleiding van prof. dr. Van Gaal en prof. dr. De Block. In zijn vrije tijd was hij vaak te vinden op het voetbalveld maar bovenal op het strand om te kitesurfen.

In 2016 studeerde Stijn magna cum laude af aan de Universiteit van Antwerpen. Ondanks dat hij in Antwerpen kon starten met de opleiding tot internist besloot hij terug te gaan naar Nederland om dichterbij zijn familie te wonen en zorg te kunnen dragen voor zijn vader. Hij begon in 2016 aan zijn promotieonderzoek in de groep van prof. dr. Nieuwdorp en bouwde samen met veel collega's het BARIA-cohort op.

In 2021 begon Stijn met de opleiding tot Maag-Darm-Lever arts in het LUMC en startte hij met de vooropleiding interne geneeskunde in het Groene Hart Ziekenhuis. In 2023 gaat Stijn samen met zijn vrouw Charlotte en hun twee kinderen Hugo en Louise naar San Diego verhuizen. Stijn gaat aan de University of California San Diego als postdoctoraal onderzoeker werken in de groep van prof. dr. Bernd Schnabl voor een jaar. Na dit jaar zal hij terugkeren om zijn opleiding tot Maag-Darm-Lever arts af te ronden en zijn onderzoek voort te zetten.