Obesity: with a focus on diet quality and gut microbiota

Brunkwall, Louise

2019

Document Version:
Publisher's PDF, also known as Version of record

Obesity; with a focus on diet quality and gut microbiota

LOUISE BRUNKWALL | FACULTY OF MEDICINE | LUND UNIVERSITY

Louise started in Marju Orho-Melanders group when conducting her bachelor thesis in Nutrition 2011. She continued to work in the group while studying the Masters of Public Health at Lund university which she completed in 2014, the same year she was enrolled as a PhD Student in Orho-Melanders group. .

Printed by Media-Tryck, Lund 2019 NORDIC SWAN ECOLABEL 3041 0903

FACULTY OF MEDICINE

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2019:65
ISBN 978-91-7619-794-3
ISSN 1652-8220

Obesity; with a focus on diet quality and gut microbiota

Louise Brunkwall

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Agardh Auditorium, Clinical Research Centre Malmö.

Friday 14 June at 1 p.m.

Faculty opponent
Paul O´Toole

Professor of Microbial Genetics

University Collage Cork, Ireland

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION
	Date of issue: June 14 2019
Author(s): Louise Brunkwall	Sponsoring organization

Title and subtitle: Obesity; with a focus on diet quality and gut microbiota

Abstract

Background:

Obesity is one of our times major public health issues as it is increasing world-wide and is a risk factor for several severe diseases and conditions. The risk of developing obesity is multifactorial; genetic predisposition, lifestyle and the microbiome all affect the possibility to keep the balance between energy intake and energy expenditure..

Aim:

The aim of this doctoral thesis is to investigate obesity and its complications by exploring the role of gene lifestyle interaction, diet quality, metabolites and the gut microbiota in a population-based setting.

Methods:

We have conducted the studies within Malmö Diet and Cancer Cohort and Malmö Offspring Study (MOS). Individuals in both cohorts have provided blood samples, anthropometric measurements, lifestyle information and dietary information. Additionally, individuals in MOS have provided faeces samples.

Results: The interaction between sugar-sweetened beverages and genetic predisposition of obesity on BMI was replicated in the first paper, in the second paper association between beverage intake and food groups were systematically investigated. BMI related plasma metabolites were associated with four gut microbes in paper four and in the last paper six we observed associations between gut microbes, especially *Roseburia* with a healthy dietary pattern and prediabetes. Additionally, this thesis includes a review article about microbiota and diabetes and a descriptive paper of MOS, including interim analysis.

Conclusion:

In this thesis we have by studying human genetic variation, gut microbiota, circulating metabolites and beverages intake and diet quality added to the existing body of knowledge of that the risk of developing obesity is multifactorial and highly individual. The question of how to decrease the risk of obesity in individual- and population level by targeting the modifiable risk factors is among the top priorities for public health and need to be addressed by future studies.

Key words		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language: English
ISSN and key title 1652-8220		ISBN: 978-91-7619-794-3
Recipient's notes	Number of pages 75	Price
	Security classification	

Signature  Date 2019-05-09

Obesity; with a focus on diet quality and gut microbiota

Louise Brunkwall

LUND
UNIVERSITY

Coverphoto by Camilla and Linus Kandefeldt

Faculty of Medicine
Department of Clinical Science, Malmö

ISBN 978-91-7619-794-3
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2019

Till Mormor – för att mat är livet

Table of Contents

List of Papers

This doctoral thesis is based on the following six papers

I. Brunkwall L, Chen Y, Hindy G, Ericson U, Barosso I, Johansson I, Frank PW, Orho-Melander M and Renström F ***Sugar-Sweetened beverages and genetic predisposition to obesity in two Swedish cohorts*** American Journal of Clinical Nutrition 2016

II. Brunkwall L, Almgren P, Hellstrand H, Orho-Melander O and Ericson U ***Commonly consumed beverages associate with different lifestyle and dietary intakes.*** International Journal of Food Sciences and Nutrition 2018

III. Brunkwall L and Orho-Melander O. ***The gut microbiome as a target for prevention and treatment of hyperglycaemia in type 2 diabetes: from current human evidence to future possibilities***. Diabetologia 2017

IV. F Ottosson*, L Brunkwall*, U Ericson, P Nilsson, P Almgren, C Fernandez, O Melander and M Orho-Melander ***Connection between BMI related plasma metabolite profile and gut microbiota*** Journal of Clinical Endocrinology and Metabolism 2018 (*= equal contribution)

V. Brunkwall L, Jönsson D, Ericson U, Hellstrand S, Östling G, Melander O, Engström G, Nilsson J, Ohlsson B, Orho-Melander O, Persson M and Nilsson PM ***Malmö Offspring Study – Interim analysis*** Submitted to European Journal of Epidemiology 2019

VI. Ericsson U, Brunkwall L, Hellstrand S, PM Nilsson and Orho-Melander M ***Food patterns in relation to prediabetes and gut microbiota in the Malmö Offspring Study*** Submitted to American Journal of Clinical Nutrition 2019

Papers not included in the Thesis

Brunkwall L, Ericson U, Hellstarnd S, Gullberg B, Orho-Melander M and Sonestedt E ***Genetic variation in the fat mass and obesity-associated gene (FTO) in association with food preferences in healthy adults.*** Food and Nutrition Research 2013

Ericson U, Hellstrand S, Brunkwall L, Schulz CA, Sonestedt E, Wallström P, Gullberg B,Wirfält E and Orho-Melander M ***Food sources of fat may clarify the inconsistent role of dietary fat intake for incidence of type 2 diabetes.*** American Journal of Clinical Nutrition 2015

Ericson U, Brunkwall L, Alves Dias J, Drake I, Hellstrand S, Gullberg B, Sonestedt E, Nilsson PM, Wirfält E and Orho-Melander M ***Food patterns in relation to weight change and incidence o type 2 diabetes, coronary events and stroke in the Malmö Diet and Cancer cohort.*** European Journal of Nutrition 2018

Enhörning S, Brunkwall L, Tasevska I, Ericson U, Tholin JP, Persson M, Lemetais G, Vanhaecke T, Dolci A, Perrier ET and Melander O ***Water supplementation reduces copeptin and plasma glucose in adults with high copeptin: the H2O Metabolism pilot study.*** Journal of Clinical Endocrinology and Metabolism 2018.

.

Populärvetenskaplig Sammanfattning

Övervikt och fetma är ett växande problem över hela världen och en viktig riskfaktor för många sjukdomar såsom diabetes, hjärtkärlsjukdom och vissa cancerformer men även för mentalt välmående. Risken för att bli överviktig beror till största delen på vår livsstil, men våra gener och tarmbakterier spelar också en avgörande roll. I min avhandling har jag försökt besvara frågor kring detta i fyra studier, varav två är publicerade.

En av de livstilfaktorer som studerats mycket i samband med övervikt är läsk. Många befolkningsstudier har funnit att läsk ökar risken för övervikt samtidigt som interventionsstudier där man byter ut läsk mot vatten eller sockerfria drycker ser att deltagarna går ner i vikt. I en stor Amerikansk studie tog man det ett steg längre och tittade på om genetisk risk för övervikt och läsk tillsammans ökade risken för övervikt. De fann att de som hade fler riskgener för övervikt och drack läsk hade en större risk att var överviktiga än de som hade färre antal riskgener eller de som drack mindre läsk. Som alla vet är USA ett extremt land och man dricker mycket mer läsk, men det visade sig att när vi tittade på samma sak i en kohort från Malmö (Malmö kost cancer) och en från Västerbotten (GLACIER) såg vi samma samband.

Kost är otroligt spännande att studera, men det är inte alltid helt enkelt eftersom att vi äter och dricker saker i kombinationer och mönster. Efter vårt fynd med läsk gick vi vidare och tittade på de individer som dricker mycket läsk (mer än en burk per dag), inte så förvånande visade det sig att de åt betydligt sämre generellt än de som drack lite eller ingen läsk alls. Individer som drack mycket läsk åt mindre frukt, grönsaker, fisk och yoghurt men mer glass och pommes frites. Vår slutsats från den studien var att det kanske inte är så enkelt som att läsken ensam orsak en stor riskökning för övervikt, utan kostens helhet. Det är därför viktigt att ta hänsyn till den totala kosten när man studerar enskilda livsmedel.

Ett nytt hett forskningsområde är tarmbakterier och vi har möjligheten att i en annan befolkning i Malmö (Malmö Offspring Study) titta på just tarmbakterier. I en studie tillsammans med en annan forskargrupp undersökte vi tarmbakteriernas samband med grupper av metaboliter i blodet starkt kopplade till övervikt i 900 individer. Vi såg då att några tarmbakterier korrelerade med dessa grupper av metaboliter. Många av de tarmbakterierna är sådana som hjälper oss att bryta ner fiber i kroppen. Man vet att äter man mer fiber så får man också fler av dessa i tarmen och i vår studie såg vi nu att om man hade mycket av de metaboliterna i blodet som man vet är kopplade till ökad risk för övervikt hade man färre av de fibernedbrytande bakterierna, som tidigare kopplats till ökad risk för många sjukdomar. Så igen fick vi en indikation på att kosten spelar en central roll för risken att utveckla sjukdom.

Denna text har publicerats i sin helhet i DOKTORN 2 Maj 2019

Abbreviations

ASB - Artificially Sweetened Beverages

BMI - Body Mass Index

GLACIER - Gene-Lifestyle interactions And Complex traits Involved in elevated disease Risk

GWAS – Genome wide association studies

LC-MS – Liquid Chromatography Mass Spectrometry

LTPA - Leisure Time Physical Activity

MAC - Microbial assessible carbohydrates

MDC - Malmö Diet and Cancer

MDC-CC – Malmö Diet and Cancer Cardiovascular cohort

MOS-Malmö Offspring Study

OPLS – Orthogonal partial least square

PA - Physical Activity

PCA – Principal component analysis

PLS – Partial least square

SCFA – short chain fatty acid

SNP- Single Nucleotide Polymorphism

SSB - Sugar Sweetened Beverages

T2D – Type 2 diabetes

WHO- World Health Organization

Introduction

Obesity is a growing and one of our times most challenging public health issues. It was for a long time a condition connected to wealth and a result of having access to more food than needed. Today, obesity is spread around the globe and the association with wealth is weakening, instead the prevalence of obesity is increasing in developing countries. In the most recent Global Burden of Disease study, non-communicable diseases were found to account for 73% of all deaths world-wide. Half of them were attributed to high blood pressure, smoking, high blood glucose and high body-mass index (BMI). Obesity (BMI<30kg/m^2) is calculated to cause more than 1 million type 2 diabetes related deaths every year [1]. Except for diabetes, obesity is also a risk factor for cardiovascular disease, several cancer forms and mental disorders.

Obesity is always a result of an imbalance between energy intake and energy expenditure. The research question that remains is why individuals seem to have different abilities to maintain this balance. Today it is well known that the ability to keep the balance is highly multifactorial, genetic predisposition and an unfavorable lifestyle are the most established risk factors [2]. Additionally, the gut microbiome might also be a key player in both the pathogenesis of obesity and the risk of developing obesity [3]. Due to that obesity increases the risk for so many diseases it is of absolute importance to understand what determines a person's risk for becoming obese to be able to find the best possible prevention strategies. This would benefit not only the individual; the society would save an enormous amount of resources.

Background

Obesity

Obesity is the result of an unbalance between energy intake and energy expenditure over time resulting in excess adiposity. The World Health Organization defines obesity as having a BMI of more than 30kg/m^2, as an example a person who is 1,75m tall and weighs 92kg has a BMI of 30 [1].

BMI is one measurement of obesity, there are others such as waist-hip ratio and body-fat mass, however in this thesis BMI is used as a proxy of obesity throughout the four studies. When the energy intake is superior to the energy expenditure the excess energy is stored as fat in the adipose tissue. The adipose tissue contains immune cells and the number of cells increases with the volume of the adipose tissue, these cells produce proinflammatory cytokines, which promote low-grade inflammation and contribute to insulin resistance [2]. Additionally, there is a linear relationship between BMI and insulin secretion [4].

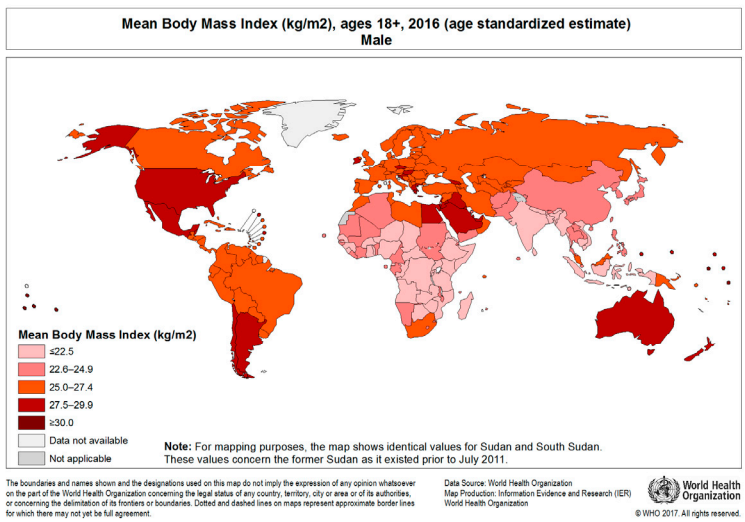


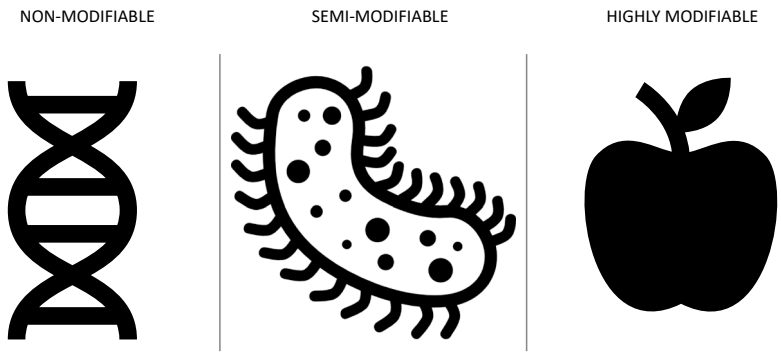
Figure 1 Mean body mass index around the world 2016

Epidemiology

Since 1980, the prevalence of obesity has more than doubled and today ~12% of adults globally are obese (figure 1). This is a real threat to the public health as obesity associates negatively with almost all physiological functions and diseases such as type 2 diabetes, cardiovascular disease some cancer forms and mental conditions. America and Europe are the continents with the highest prevalence of obesity [1]. The trends in obesity progression differ between developed and developing countries and between men and women, globally more women are obese than men while in developed countries men are more obese than women. This might be due to that in developing countries obesity is more strongly linked to the nutrition transition and change in physical activity, while in developed countries the sedentary lifestyle is similar in men and women and therefore diet explains most of the increased prevalence of obesity in these countries. It is well established that women in developed countries in general eat more health associated foods such as fruit and vegetables while men in general eat more meat and drink more alcohol, this can partly explain why men are more obese in developed countries. Additionally, there are biological differences between men and women, such as adiposity storage and lean mass that also affect the risk of developing obesity [5, 6].

In Sweden, a developed country by definition, 15% were reported to be obese in a national survey 2016 [1].

Risk factors



The risk factors for developing obesity can be divided into three categories; non-modifiable, semi-modifiable and highly modifiable. Our genetic make-up is non-modifiable, our gut microbiota is semi-modifiable and our lifestyle is highly modifiable.

Genetics

Obesity is a multifactorial condition where genetic predisposition is one factor. Family-, migration- and twin studies provided the first evidence for that obesity is genetically heritable [7, 8]. A breakthrough in genetic research has been the genome-wide association studies (GWAS) that were introduced in 2005 and investigate the association between millions of single nucleotide polymorphisms (SNPs) in relation to an outcome in a case-control (for dichotomized traits as obesity) or population-based (for continuous traits as BMI) set-up [9]. Today, more than 300 BMI associated SNPs have been identified. Of these, the Fat Mass and Obesity-associated gene (*FTO*) locus is still the one with the largest effect size, where individuals carrying both copies of the risk allele weigh on average 3kg more than those without any risk allele[9]. When added together, the SNPs identified to associate with BMI explain approximately 4% of the variation in BMI [10]. Even though this is a very low number, the genetic predisposition still seems to matter in development of obesity and the susceptibility to environmental aspects.

BMI is a measure of the mass of a person in relation to height, while waist hip ratio (WHR) is a marker of body fat and a measure of where the fat is stored. Interestingly, genes close to the BMI associated SNPs tend to be expressed in the central nervous system while genes close to SNPs associated with WHR tend to be

expressed in the adipose tissue [11]. Many of the genes in the BMI-associated loci are expressed in the hypothalamus indicating that the mechanisms or pathways behind the associations may be connected to appetite regulation. Another suggested mechanism for how the genetic variants may affect obesity risk is related to the browning of fat; individuals with the risk alleles have less browning of the adipose tissue resulting in that the excess fat is not converted to heat, but instead is stored in the adipose tissue [12].

Gene-lifestyle interaction

Even though obesity has a well-established genetic component, environmental triggers are required for its manifestation [13]. It is the interplay between this unmodifiable genetic component and the modifiable lifestyle that determines the risk of obesity. Individuals with a high genetic predisposition but with a favourable lifestyle do not have the same risk as individuals with the same genetic risk but with an unfavourable lifestyle. Wang et al investigated the weight change over time depending on genetic predisposition to obesity and three different healthy diet indices. They found that the risk of weight gain over time was significantly attenuated when adherence to a healthy diet increased. Additionally, they found this effect to be greater in individuals with a high genetic predisposition [13]. Similar interactions have been observed for overall lifestyle but also for individual risk factors such as physical activity and intake of sugar-sweetened beverages (SSB) [14, 15]. Qi et al were the first to observe a more pronounced increase in BMI per serving of SSB among individuals with high genetic risk for obesity than among those with a lower genetic risk. This result suggests that individuals with high genetic predisposition for obesity may be more sensitive to certain environmental risk factors of obesity, such as high consumption of SSB. Additionally, genetic predisposition to obesity might also affect the individuals´ ability to lose weight, according to results from Food4Me, a weight loss trial based on genetic predisposition to obesity, individuals with risk alleles in *FTO* had a more successful weight loss than those without risk alleles [16].

Gut Microbiota

History

Two commonly used quotes by Hippocrates are "*Let food be thy medicine and medicine be thy food*" and "*All disease starts in the gut*". It is almost humoristic that we are still working to figure out how to eat right and what role the gut has for human diseases, however we have never been closer.

The human gut microbiome (which refers to the genomes of bacteria, archea, fungi, viruses etc), in particular faecal bacterial microbiome, has become one of the major focus areas in medical research the last decade. This is largely due to technical progress. In 1995, the first bacterial genome was sequenced [17] and approximately a decade later the sequencing techniques were more established and the prices feasible for "larger" studies. Previously, the main technique to investigate microbiome was culturing. However, culturing is very laborious, and a majority of gut bacteria cannot be cultured, like the anaerobic microbes that are extremely difficult to grow and study. The 16S RNA sequencing was a great breakthrough in the microbiome field and enabled sequencing of larger studies. With this technique, the bacterial DNA is extracted and a specific hypervariable region of the 16S RNA gene is sequenced. The sequences are then binned together and used to identify operational taxonomical units (OTU). This can be done by either matching the reads to a database of known microbial sequences or by letting the reads to cluster together. From the matched reads, OTUs are identified at different taxonomical levels. With 16S sequencing the data is reliable down to genus level [18].

Obesity and the gut microbiome

The gut microbiome was first connected to obesity in germ-free mice studies, from which the following conclusions were drawn; 1) the microbes have a causal effect on development of obesity, 2) a more diverse microbiome is associated with a leaner phenotype, and 3) the bacterial composition determines how much energy that can be absorbed from food [19-21].

In the first studies faecal matter from obese mice were transplanted into germ-free mice, causing them to become obese which was not observed when faecal matter from a lean mouse were transplanted. They identified that the bacterial composition could determine the amount of energy absorbed from the food, if the abundance of fibre degrading bacteria were higher more energy was absorbed through that production of short chain fatty acids (SCFA) [21]. Studies on discordant twins where one was obese and the other lean showed that the obese twin hade fewer bacterial species and a less rich bacterial composition [20]. These conclusions have now been refined and even proven wrong. The association between obesity and the gut microbiota composition exists, however the associated effect is much smaller than these initial studies showed. The effect is not as general and broad as first thought but rather specific and precise [22].

Today

In 2016, the first two population-based studies were published in Science with study populations above 1000 individuals [23, 24]. They observed the microbiome to be associated with an extensive number of the variables they investigated, however the associated effect of each was modest. They observed categories of medication, blood parameters, bowel associated variables, dietary information, health associated variables, anthropometrics and lifestyle to affect the gut microbiome composition in ascending order. Together they explained 16.4% of the variation in gut microbiome composition assessed by the Shannon index, which means that 16.4 % of the variation in the richness of the bacterial composition could be explained by the assessed variables [23]. This and other large studies have been instrumental in understanding what may affect the gut microbiota on a population level and although conclusions may not be drawn because of the cross-sectional design they are highly hypothesis generating and could lead the way to other studies where the associations could be studied more in detail. Additionally, they provide information on the potential confounders to take into account in future studies [25].

There has now been a shift towards deeper sequencing named shotgun metagenomics (figure 2). This is where the DNA is extracted from all micro-organisms within the sample, and deeply sequenced, giving a much more comprehensive picture of the gut microbiome composition. The drawbacks of this technic are the extensive bioinformatic competence needed to process the data properly, many microbes that are identified are unknown and the price is still very high.

One of the aspects to why the microbiome is so complex and interesting is the variation, it has been suggested that individuals might only share a very small part of its microbiome, while 99.9% of the human genetic sequence is the same. Although the technical aspect of microbiome analysis is crucial many studies have shown that the inter-individual variation is still much larger than any technical aspect [26, 27]. In figure 3 one individuals microbiome is visualized from phylum to genus level.

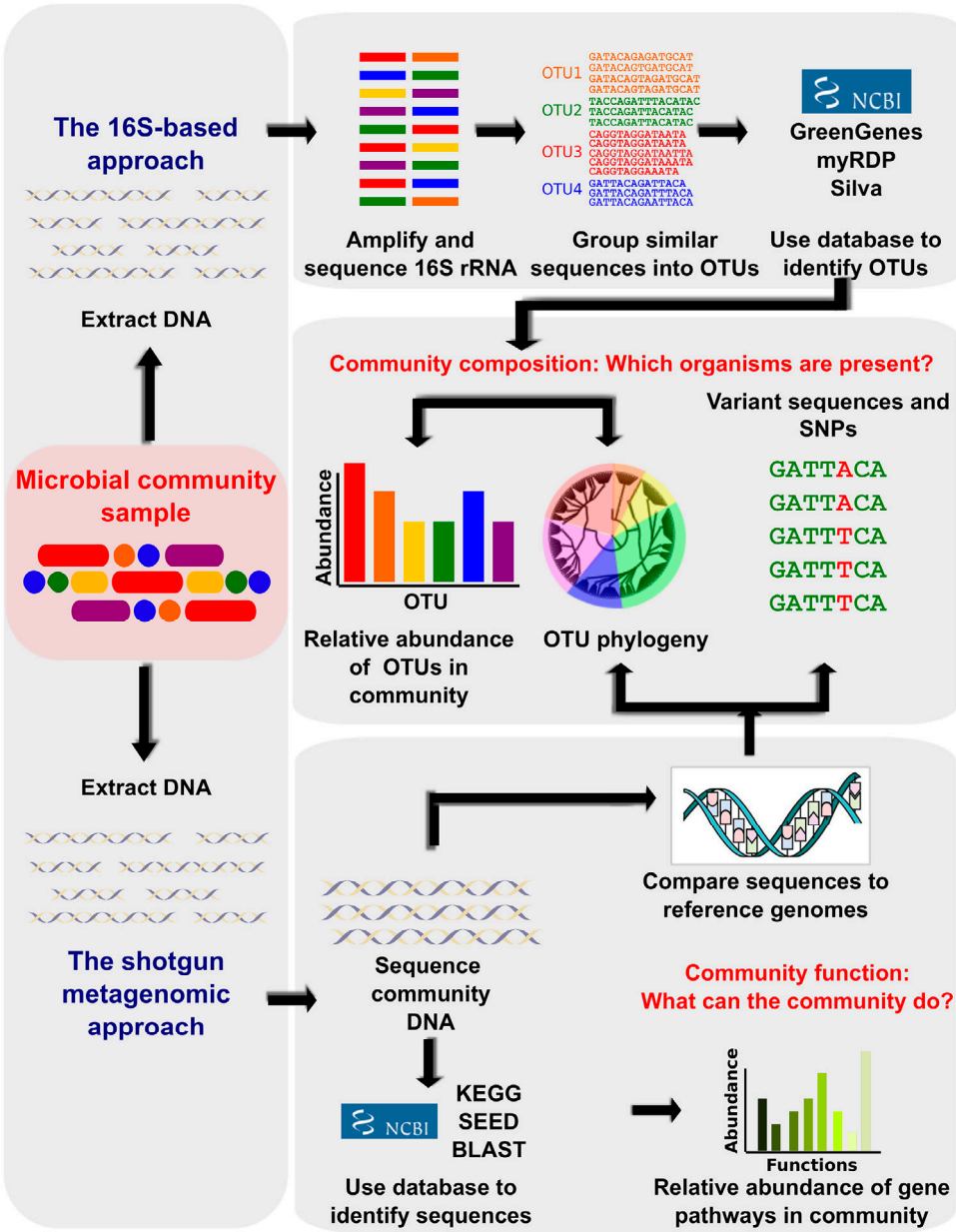
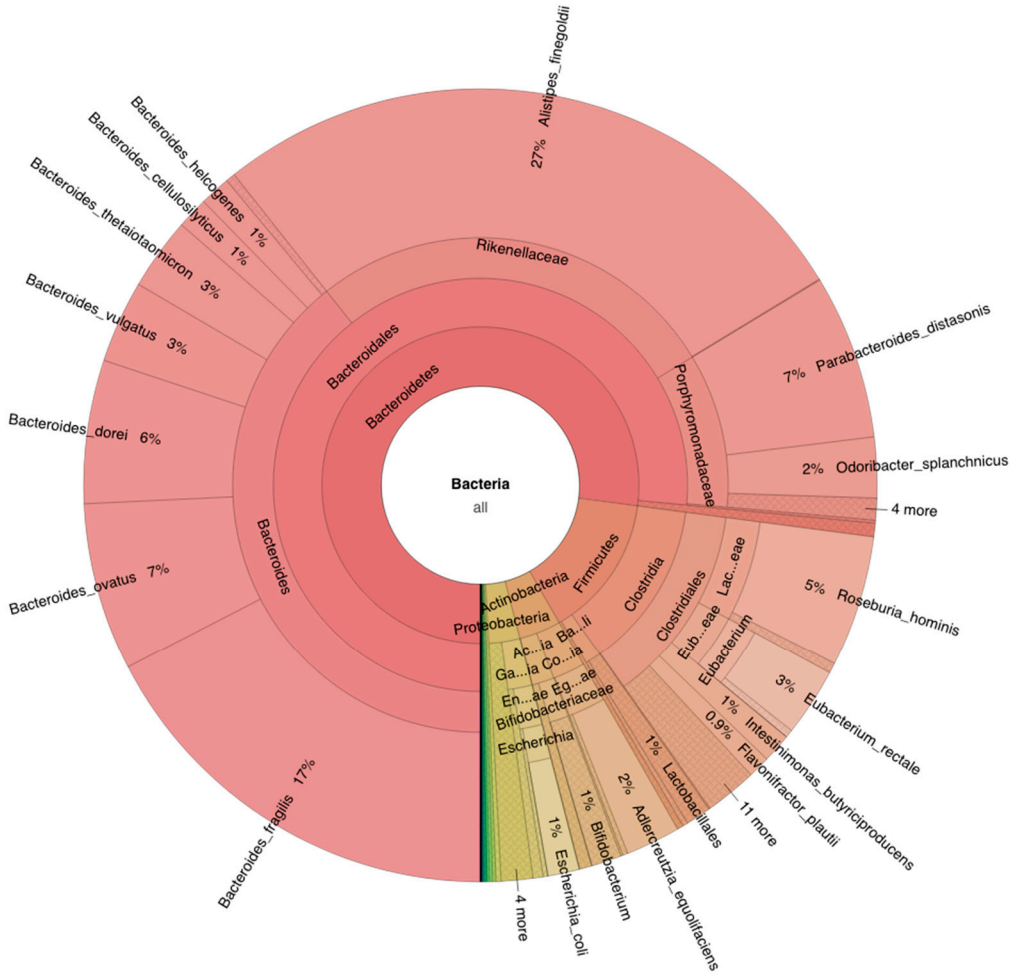


Figure 2 - Human microbiome analysis – difference between 16S RNA and shotgun metagenomic. Figure adapted from Morgan et al[18]



Kingdom	Phylum	Class	Order	Family	Genus	Species
Bacteria	**Firmicutes**	**Clostridia**	**Clostridales**	**Lachnospiraceae**	***Blautia***	***Blautia Obeum***

Figure 3 - One person from MOS – plot visualising the bacterial composition from phylum to species level

The "healthy and unhealthy microbiota"

There is still not enough knowledge to determine what is a healthy gut microbiota, it is most probably dependent on a range of different aspects including age, geography, ethnicity, genetics and lifestyle [28-31]. What is known is that certain bacteria have health promoting aspects such as SCFA production and mucus generation. These two biological processes are essential for many other aspects such as glucose metabolism and intestinal integrity in the host [32]. The bacteria that produce SCFA are fibre degrading bacteria and to keep them in the intestine and to increase their abundance a sufficient fibre intake is extremely important. Additionally, bacteria producing mucus are of importance due to that they help keep the intestinal integrity by producing enough mucus on the inside of the intestinal wall, if the mucus layer is not thick enough macromolecules might leak out from the intestine out in the blood stream, and potentially mediate autoimmune disease [33, 34].

Lifestyle

Our lifestyle might be the most important risk factor for obesity due to its highly modifiable character. Independent on genetics a healthy lifestyle can prevent individuals from developing obesity.

An energy dense diet and an inactive lifestyle increase the risk of developing obesity. Energy dense diets do not only include fat and sugar rich foods. They commonly also lack foods high in fibre such as vegetables and whole grain products. The association between diet and body composition has been observed in both large global cohorts and in controlled intervention studies [35].

Physical activity

A sedentary lifestyle, generally defined as a lack of physical activity (PA) is a well-established risk factor for obesity. PA is usually comprised of the combination of work-related PA and leisure time physical activity (LTPA). PA at work is decreasing in most sectors due to digitalization. Then the LTPA becomes more important to keep the balance between energy intake and energy expenditure. PA helps to keep the balance by consuming energy, primarily while performing the activity and secondly. PA builds muscle mass, which consumes more energy than fat mass and increases the metabolic rate resulting in higher energy consumption even while resting.

Socioeconomy

Our lifestyle is highly influenced by our socioeconomy, it is well established that low education level is a risk factor for developing obesity. Additionally, studies have shown that food availability and food systems affect the choices made by the individual and affect the risk of becoming obese. Recent studies have found the gut microbiome to associate strongly with socioeconomic status [36].

Other lifestyle aspects that increase the risk of obesity but are not within the scope of this thesis are stress and sleeping problems.

Diet

As an exposure

If not living in a war or famine zone, everybody eats, several times a day. Diet is therefore one of the most interesting yet complex exposures. It is continuous, individual, cultural and essential. The food choices made are highly driven by socioeconomy and culture in parallel with day to day variation. These aspects make diet a complex exposure to assess. In nutrition epidemiology the aim is to study habitual dietary intake, this can be done in different ways, food frequency questionnaires, 24-hour recall and food records are among the most common tools used. As with all types of measurement tools there are pros and cons, this has to be taken into account when analysing the data and interpreting the results. In this thesis diet is assessed with two different combinations of a questionnaire and recording methods, with the aim of capturing the habitual dietary intake within the study population. To account for some of the major drawbacks of the measurement methods, energy adjustment and identification of miss-reporters are applied. To divide the food variable of interest with the reported total energy intake is a way to energy adjust, which should be done to control for confounding of the energy intake and to account for potential measurement errors. Identification of missreporters in epidemiological research could be done by calculating the individual biological needed energy intake and relate it to the reported energy intake, this identifies individuals who have reported a too low or too high energy intake [37].

Due to its continuous nature the effects of diet are commonly seen in the extreme groups of excess consumption or lack of consumption. This makes it tempting to investigate one food or beverage at a time. An example is SSB where the amount per day is easily estimated and the variation over weeks and season is fairly small compared to that of other foods. SSB have been observed to associate negatively with many diseases and phenotypes in several populations. However, SSB consumption is highly correlated to lifestyle and other unhealthy foods. This makes the associations between SSB and the phenotype of interest more complicated. Additionally, SSB is a highly accessible product with a low price which also contribute to the consumption pattern. Interestingly, other beverages such as juice that contains the same amount of energy per litre, does not associate with increased risk of disease. This raises the question if it is SSB in itself that cause the increased risk.

Fibre

One of the main components in a healthy diet is a sufficient amount of dietary fibre. In general fibre is divided into two groups, soluble and insoluble. The soluble fibre dissolves in water and are fermented in the colon, the fermentation results in gas and other by-products such as SCFA. The insoluble fibres absorbed water and contribute to what is referred to as the bulking effect, whish leas to an increase feeling of fullness. Fibre has been observed in a countless number of studies to be associated with a lower risk for disease, and the global food recommendations is to increase the intake of fibre [38].

To study dietary fibre has become even more interesting now when the connection to the gut microbiome can be investigated in more detail. Recently, a new classification of fibres has been suggested, microbial assessible carbohydrates (MAC). This definition does not consider the difference between soluble and un-soluble fibres, it focuses on the fibres that could be utilized by the microbes. Which fibers are MACs depends on the individual bacterial composition, if the individual have the bacteria with the ability to ferment the different fibres [39]. When investigating differences between modern diet and traditional diets consumed by rural and isolated populations the difference in fibre intake stands out and it is the lack of fibre in the modern diet that has been suggested to be one of the risk factors for several disease and mortality [40, 33]. Additionally, it has been suggested from animal studies that if the consumption of fibre is too low across generations microbes can be extinct [41].

Dietary patterns

Food and beverages are consumed in combinations and in many cases in specific patterns. Many studies have shown that when investigating which foods correlate with one another clear patterns emerge where individuals tend to eat not just one "good" or "bad" food but a combination of foods associated with either reduced or increased risk of disease.

Diet patterns can be extracted in various ways. However, almost independently of mathematical method the patterns are similar; fruit and vegetables are frequently consumed in combination with other fibre rich products and health associated products such as yoghurt and fish. Additionally, individuals that consume healthy foods commonly consume less of disease-associated products such as SSB, fried food and meat products [42].

Dietary patterns used in epidemiological studies can be data driven food patterns commonly extracted from large data sets with dimensionality reduction or cluster algorithms. Additionally, there are dietary patterns and indices constructed based on previous knowledge such as the Mediterranean food pattern. Consequently, these

patterns all are usually associated with a decreased risk of obesity and cardiometabolic disease[43].

Type 2 diabetes

Hand in hand with obesity, type 2 diabetes (T2D) is one of the fastest growing diseases globally.

In healthy individuals, blood glucose is strictly regulated by homeostatic mechanisms, where insulin and glucagon are the main hormones. Insulin is secreted after a meal and lowers blood glucose levels and glucagon increases the glucose levels between meals. All forms of diabetes are a consequence of an inadequate glucose control. The most prevalent and probably the most complex form of diabetes is T2D, which is characterized by insulin resistance and relative insulin deficiency [44]. In insulin resistant stage, pancreatic beta-cells must produce very high levels of insulin to keep the glucose levels within the normal range and when the need of insulin outruns individuals´ insulin secretion capacity, hyperglycaemia and T2D is developed. In individuals with T2D, insulin production is insufficient which results in hyperglycaemia especially after a meal. To have too high blood glucose over time is a risk factor for several severe conditions such as kidney disease, renal failure, cardiovascular disease etc. If T2D diagnosis is made early and lifestyle changes are successful, the disease can often be reversed or at least kept in good control [45, 46].

Due to that T2D is often a consequence of obesity the risk factors are similar; genetic, lifestyle and microbiota.

Aims

General aim

The general aim of this doctoral thesis is to investigate obesity and its complications by exploring the role of gene lifestyle interaction, diet quality, metabolites and the gut microbiota in a population-based setting.

Specific aims

Paper I: To replicate the interaction between genetic predisposition to obesity and sugar-sweetened beverages (SSB) originally found by Qi et al. Additionally, to investigate how the genetic predisposition, SSB and artificially sweetened beverages (ASB) *per se* associate with obesity in two large Swedish cohorts.

Paper II: To systematically investigate how five commonly consumed beverages (SSB, ASB, juice, coffee and tea) are associated with overall diet and lifestyle in the Malmö Diet and Cancer Study.

Paper III: To review the current and future possibilities of the human gut microbiota in the development of preventive and therapeutic strategies for hyperglycemia and type 2 diabetes.

Paper IV: To identify a BMI associated metabolic profile and to investigate how such profile is associated with the human gut microbiota.

Paper V: To provide a comprehensive reference to the Malmö Offspring Study including study design, detailed method description and interim analysis.

Paper VI: To identify data driven food patterns and investigate how they are associated with prediabetes and the human gut microbiota.

Participants and methods

Cohorts

Table 1 – Baseline characteristic of cohorts applied in this thesis

	MDCS	*GLACIER*	*MOS*
Baseline (years)	1991-1996	1991-2007	2013-2020
N	28,098	4,902	2,644
Participation rate (%)	40	68	46
Mean age (years)	57.9	49.0	39.0
% of women	62	62	52

Malmö Diet and Cancer Study

Malmö Diet and Cancer Study (MDCS) is a population-based cohort, with the original aim of studying how diet high in energy and fat, but low in fibre and vitamins, associates with various types of cancer. All men born between 1926-1945 and all women born between 1923-1950 in the city of Malmö (n=74,138) were invited, via letters and public advertisement, to participate in MDCS. The expanded timespan was used for women to facilitate studies of postmenopausal breast cancer. Lack of mental capacity or limited Swedish language skills were the only exclusion criteria. In total 68,905 individuals were eligible for participation and of those, 30,146 were enrolled and 28,098 completed the baseline examinations. Additionally, a random sample (n=6,103) of participants who joined between October 1991 and February 1994 were invited to participate in an elongation of the MDCS, called MDCS-Cardiovascular Cohort (MDCS-CC), where additional phenotyping for cardiovascular risk factors was performed. Details of the MDCS are described elsewhere [47, 48].

MDCS is used in paper I and II.

GLACIER

The Gene-Lifestyle interactions And Complex traits Involved in elevated disease Risk (GLACIER) is a nested cohort within the Västerbotten Health Survey, which is a population-based cohort in northern Sweden where participants were examined around their 40th, 50th and 60th birthday. Altogether 4,902 GLACIER participants, born in 1932-1957, entered the study between 1991-2007 and have complete genotype and phenotype data [49].

GLACIER is used in paper I

Malmö Offspring Study

Malmö Offspring study (MOS) is a population-based family study where adult children and grandchildren (>18 years old) to the participants in MDC-CC living in southern Sweden are personally invited to participate. In 2013, 10,202 individuals were identified as the source population to MOS, and this number has increased during the study as more individuals have turned 18 years old. Participants visit the study centre on two occasions. During the first visit blood samples were drawn, anthropometrics were measured, and the participants were instructed on how to collect the stool and urine samples at home. Additionally, they were instructed on how to report their diet intake with two different methods. Lifestyle questionnaires were distributed via email.

In 2016, the study had reached "half-time" with 2,644 enrolled individuals (participation rate 46%). All data collected at that point were quality controlled and included in paper V in this thesis.

MOS is used in paper IV-VI

Assessment of main exposures

Anthropometrics

Weight (kg) was assessed by participants standing on a balance-beam scale in light clothing with no shoes. Height (cm) was assessed by a fixed stadiometer. BMI was calculated by dividing the weight in kilograms with the squared height in metres.

Diet

Malmö Diet and Cancer Study Cohort

Diet was assessed with a combined diet history method that comprised of three parts, one 168-item dietary questionnaire, a 7-day menu book and a 45-60 min diet history interview. The method was specially designed for the MDCS[50]. The aim of the questionnaire was to cover habitual dietary intake not assessed by the menu book, which aimed at assessing meals that generally varied between days. The interview then aimed at adding details to the menu book such as portion sizes and cooking methods and to check for overlap between the menu book and the questionnaire. The average daily intake of foods and nutrients (g/day) were calculated from the PC KOST-93 database created by the Swedish National Food Agency. Validity and reproducibility of this method has been investigated and published [51, 52].

Malmö Offspring Study

In MOS diet is assessed by a combination of a 4-day web-based food record and a food propensity questionnaire (FPQ). At the first visit to the research clinic the participants were instructed on how to fill in the 4-day food record and they received a link via email to the FPQ. All participants started the consecutive day after the first visit to get an even spread over the week days. The web-based food record was developed by the Swedish national food Agency (Riksmaten2010) and has been validated by the double-labelled water technique, where the reported energy intake was compared to the energy expenditure [53]. In addition, due to that 4 days do not capture the entire habitual diet, the participants filled out a FPQ, with questions about foods that might not be consumed within the 4 days, but are part of the habitual diet such as fish. The average daily intake of foods and nutrients (g/day) were calculated from the National food database (Riksmaten Vuxna 2010).

Genetics

In paper I, 31 BMI associated SNPs, identified in 2010 by Speliotes et al [54], were studied. An SNP in ZNF608 was excluded from the analysis due to failed genotyping. In MDCS, the SNP's were genotyped by TaqMan, KASPar or Sequenom iPLEX, depending on the availability of the assays and reagents. The averages successful genotyping rate was 98.4% and all SNPs were in Bonferroni corrected Hardy-Weinberg equilibrium (HWE) ($P>0.001$ for 31 independent tests $\propto=0.05$). In GLACIER, MetaboChip, was used to perform genotyping with a success rate of 96.0%, all SNPs were in HWE. In GLACIER, SNP LRP1B rs206936 was not in Bonferroni correlated HWE and no proxy was available, therefor it was excluded and the total number of SNPs used in paper I was 30.

Genetic risk scores (GRS) were calculated with the 30 SNPs, both weighted and unweighted. In the unweighted score the BMI-increasing alleles were summed. In the weighted GRS each BMI-increasing risk allele was weighted by its previously reported effect size and to facilitate interpretation the score was rescaled by using the method published by Cornelis et al [55].

Gut microbiota

Participants in MOS were instructed on their first visit to the research clinic on how to collect the stool samples. They collected them at home in four aliquots (plastic tubes 54*28 mm Sarstedt AB). In the first 2,351 samples (2013-2017) in MOS the 16S gene (V1-V3 region, 300bp×2) was sequenced on a HiSeq Illumina. DNA was extracted by QIAamp column Stool Kit. To process the sequencing data bioinformatically a web-based pipeline was developed where the sequenceses (fastq files) were binned together using FLASH (fast Length Adjustment of Short reads)[56].

To determine which bacteria were available in the samples QIIME was used (Quantitative Insights Into Microbial Ecology), followed by a "closed operational taxonomical unit picking", where sequences were matched to a reference database (Greengenes 13.8)[57, 58].

The output from the pipeline was absolute counts of bacteria on 6 taxonomic levels, from phylum to genus. For further analysis (paper IV and VI) the absolute counts were normalized by Cumulative Sum Scaling (*metagenomicSeq* package in R)[59].

Metabolomics

Plasma metabolites were assessed by a targeted metabolomics approach using mass spectrometry. An in-house developed metabolomic platform has been used in paper IV. Two separate liquid chromatography – quadrupole-time-of-flight (LC-QTOF) methods (one positive and one negative ion mode) were combined to measure 48 metabolites, using spiked-in isotope-labelled internal standards. Using the positive ion method, samples were separated on an Acquity UPLC BEH Amide column, 1.7 µm, 2.1 * 100 mm (Waters Corporation, Milford, MA, USA). The negative ion method utilized an ACE C18 column, 1.7µm, 2.1*100 mm (Advanced Chromatography Technologies Ltd., Aberdeen, UK). In both methods mass spectrometry analyses were performed on an Agilent 6550 QTOF (Agilent Technologies, Santa Clara, CA, USA).

Assessment of other exposures

Lifestyle

All questions about lifestyle were assessed by questionnaires. In MDC it was on paper and in MOS it is web-based.

Smoking

Smoking status was defined as 1) regular smoker, 2) irregular smoker, 3) ex-smoker and 4) never-smoker .

Alcohol

Alcohol intake was assessed in both in the lifestyle questionnaires (frequency) and in the food records (g/day).

Education

Education level was assessed in all papers by questions from the questionnaires. The four levels of education were; 1) primary school (less than 9 years), 2) secondary (9 years), 3) upper-secondary/high school (12 years) and 4) degree from university or college.

Physical activity

In MDC LTPA was assessed by 17 predefined activities and one open activity option, where the participants would estimate the number of minutes per week they spend on each activity depending on the four seasons (e.g swimming, running, gardening, walking). To generate the LTPA score, the time spend on each activity was multiplied with a predefined intensity level. This was adapted from the [60].

In MOS LTPA was assessed with one question; the participants were asked to select the level of activity on his/her leisure time as 1) sedentary (<2h a week of PA), 2) moderate activity (at least 2h of PA per week), 3) regular activity (exercise 1-2 times a week), 4) regular training (exercise <3 times per week).

Other variables

Antibiotics

Antibiotic use was assessed within MOS by a question in the lifestyle questionnaire "Have you been taking any antibiotics within the last 6 months?" (paper IV and VI)

Probiotics

Probiotic use was assessed within MOS by a question in the FPQ "How often do you eat probiotics (in ex drinks, yoghurt, tablets/capsules)?" The participants answered by filling in 1 of 9 frequency categories, from seldom/never to several times per day, in the instructions they were asked to think of their average intake the last 6 months (paper IV and VI).

Prediabetes

Individuals in MOS with a fasting glucose >6.0 mmol/L and/or a HbA1c >42 mml/L at baseline were classified as having prediabetes.

Statistical analysis

Distribution and transformation

In Bayesian statistics normal distribution is assumed, however many variables used were not normally distributed.

Food and beverages variables are in general very skewed, to make it possible to use the data in equations assuming normal distribution all food and beverage variables were logarithmically transformed. Additionally, in paper II all associations were additionally analysed with a non-parametric model, the results were similar and therefor only parametric results were presented.

Gut microbiota data is used in paper IV and VI, in both papers the absolute counts were normalized with cumulative sum scaling (CSS), where the raw counts were divided by the cumulative sum of counts, this is a further development from total-sum normalization where feature read counts are divided by the total number of reads in each sample. CSS is believed to be better suited for 16S data (marker gene technic)[59].

Regressions

In paper I and II linear regression was applied to test for association between a dependent (continuous) variable and one or more independent variables, assuming a linear relationship. When investigating associations for dichotomized depending variables in paper VI logistic regression was applied.

In paper IV associations between microbes and a principal component were analyzed by applying a negative binominal regression model, due to the distribution of the microbes [61].

Meta analysis

In paper I results from the two cohorts were meta analysed to combine the cohort-specific effect estimates from the interaction analysis. Inverse-variance weighted fixed effect meta-analysis was performed by using the *metan* command I STATA.

Dimension reduction

In the medical field today the amount of data and the number of variables are increasing. Correction for multiple testing is one way of handling this, another is to use dimension reduction techniques, which allow us to capture the information in fewer variables.

Coordinates is a way of dimension reduction when it comes to describing a location; instead of describing the position using both east and north of something you use degrees instead. Instead of using two variables to explain the position, one is enough (Figure 4).

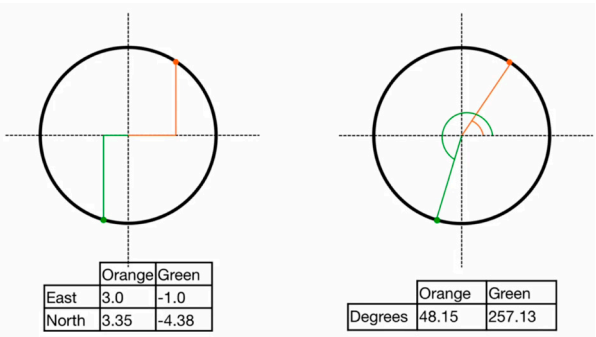


Figure 4 – Dimention reduction

Principal component analysis

To reduce the number of variables (dimensions) principal component analysis (PCA) can be applied. PCA is an unsupervised method where PCs explaining the variation are generated from one matrix. To find out the number of PCs that should be generated one could apply the ”elbow” technic; the possible components are plotted in a scree plot and one can visually see how many of them contribute to the variation. In paper VI, PCA was applied to extract food patterns, 43 food groups were put in the matric and two components were extracted.

Partial least squares regression

Partial least squares regression is a supervised multivariate statistical model that aims in finding a relationship between two matrices (X and Y). The model will try to find what in the X matrix that explains most of the variance in the Y. The results are principal components (PC) explaining variation in Y. Orthogonal partial least square regression (OPLS) additionally separates systematic variation in X to either predictive components correlating to Y or to orthogonal components unrelated to Y. OPLS was applied in paper III; X was 48 metabolites and Y was BMI.

Results

Paper I

This paper is a replication of a paper published by Qi et al in 2012, that investigated the association between SSB intake and obesity depending on genetic susceptibility for high BMI.

Background characteristics

SSB intake was assessed in both cohorts, while ASB was only assessed in MDCS. Due to that SSB was assessed using different methods in the two cohorts, it was harmonized into four categories where one time per day in GLACIER was set to 250 ml/day in MOS, where the first categories was zero or seldom consumption while the last category represented the highest intake of at least one can a day (33 cl).

Individuals in both cohorts were slightly overweight, middle aged and >60% were women. The average intake of SSB was 0.3 servings/day and the average number of obesity associated risk alleles per individual was 27.

SSB intake associates with obesity

In a fixed effect meta-analysis of MDCS (n=21,824) and GLACIER (n=4,905), BMI was 0.18 kg/m^2 higher per increment of SSB intake categories (quartiles) in a fully adjusted model. Additionally, in MDCS each increment of ASB intake was associated with a 0.64 kg/m^2 higher BMI.

Genetic predisposition and intake of SSB interacts on BMI

As expected, the GRS of 30 BMI associated SNPs was associated with BMI in both cohorts. Results using the weighted and unweighted GRSs were very similar and therefore only the results using the weighted GRS were presented.

We observed an interaction in the pooled analysis between the GRS and SSB on BMI, where the magnitude of the association between SSB intake and BMI increased by GRS quartile (figure 5). The interaction analysis was performed on both quartiles of SSB intake and on dichotomized intake of SSB. In the

dichotomized analysis, individuals in the group with a higher intake had on average a 1.29 kg/m^2 higher BMI for each 10 unit increase in GRS. The group with a lowest intake had a 0.83 kg/m^2 higher BMI for each 10 unit increase in GRS, *p for interaction* 0.01.

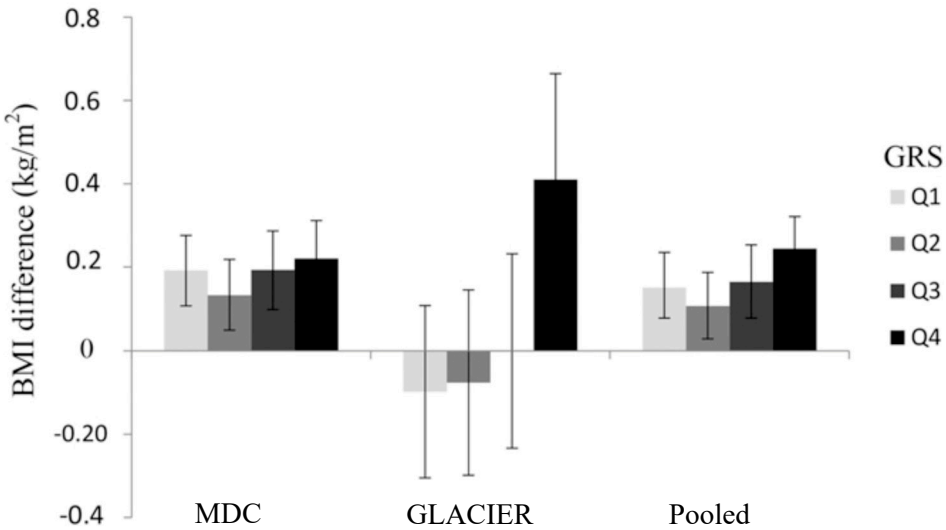


Figure 5 – BMI per 1-increment of SSB intake in MDC, GLACIER and pooled analysis.

Paper II

In this paper we investigated how reported intake levels of SSB, ASB, juice, coffee and tea associates with lifestyle factors and 24 food groups.

Lifestyle associates with beverage consumption

High intake of SSB, ASB and coffee was associated with a higher BMI and a lower frequency of individuals with a university degree, while a higher intake of juice and tea was associated with a lower BMI and a higher frequency of individuals with a university degree. Associations with lifestyle were similar in men and women.

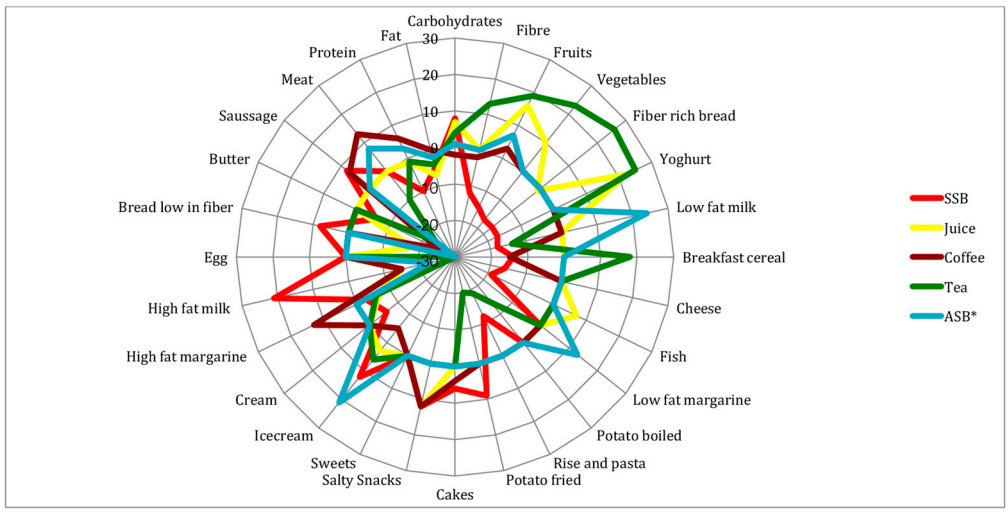


Figure 6 – Difference in % of food and macronutrients intake between the highest and lowest levels beverage intake.

In figure 6 the differences in intake level of each food and macronutrient in percent between the highest and the lowest intake level of each beverage are visualized.

High intake of SSB was associated with a lower intake of foods commonly associated with health such as fruit, vegetables, yoghurt and fish while high consumption of juice and tea associated with a higher intake of these foods. ASB intake associated with other "calorie-reduced" foods such as low-fat margarine and low-fat milk. The intake of coffee did not, as clearly as the other beverages, associate with a specific food pattern.

Paper III

The third paper of this thesis is a review, here we went through the literature to summarize the existing evidence for the link between the gut microbiota and hyperglycaemia with a special focus on human studies. In figure 7 we have tried to encapsulate what was known at that time (March 2017). The main findings form this review was that there is evidence for the link between gut bacteria and hyperglycaemia and T2D, however larger studies are needed to take diet, medication and other potential confounders into account. Knowledge about the gut microbiome and T2D could potentially be used in future personalized treatment strategies.

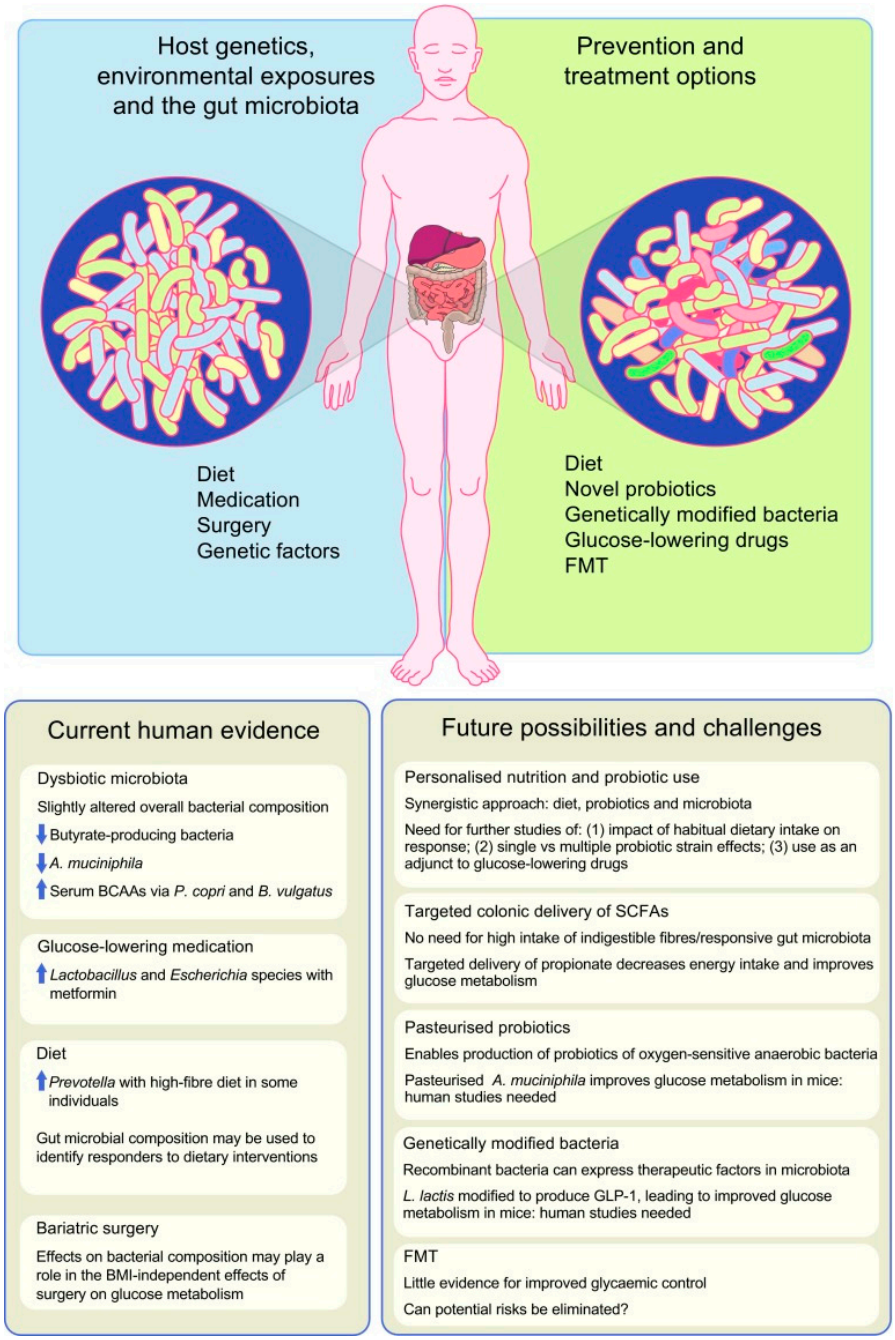


Figure 7 – Main findings from the review "The gut microbiome as a target for prevention and treatment of hyperglycaemia in type 2 diabetes: from current human evidence to future possibilities

Paper IV

In this study the aim was to identify a component constructed of fasting plasma metabolites that explain the variation in BMI, and to associate such component with gut microbes.

BMI associated component

The statistical model OPLS was used, where 48 metabolites were put in the X matrix and BMI as a continuous variable was the Y matrix. From this model, a component (PC_{BMI}) that explained 36.5% of variation in BMI was extracted. The PC_{BMI} is plottet against the orthogonal component in figure 8 colored by BMI. In total, 25 of the 48 metabolites contributed to PC_{BMI} with significant loadings. The strongest contributors to PC_{BMI} were high glutamic acid and branched chain amino acid levels.

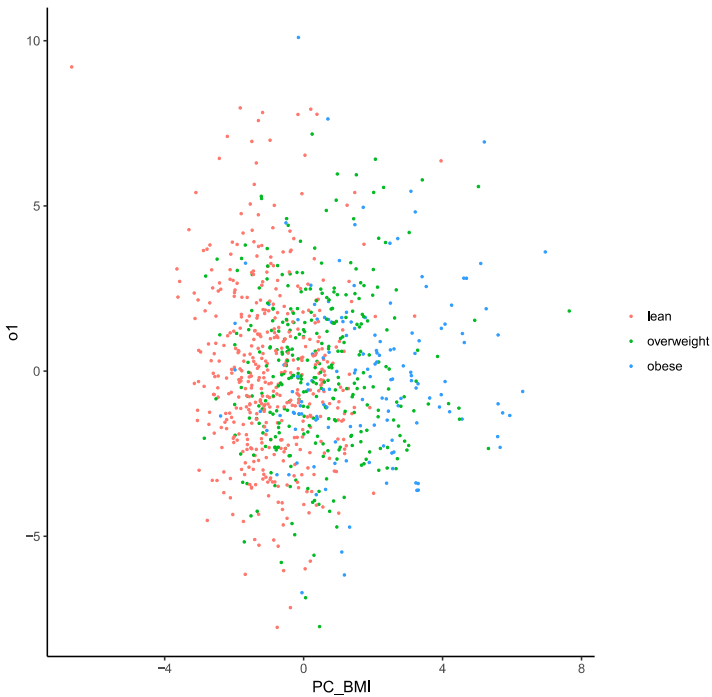


Figure 8 – The predictive BMI component plotted against the orthogonal component, colored by BMI category

The BMI associated component associates with classic risk factors and gut microbes

The PC_{BMI} associated positively with waist circumference, fasting glucose, LDL cholesterol, triglycerides and systolic blood pressure. The PCBMI associated negatively with fiber intake.

The PCBMI was examined in relation to 62 gut bacterial genera and after adjustments for potential confounders three bacteria in the *Lachnospiraceae* family showed significantly positive associations with the component (**table 2**) and SHA-98 showed a negative association. Additionally, we observed a negative association with the Shannon index and positive association with the Bray-Curtis dissimilarity matrix.

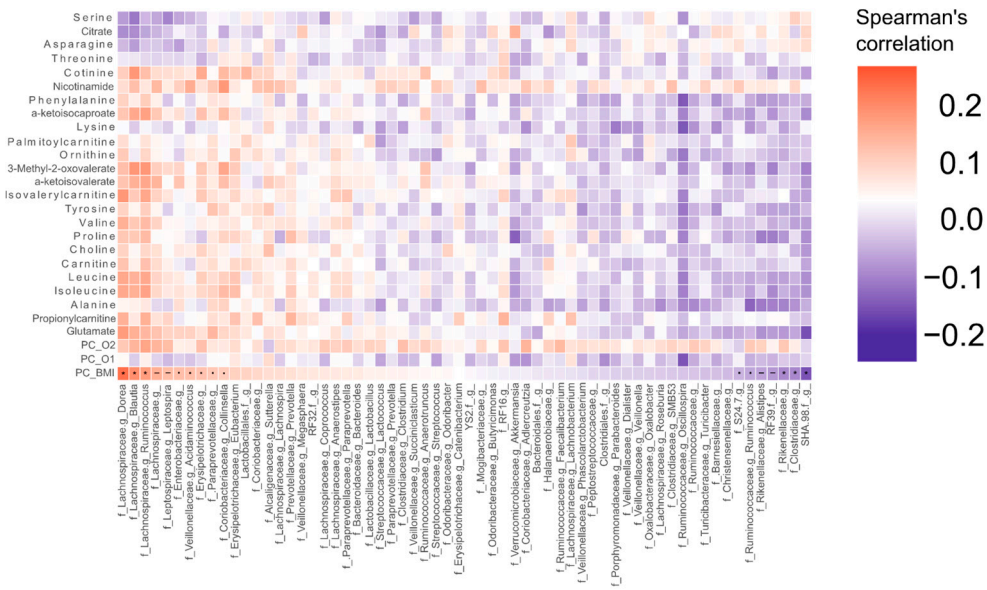


Figure 9 – Correlations between 62 gut bacteria genera and the 25 plasma metabolite that contributed to PC_{BMI}. Additionally, correlations between the predictive component and the orthogonal components with the metabolites.

Table 2 – Associations between gut microbiota genera, BMI and the metabolite PC in MOS (n=674)

Bacteria	Metabolite PC		BMI	
	Beta	P	Beta	P
Blautia	0.20	<0.001	0.06	0.002
Ruminococcus	0.17	<0.001	0.04	0.02
Dorea	0.14	0.004	0.06	<0.001
SHA_98	-0.18	0.011	-0.08	0.004

Paper V

MOS is at the time of this thesis an ongoing cohort. In 2016, MOS had enrolled half of the study population (n=2644). All data collected until then was examined and quality control of the data were done.

Of the 2644 participants, 1326 were children (generation 2) and 1321 were grandchildren (generation 3) to individuals that had participated in MDCS-CC.

Table 3 – Selected backgrounds characteristics from MOS

Characteristics of Malmö Offspring Study participants (n=2644)	
	Means (SD) or %
Age (years)	39.8 (13.9)
Sex (% women)	52.1
BMI (kg/m^2)	25.8 (4.7)
Antibiotics use the last 6 months (%)	14.2
Regular smokers (%)	6.7
University degree (%)	37.9

Paper VI

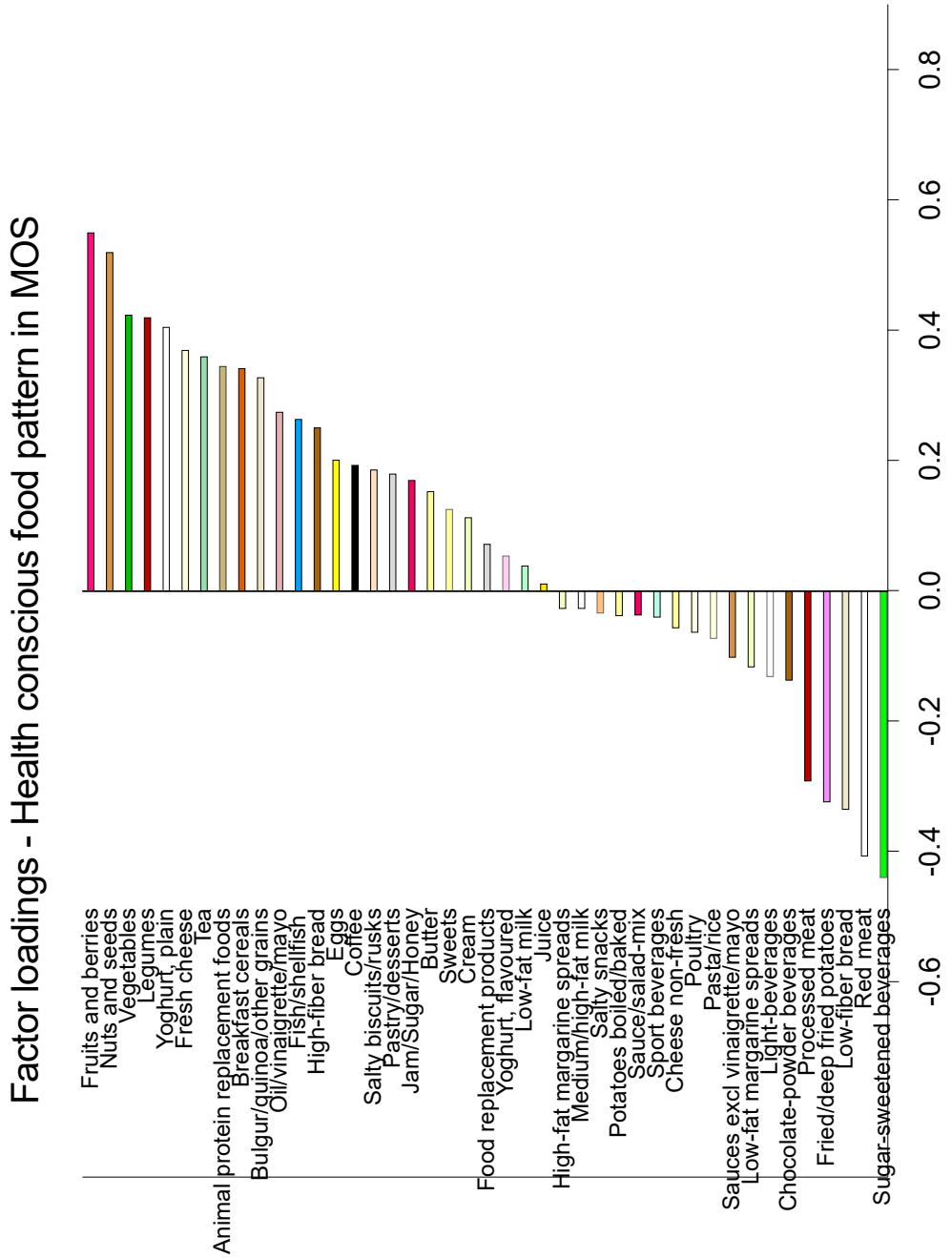
In this study the aim was to extract data driven food patterns and to investigate their association with prediabetes and gut bacteria.

From the included 43 food groups, two patterns were extracted by PCA, *the Health-conscious pattern* and the *Sugar and dairy-fat pattern*.

The Health-conscious pattern associates with prediabetes and gut microbiota

The *Health-conscious pattern* explained 6.8% of the variation in the diet data and was characterized by high intake of fruits, berries, nuts, seeds, vegetables and yoghurt and low intake of SSB, red meat, low-fiber bread and fried potatoes (figure 10).

Figure 10 - Factor loadings of the healthy food patterns in MOS



Individuals with a high adherence to the healthy pattern had a higher education, lower BMI and higher LTPA. The pattern was divided into quintiles and a higher adherence to the pattern was associated with a lower risk of prediabetes. Individuals in the highest quintile were 50% less likely to have prediabetes. However, after adjustment for BMI the association was attenuated, suggesting that the association may at least partially be mediated by BMI.

Additionally, the pattern was associated with a higher abundance of *Roseburia*, *Lachnospira*, a genus in the Streptophyta order and a genus in the RF39 order (figure 12a) as well as with lower abundance of *Blautia*, *Eubacterium* and *Anaerotruncus* (figure 12b).

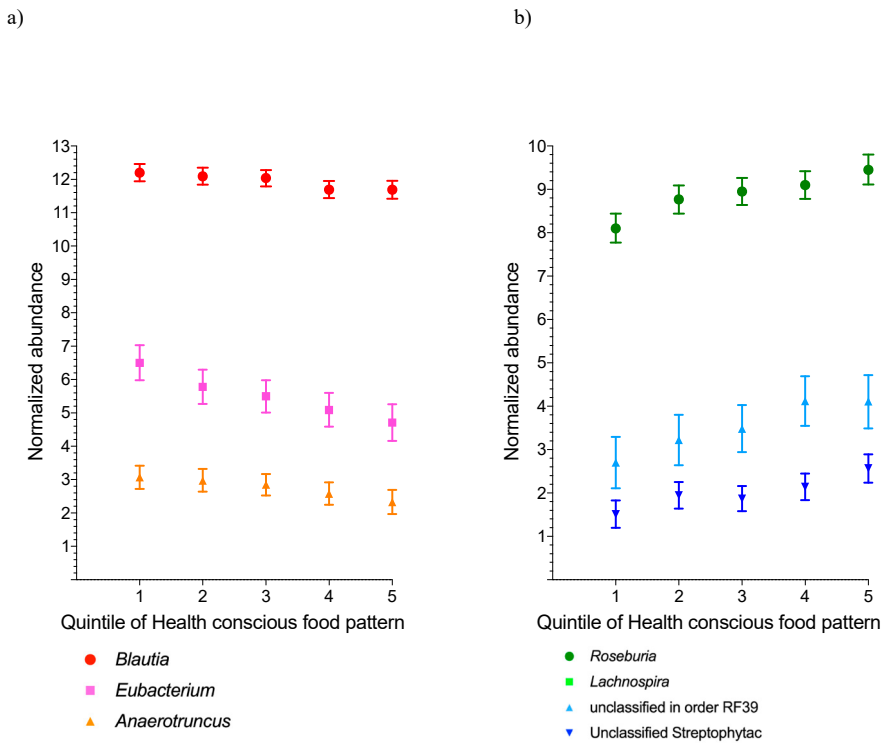


Figure 11a and b – abundance of bacteria associated by the healthy food pattern in quintiles. a) bacteria that associated negatively and b)bacteria that associated positively

The Sugar and Dairy-fat pattern

This pattern did not associate with prediabetes in the overall study population, however an interaction with sex was observed where women had an increased risk of prediabetes if adhering to this pattern, while this was not observed in men. After adjustments no association between the pattern a microbiota remained.

Roseburia associates with prediabetes and the healthy food pattern

Roseburia was inversely associated with prediabetes (figure 13) and the association remained significant after adjustment for the healthy food pattern. However, the association between the healthy food pattern and lower risk of prediabetes was attenuated after adjustment for *Roseburia.*

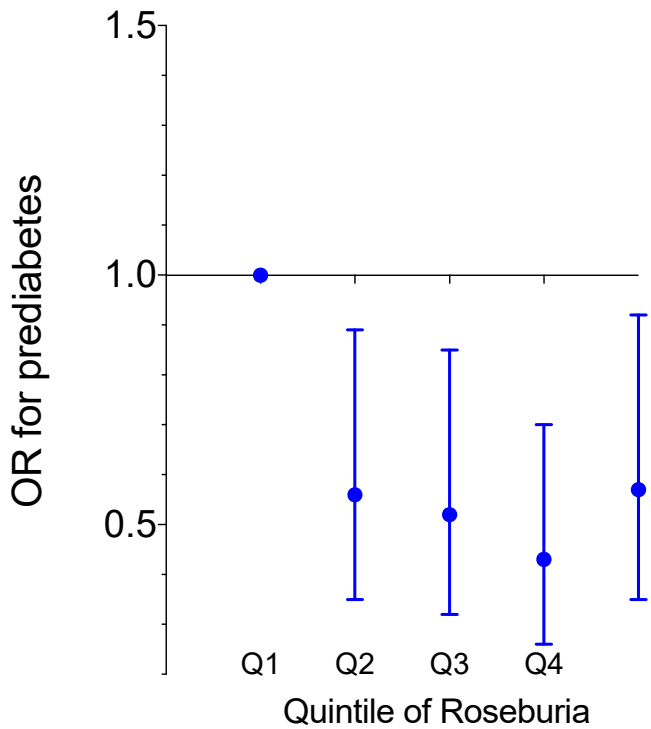


Figure 12 – Odds ratio for the risk of developing prediabetes by quintiles of Roseburia abundance.

Discussion

Main findings

In this thesis we have through six projects, including one review and one design paper, studied obesity, genetic predisposition, diet quality, gut microbiota and prediabetes in a variety of ways. We started out with a project (Paper I) to replicate findings of a study that reported interaction between intake of SSB and genetic predisposition to obesity on BMI. In line with the original findings, we observed that individuals with a higher genetic predisposition to obesity were more sensitive to the BMI-increasing associated effect of high consumption of SSB [62, 15]. To further understand this result, we conducted the second study (Paper II), where we in detail investigated the reported diet and other lifestyle characteristics of high consumers of SSB and four other beverages, and observed that high consumption of SSB, ASB, juice and tea clearly associated with distinct overall dietary and lifestyle patterns, while for high-consumers of coffee such patterns were less apparent [63].

During the course of this thesis the field of gut microbiota has greatly evolved, and an enormous number of reviews has been written. However, in most of these reviews it often remains unclear if the associations from the original papers were observed in animal models or humans. Therefore, our review (Paper III) focused on the current evidence for a connection between the gut microbiota and hyperglycaemia in humans [64]. In 2016, the first microbiota data from MOS was ready and we could initiate analyses for the first microbiota paper. This study (Paper IV) focused on BMI and how BMI related plasma metabolites were associated with gut microbiota, In this study, we observed patterns of BMI related plasma metabolites to associate with four specific microbes [65].

The MOS cohort reached half-way in 2017 and all data collected up until then was quality controlled and that was the basis for the fifth project (Paper V) where the aim was to give a comprehensive description of the cohort and study related methodologies.

With the knowledge from the review (Paper III) and a paper not included in this thesis investigating food patterns in relation to T2D [66], the last study of this thesis

(Paper VI) was designed to investigate how data driven food patterns associate with prediabetes and the gut microbiota. In this study, we observed *Roseburia* to be associated with both prediabetes and a healthy food pattern.

Interpretation of main findings

In this thesis we have replicated and added to the body of knowledge that SSB is strongly associated with obesity, additionally we observed that the magnitude of the association is different depending on the genetic predisposition to obesity, suggesting that there are group of individuals that are more sensitive towards environmental risk factor of obesity. This could potentially be explained by a less fine-tuned appetite regulation because of that many of the BMI associated SNPs examined in these studies are in or nearby genes that are expressed in the hypothalamus. Another aspect might be that the genetic predisposition of obesity affects the food preference, which could mean that individuals with a high genetic predisposition not only consumed more SSB but may also in general consume more unhealthy foods. In the second paper we identified a clear association between SSB intake and overall diet. However, we did not investigate if the association between SSB and unhealthy diet was more pronounced in genetically susceptible individuals.

The role of the gut microbiota in development of obesity has been debated and more research is needed. However, in paper IV we identified an association between plasma metabolites and especially glutamate and BCAA and BMI that associated with four gut microbes suggesting that there is a link between the gut microbiota and BMI that could partly be mediated by the plasma metabolites and especially glutamate and BCAA. Additionally, there might be a link to appetite regulation that could explain the associations between BMI and microbiota, where acetate (a SCFA produced by fibre degradation bacteria) has been found to directly affect the appetite in mice. In the last paper we observed an association between a healthy food pattern and decreased risk of prediabetes that was potentially mediated by *Roseburia*, indicating that a healthy dietary pattern may facilitate growth of bacteria with positive health effects.

Specific elements of the thesis

Sugar sweetened beverages

SSBs contain water, flavour and sugar and have become one of the most debated products of our time. In population-based studies, SSBs have been associated with an increased risk of obesity, T2D, coronary heart disease and several other chronic diseases. SSBs do not contain anything strange or unique that would make them especially harmful. It may rather be the high content of sugar in liquid form that makes them "dangerous". The high sugar content results in an energy dense product that is very easily accessible and consumable. Studies have shown that the amount of energy consumed from SSBs is not fully compensated for by eating less, it just adds energy to the total daily intake. In contrast to candy or other types of sweet products, the liquid form of SSBs makes their consumption very rapid, in just a few minutes one can easily consume a substantial amount of extra energy while it commonly is more time-consuming to ingest solid products, which is an important aspect in appetite regulation. The observation in Paper I, that individuals with a higher genetic predisposition to obesity were more sensitive to the increasing associated effect of [15]SSBs on BMI, was in line with the observations in a previous publication by Qi *et al*, while Tyrell *et al* did not observe similar interaction in a British data set[67]. This could be due to that they adjusted the analyses for an overall Western diet, which is a potential confounder for the interaction. This is in line with our findings that SSBs are commonly consumed in combination with other unhealthy foods, which we observed in Paper II, where we systematically analysed SSB intake in relation to food groups, and in Paper VI where SSB was correlated with intake of other unhealthy foods in a data driven food pattern. In contrast to SSBs, the intake of juice, which contains the same amount of energy and is in the same liquid form, does not associate with BMI or T2D, which is most likely because it is consumed differently. The intake levels are generally lower, and juice is generally part of a more health conscious diet, as observed in Paper II. Additionally, juice is commonly consumed at breakfast, which in itself is associated with a decreased risk of metabolic disease.

Gut microbes

The gut microbiome has become a new piece in the puzzle of metabolic health, and the research to where it will fit is ongoing. Delivery mode, early life dietary intake, infections, antibiotics, other drugs, lifestyle and diet are believed to affect the microbiome composition. It has been hypothesized that there is a "window of opportunity" in early childhood where the foundation for the microbiome is initiated and it is then valuable not to have many severe infections nor to have several antibiotic treatments. However, what the effect of this is later in life and how much it affects the microbiome compared to the adult lifestyle is still debated.

Genetics

In Paper I, we used a GRS for BMI based on 32 BMI-associated SNPs. Even though studies have estimated BMI to be highly heritable, the associated effects of the identified individual SNPs are limited, and therefore using risk scores is one way to increase the effect size by summing the effect of each SNP to one score. This approach has been widely used to increase the statistical power when investigating genetic risk for complex diseases, conditions and quantitative traits.

Many of the genes associated with obesity have been found to be expressed in the hypothalamus and thereby potentially effecting the appetite regulation. Interestingly, gut bacteria have also been linked to appetite regulation through acetate, in mice. Non-digestible carbohydrates are fermented by bacteria in the gut and SCFAs, including acetate, are produced. Acetate has been found to associate with hypothalamic activity suggesting a direct link between SCFA and appetite regulation [68].

Methodological considerations

Epidemiology

Nutrition epidemiology has always been an easy target for criticism; dietary data is commonly self-reported and often based on FFQs, the effect sizes are usually small and the confounders are uncountable. However, this research area has despite all these problems provided important knowledge [40]. An example of this are the epidemiological findings of that trans fats increase the risk of coronary heart disease; observations which were further tested and confirmed in randomized controlled studies, and this knowledge have now led to a significant reduction in consumption of trans fats globally. Additionally, coffee intake has been observed in smaller studies to be hazardous and in California there was a suggestion to put warning labels on coffee. However, after investigating the intake of coffee within populations in epidemiological studies, clear evidence now exists for that coffee intake instead is beneficial and reduces the risk of developing T2D [69].

Epidemiology is an important resource when studying research questions that cannot be examined within an intervention study [70]. Studying individuals asked to have a high daily SSB intake for several years would for example not be approved by an ethics committee, given the knowledge we have today regarding SSB intake. However, in an epidemiological setting we have the potential to challenge questions regarding the consequences of high SSB intake. Knowledge about the actual intake in the general population is of importance when trying to understand what is healthy, while intervention studies and other types of more controlled studies are important for generating mechanistic understanding. Controlled studies are commonly limited in size, and the diets investigated are more extreme and the time of exposure is much shorter than what can be studied in an epidemiological setting which limits generalizability from controlled studies to real life settings. The aspect of time is one of the main strengths in epidemiological studies, as many diseases take a long time to develop, which means that the optimal way to study this is to examine individuals and collect data about their lifestyle while they are healthy and then follow them to be able to investigate if the incidence of disease differs depending on exposure levels of different lifestyle factors.

Recently it has become possible to study the human gut microbiome within an epidemiological setting, which can add to the understanding of what characterizes a healthy microbiome, and to provide a more nuanced picture of what is potentially harmful for the microbiome and what promotes a “healthy microbiome”.

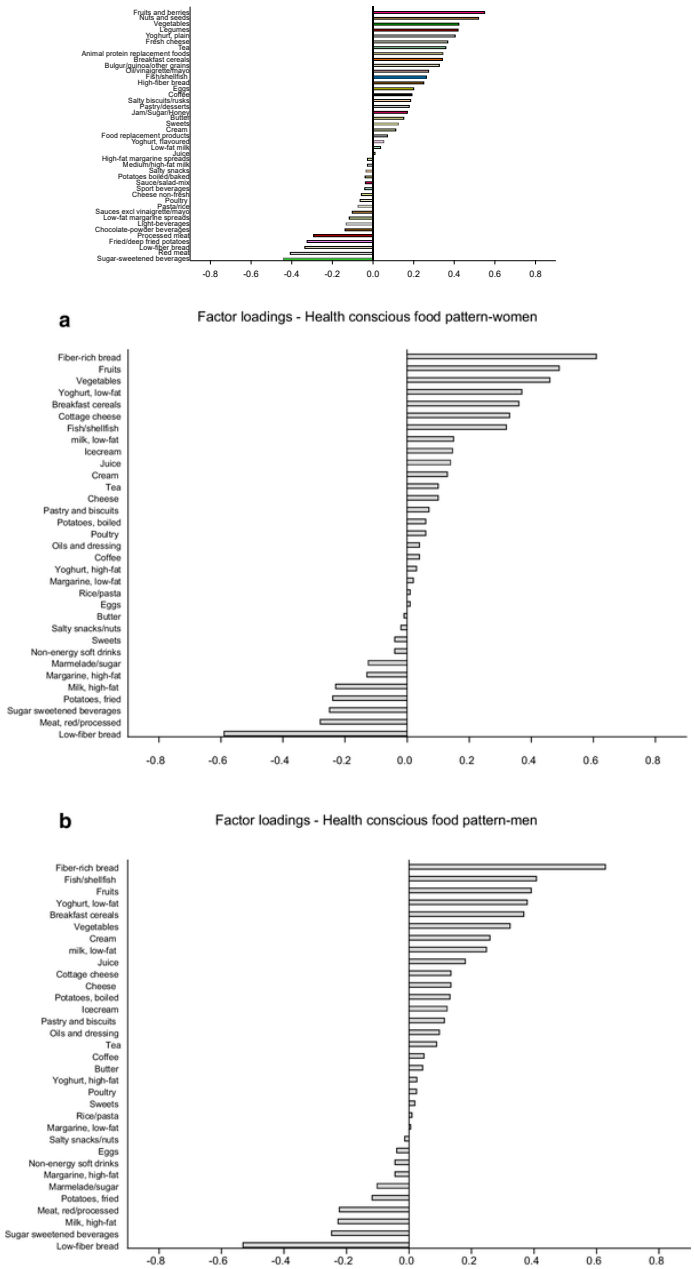
Diet data

Diet is difficult to assess properly and even though much effort has been put into improving diet assessment tools, several considerations and limitations remain and need to be taken into account when analysing dietary data. The "relative intake", the relationship between the food of interest and other food components and the accuracy of the reported intake are three important aspects to take into consideration when working with self-reported diet data. Energy adjustment is one of the most basic things in nutrition epidemiology and in all papers containing diet data in this thesis, we have used intakes relative to the total energy intake. Due to that many foods are correlated, effort is made to try to take other aspects of the diet into account that might confound the potential association. Fibre intake can be used as a proxy of an overall "healthy" diet and treated as a confounder in statistical analyses, due to that most of our fibre intake comes from fruit, vegetables and whole grain products, which are known components of a healthy diet. If the reported intake is far from the true intake the potential associations found could be biased, therefor suspected miss reporters can be excluded from the study population. Miss reporters are identified by dividing the reported energy intake with the assumed energy intake (based on metabolic rate and reported physical activity). It is known that overweight and obese individuals tend to underreport to a higher extent than normal weight.

Generalizability

The MDCS was conducted between 1991-1996 and a common question by reviewers and others concerns the almost 30 years old data and how generalizable the dietary habits and consequently the results from the MDCS are to today´s population. This is a difficult question to answer. However, MOS is an ongoing study, in which dietary data is collected. Interestingly, the diet patterns generated by principal component analysis in MOS in Paper VI are almost identical to the patterns generated in a recent paper in MDCS using the same method to generate dietary patterns (figure 13) [66]. Not claiming that the diet in Malmö today is the same as it was in the 1990s, but foods seem to be consumed in similar combinations and healthy foods are consumed together with other healthy foods. It should be mentioned that the individuals in MOS are relatives to the individuals in MDCS, however the patterns extracted in both cohorts are similar to what has been observed in other cohorts. This strengthens the generalizability of findings using dietary patterns in MDCS.

Figure 13-Loading plots of the healthy food patterns generated in MOS (see figure 11) and the healthy food patterns generated in MDC a in women and b in men.



Patterns vs single variables

In Paper I, we examined intake of a single beverage, SSB, in Paper II we examined how SSB and four other beverages related to other foods and in Paper VI data driven diet patterns were examined. How a single food associates with a disease or a condition is difficult to analyze, due to that it correlates with the intake of several other food and beverages which increases the probability of an over adjusted statistical model, if trying to include the correlated foods as confounders. The food patterns are therefore an option when trying to investigate the association of healthy or unhealthy foods and can also be used in a statistical model as a confounding factor to adjust for the overall diet. However, if there is one food or beverage that is responsible for the association with a disease or condition this could potentially be missed in a dietary pattern analysis. Therefore, the preferred approach depends on the research question.

Microbiota data

In this thesis all microbiota data was generated by 16S sequencing. There are other methods such as shotgun metagenomics that can be used to assess the microbiome data, however at the time of the analysis of the samples included in this thesis, the price was much higher than it is today and the knowledge on how to work with such data was limited. Independent of sequencing method, the DNA extraction method is a very sensitive step and can greatly affect the final results. The cell wall of the bacteria needs to be broken to extract the DNA and the structure of the cell wall determines if the bacteria is a gram positive or -negative. To break the cell walls, different methods can be used including mechanical (small beads) or chemical (lysis) breakage. These processes affect the amount of DNA extracted from the different bacteria and if the process is too rough there is a risk of destroying DNA while if the process is to weak there is a risk of not detecting bacteria with thicker cell wall.

Other factors that can affect the final results are how the sample is handled, which region of the 16S gene is sequenced and how long the sequencing read is, how the OTUs are identified and how the data is normalized. However, independently on all these factors, studies have shown that the inter-individual variation of gut microbiota exceeds the technical.

Additionally, the microbiome is affected by for example antibiotics and diet and to take this into account we have performed sensitivity analysis where individuals reporting to have taken antibiotics and/or probiotics within the last six months were excluded. In general, this has not markedly changed the results, which could indicate that the use of antibiotics or probiotics does not affect the associations. However, we cannot exclude that this could be due to that the antibiotics were taken rather

long time before the sample was collected or that an antibiotic with local effect was used, such as antibiotics for urinary infections. More information about the type of antibiotics and a more exact time of when the participants took the drug is generated at the moment utilizing data from the *Swedish Prescription Register* and will be important aspects to study.

Generalizable – how much of the microbiota is universal

Most probably a very small part of the microbiome is universal thereby limiting the generalizability of microbiome research and results overall. Geographical region is one of the aspects that has been found to associate strongly with the overall microbiome composition. Hence, there are studies that have identified a core microbiota in a number of populations. However, the core is very limited in relation to the number of bacteria identifiable in the gut. The variability and DNA extraction technics and sequencing methods adds to the difficulty of generalizing the results across other populations.

We study the gut microbiota by using stool samples although the microbiome differs all the way through the GI tract. However, the stool sample is clinically the most feasible and it is also where most bacteria are found.

General limitations

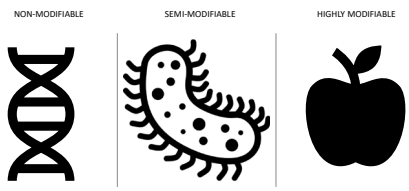
The major limitation of this thesis is the cross-sectional study design where we are unable to draw any firm conclusions and understand potential causal relationships. The self-reported diet and lifestyle data is an additional limitation, as it is well known that there are potential problems with for example underreporting of energy and overreporting of physical activity. The 16S sequencing used to identify bacteria from stool samples also limits us in what we are able to detect in the sample. A deeper sequencing method would enable investigating bacteria further down the taxonomical tree.

Effect sizes

Both genetic and diet studies in general have problems with small effect sizes even though the associations may be highly statistically significant. Further, the cohorts used in this thesis most probably suffer from "the healthy worker" effect, where individuals taking part in a study are healthier than the once that do not participate. This could contribute to underestimation of effect sizes, which would mean that the effect of diet might be larger than what is observed in the studies.

Future perspectives and challenges

Non-modifiable, semi-modifiable and highly modifiable risk factors in future public health strategies to prevent obesity and type 2 diabetes



The risk of becoming obese or developing T2D is dependent on several risk factors where some are very static and non modifiable such as the gene sequences one is born with, while others are more dynamic and modifiable such as the gut microbiome and lifestyle.

One could argue that even the genetic risk is modifiable, due to epigenetic changes, and that certain aspects of once lifestyle such as physical activity can impact the level of expression of some genes. However, we are born with a specific genetic make-up which we in general cannot change, that also in part determines which epigenetic changes that are possible[71].

Based on the knowledge we have today, the gut microbiome is more modifiable than our genes but less modifiable as compared to lifestyle. What shapes a person's microbiome is most likely a combination of genetics, early life aspects and current lifestyle including medication and diet. Animal studies have found diet to be more influential on the microbiome than genetics and genome wide association studies for gut microbiota have found evidence for that only a few bacteria may be connected to the human genome[72]. However, the studies so far have been very limited in size and have mostly not included species or subspecies level of data. Therefore, further studies are needed to understand to what extent the microbiome

is “pre-programed” from our human genes and what is a result of our lifestyle. One potential way to do this is by applying Mendelian Randomization (MR), where a SNP (or usually a combination of SNPs) associated with the exposure is used as an instrumental variable. MR analysis tests if the instrumental variable infers a causal relationship between a a microbial feature such as a bacteria, a pathway or a microbial fermentation product (exposure) and a metabolic phenotype (outcome). This method was recently applied by Sanna et al who identified potential causal relationships between SCFAs butyrate and propionate and insulin secretion and risk of T2D [32]. Bi-directional MR can further challenge the question of directionality i.e. if it is the microbial feature that affects the disease or if it is the disease that affects the microbial feature.

Lifestyle, including dietary habits, is highly modifiable and one could therefore argue for that it is the most important group of risk factors. A study systematically investigating dietary habits in 195 countries where researchers have calculated that one of five deaths could be prevented by improved dietary habits, such as lowering the intake of sodium and increasing the intake of fruits and whole grains, the three food groups found to have the largest effect[73]. This study, our studies and many other studies are building knowledge needed to design large public health interventions, to increase the overall diet quality and physical activity level and thereby reduce the number of obese individuals which would decrease the risk for several diseases.

Precision medicine – will it become reality?

In almost all types of interventions a gradient of responders is observed, from those who respond well and get a beneficial effect of the intervention to those who do not respond despite participating according to the study protocol. There can be several reasons for this, but one important contributor could be the gut microbiota composition of the study participants while entering the study. There are studies observing this for both dietary interventions and drug interventions, where the response of the food or drug is dependent on the gut microbiota. If these findings can be replicated and further understood, they might open up for a new way of thinking and hopefully result in either microbiome interventions prior to the other interventions to introduce the needed bacteria, or that the interventions could include probiotics with bacteria essential for the drug or food to be metabolized in the best way.

Microbiota - what is in the crystal ball?

The gut microbiota field has two main roads ahead, the “population road” and the “mechanistic road”. There is a need for both and the combination might enable new therapies and a more personalized health care. The “population road” is where we need more information about the overall gut microbiota, how it changes over a lifespan, how different factors such as diet, medication, common diseases etc affect the overall microbiome at different ages, and what role microbiota play for future risk of diseases. In parallel there needs to be a focus on the “mechanistic road” to understand how the different microorganisms affect us, which metabolites are produced by them and which are of importance, and how these affect the risk of metabolic unbalance and may translate to health or disease. Additionally, a better understanding of the gastrointestinal pathophysiology is needed, to understand the role of the mucus, the permeability of the gut and their interplay with the rest of the body.

It is also important to understand how the gut bacteria interact with each other, and with viruses and bacteriophages. It will be important to challenge the question of if certain bacteria can have different functions and effects on our health depending on which other bacteria are there, or if viruses are present or not, as well as to understand how the abundance of different bacteria and virus are dependent on other microorganisms in the ecosystem. This would open up not only for new therapies but also for novel prevention strategies.

Our genetic makeup is non-modifiable, still it does not have to determine the risk of developing multifactorial obesity. A healthy lifestyle has the ability to reduce the risk even in individuals with a high genetic predisposition to obesity. The genetic studies investigating the association between genetic risk and obesity identifies groups who are more sensitive towards the environment that we live in now. This could be important when initiating public health strategies due to that the resources are always limited and therefore it might be important to focus on groups that will have the most beneficial effects of the intervention.

Conclusion

Obesity is always a result of an imbalance between energy expenditure and energy intake and it is a combination of non-modifiable and modifiable factors that in combination determine an individual's ability to keep this balance that determines the risk of developing obesity. In this thesis we have by studying human genetic variation, gut microbiota, circulating metabolites and beverages intake and diet quality added to the existing body of knowledge of that the risk of developing obesity is multifactorial and highly individual. The question of how to decrease the risk of obesity in individual- and population level by targeting the modifiable risk factors is among the top priorities for public health and need to be addressed by future studies.

.

Acknowledgements

It takes a village to raise a baby or to write a thesis! This thesis a team effort from all aspects. In particular I would like to thank;

All the participants in Malmö Diet and Cancer Study and Malmö Offspring Study, this would never have been possible without you. Thank you to everyone that have worked or is working with the recruitment of participants at KFE, you are of essence to everything we do!

Marju, your enthusiasm for research is contagious and you have made my Phd journey into something special. You have guided me through the research projects as expected, but more importantly you have given me the opportunity to take responsibility and to develop many other skills which I will be forever grateful for. Finally, I would like to thank you for all the trips we have taken together, to more or less "fun" destinations, in the end we have always had a really good time!

Ulrika, ever since you helped me during my Bachelors Thesis time you have been instrumental in everything I have done and I have learned so much from you. You have corrected all my mistakes, challenged my mind and been an amazing support. On top of that I consider you a friend and an inspiration!

Sophie you are not only a great colleague you are also a wonderful friend. Without you all days at the office would not have been half as fun.

Olle, I'm very thankful that I have been included in some of your projects! **Filip**, thank you for always answering my questions and for all the good times we have had, at CRC, in the slopes and various work-related dinners. **Celine**, you are a brilliant! You are intelligent, fun and dear to become really angry with me!

Current and former group members; **Malin** thank you for always having time for a talk or answering a question, **Gunilla**, what would we do without you…, **Isabel, Tanja, Stanley, George, Christina, Peter A and Anna** thank you for making me improve over the years with your experience and challenging questions.

All members of the **Nutrition Epidemiology group, especially Emily, Joana and Stina** and **Olle Melanders** group; you are all extremely smart, generous and wonderful colleges!

Johan, I sincerely hope that you realise how important you are for me and this project. Except for all practical things you have actually done, most of all you have always welcomed me with a smile when I entered you room, even though it was the 10th time that day!

Thank you **Peter N** for encouraging me and making me part of Malmö Offspring Study!

There are so many other people at **CRC** that have helped me and given me opportunities over the years and provided me with encouraging words. Thank you for making my Phd time so great!

Liberty Bells – thank you for all the fun things that we do together and for being interested in what I do at work. I hope we have many more weekend-trips, party nights and play dates ahead! Special thanks to Linus and Camilla for the lovely front page of this book and help with other graphical tasks I have handed you over the years.

Emelie – thank you for dragging me out of the office for lunch or the house for a AW and for always being a "bollplank"

Camilla, you took this lost Skånetös under your wings when I moved to Stockholm and we had a really good time during the bachelor in Nutrition, where we clearly stated that neither of us should get into research… I look forward to your defence!

And all other friends and extended family that have made life as fun as it is.

Ursprungsfamiljen, without you I'm nothing. I do not believe any other family have downloaded an EASD presentation and watch it together in the couch the same day it was available. The amount of support you give me is unbelievable!

Mamma och Pappa you have given me the best gift possible, a self-confidence. You have made me believe that I could do anything as long as I put some effort into it. This is all thanks to you!

Amelie och Pernilla you inspire me more than you could imagine, your courage, your hard work ethics and your ability to always appreciate a glass of bubbles.

Fredrik, without you there is no air… I could never have dreamt of a better husband. You are my soulmate and inspiration! I love you truly, madly, deeply

Alma, you make it so much harder to go to work, but easier to find motivation to do something that potentially matters!

References

1. Collaborators GBDO, Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K et al. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. N Engl J Med. 2017;377(1):13-27. doi:10.1056/NEJMoa1614362.

2. Heymsfield SB, Wadden TA. Mechanisms, Pathophysiology, and Management of Obesity. N Engl J Med. 2017;376(3):254-66. doi:10.1056/NEJMra1514009.

3. Cornejo-Pareja I, Munoz-Garach A, Clemente-Postigo M, Tinahones FJ. Importance of gut microbiota in obesity. European journal of clinical nutrition. 2018. doi:10.1038/s41430-018-0306-8.

4. Ferrannini E, Camastra S, Gastaldelli A, Maria Sironi A, Natali A, Muscelli E et al. beta-cell function in obesity: effects of weight loss. Diabetes. 2004;53 Suppl 3:S26-33.

5. Kanter R, Caballero B. Global gender disparities in obesity: a review. Adv Nutr. 2012;3(4):491-8. doi:10.3945/an.112.002063.

6. Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML et al. The global obesity pandemic: shaped by global drivers and local environments. Lancet. 2011;378(9793):804-14. doi:10.1016/S0140-6736(11)60813-1.

7. Stunkard AJ, Foch TT, Hrubec Z. A twin study of human obesity. JAMA. 1986;256(1):51-4.

8. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body weight and human adiposity. Behav Genet. 1997;27(4):325-51.

9. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94. doi:10.1126/science.1141634.

10. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN et al. Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. Hum Mol Genet. 2018;27(20):3641-9. doi:10.1093/hmg/ddy271.

11. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206. doi:10.1038/nature14177.

12. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C et al. FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. N Engl J Med. 2015;373(10):895-907. doi:10.1056/NEJMoa1502214.

13. Wang T, Heianza Y, Sun D, Huang T, Ma W, Rimm EB et al. Improving adherence to healthy dietary patterns, genetic risk, and long term weight gain: gene-diet interaction analysis in two prospective cohort studies. BMJ. 2018;360:j5644. doi:10.1136/bmj.j5644.

14. Kilpelainen TO, Qi L, Brage S, Sharp SJ, Sonestedt E, Demerath E et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med. 2011;8(11):e1001116. doi:10.1371/journal.pmed.1001116.

15. Qi Q, Chu AY, Kang JH, Jensen MK, Curhan GC, Pasquale LR et al. Sugar-sweetened beverages and genetic risk of obesity. N Engl J Med. 2012;367(15):1387-96. doi:10.1056/NEJMoa1203039.

16. Celis-Morales C, Marsaux CF, Livingstone KM, Navas-Carretero S, San-Cristobal R, Fallaize R et al. Can genetic-based advice help you lose weight? Findings from the Food4Me European randomized controlled trial. Am J Clin Nutr. 2017;105(5):1204-13. doi:10.3945/ajcn.116.145680.

17. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR et al. Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science. 1995;269(5223):496-512.

18. Morgan XC, Huttenhower C. Chapter 12: Human microbiome analysis. PLoS Comput Biol. 2012;8(12):e1002808. doi:10.1371/journal.pcbi.1002808.

19. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. Nature. 2006;444(7122):1022-3. doi:10.1038/4441022a.

20. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE et al. A core gut microbiome in obese and lean twins. Nature. 2009;457(7228):480-4. doi:10.1038/nature07540.

21. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027-31. doi:10.1038/nature05414.

22. Sze MA, Schloss PD. Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome. MBio. 2016;7(4). doi:10.1128/mBio.01018-16.

23. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K et al. Population-level analysis of gut microbiome variation. Science. 2016;352(6285):560-4. doi:10.1126/science.aad3503.

24. Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science. 2016;352(6285):565-9. doi:10.1126/science.aad3369.

25. Jackson MA, Verdi S, Maxan ME, Shin CM, Zierer J, Bowyer RCE et al. Gut microbiota associations with common diseases and prescription medications in a population-based cohort. Nat Commun. 2018;9(1):2655. doi:10.1038/s41467-018-05184-7.

26. Huse SM, Ye Y, Zhou Y, Fodor AA. A core human microbiome as viewed through 16S rRNA sequence clusters. PLoS One. 2012;7(6):e34242. doi:10.1371/journal.pone.0034242.

27. Ursell LK, Clemente JC, Rideout JR, Gevers D, Caporaso JG, Knight R. The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. J Allergy Clin Immunol. 2012;129(5):1204-8. doi:10.1016/j.jaci.2012.03.010.

28. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R et al. Human genetics shape the gut microbiome. Cell. 2014;159(4):789-99. doi:10.1016/j.cell.2014.09.053.

29. Turpin W, Espin-Garcia O, Xu W, Silverberg MS, Kevans D, Smith MI et al. Association of host genome with intestinal microbial composition in a large healthy cohort. Nat Genet. 2016;48(11):1413-7. doi:10.1038/ng.3693.

30. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;488(7410):178-84. doi:10.1038/nature11319.

31. Deschasaux M, Bouter KE, Prodan A, Levin E, Groen AK, Herrema H et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. Nat Med. 2018;24(10):1526-31. doi:10.1038/s41591-018-0160-1.

32. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Vosa U et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. Nat Genet. 2019;51(4):600-5. doi:10.1038/s41588-019-0350-x.

33. Makki K, Deehan EC, Walter J, Backhed F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. Cell Host Microbe. 2018;23(6):705-15. doi:10.1016/j.chom.2018.05.012.

34. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M et al. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell. 2016;167(5):1339-53 e21. doi:10.1016/j.cell.2016.10.043.

35. Collaborators GBDD. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019. doi:10.1016/S0140-6736(19)30041-8.

36. Bowyer RCE, Jackson MA, Le Roy CI, Ni Lochlainn M, Spector TD, Dowd JB et al. Socioeconomic Status and the Gut Microbiome: A TwinsUK Cohort Study. Microorganisms. 2019;7(1). doi:10.3390/microorganisms7010017.

37. Mattisson I, Wirfalt E, Aronsson CA, Wallstrom P, Sonestedt E, Gullberg B et al. Misreporting of energy: prevalence, characteristics of misreporters and influence on observed risk estimates in the Malmo Diet and Cancer cohort. Br J Nutr. 2005;94(5):832-42.

38. Organization WH. Healthy Diet. World Health Organization, who.int. 2018. https://www.who.int/en/news-room/fact-sheets/detail/healthy-diet. Accessed 8 May 2019.

39. Sonnenburg ED, Sonnenburg JL. Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell Metab. 2014;20(5):779-86. doi:10.1016/j.cmet.2014.07.003.

40. Schwingshackl L, Schwedhelm C, Hoffmann G, Lampousi AM, Knuppel S, Iqbal K et al. Food groups and risk of all-cause mortality: a systematic review and meta-analysis of prospective studies. Am J Clin Nutr. 2017;105(6):1462-73. doi:10.3945/ajcn.117.153148.

41. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. Diet-induced extinctions in the gut microbiota compound over generations. Nature. 2016;529(7585):212-5. doi:10.1038/nature16504.

42. Willett WC. Nutritritional Epidemioogy. Third ed. Oxford University Press Inc; 2012.

43. Wirfalt E, Drake I, Wallstrom P. What do review papers conclude about food and dietary patterns? Food Nutr Res. 2013;57. doi:10.3402/fnr.v57i0.20523.

44. Samuel VT, Shulman GI. Mechanisms for insulin resistance: common threads and missing links. Cell. 2012;148(5):852-71. doi:10.1016/j.cell.2012.02.017.

45. Tuomilehto J, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, Ilanne-Parikka P et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med. 2001;344(18):1343-50. doi:10.1056/NEJM200105033441801.

46. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med. 2002;346(6):393-403. doi:10.1056/NEJMoa012512.

47. Berglund G, Elmstahl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. J Intern Med. 1993;233(1):45-51.

48. Manjer J, Carlsson S, Elmstahl S, Gullberg B, Janzon L, Lindstrom M et al. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation. 2001;10(6):489-99.

49. Kurbasic A, Poveda A, Chen Y, Agren A, Engberg E, Hu FB et al. Gene-Lifestyle Interactions in Complex Diseases: Design and Description of the GLACIER and VIKING Studies. Curr Nutr Rep. 2014;3(4):400-11. doi:10.1007/s13668-014-0100-8.

50. Callmer E, Riboli E, Saracci R, Akesson B, Lindgarde F. Dietary assessment methods evaluated in the Malmo food study. J Intern Med. 1993;233(1):53-7.

51. Elmstahl S, Gullberg B, Riboli E, Saracci R, Lindgarde F. The Malmo Food Study: the reproducibility of a novel diet history method and an extensive food frequency questionnaire. European journal of clinical nutrition. 1996;50(3):134-42.

52. Elmstahl S, Riboli E, Lindgarde F, Gullberg B, Saracci R. The Malmo Food Study: the relative validity of a modified diet history method and an extensive food frequency questionnaire for measuring food intake. European journal of clinical nutrition. 1996;50(3):143-51.

53. Nybacka S, Berteus Forslund H, Wirfalt E, Larsson I, Ericson U, Warensjo Lemming E et al. Comparison of a web-based food record tool and a food-frequency questionnaire and objective validation using the doubly labelled water technique in a Swedish middle-aged population. J Nutr Sci. 2016;5:e39. doi:10.1017/jns.2016.29.

54. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet. 2010;42(11):937-48. doi:10.1038/ng.686.

55. Cornelis MC, Qi L, Zhang C, Kraft P, Manson J, Cai T et al. Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. Ann Intern Med. 2009;150(8):541-50.

56. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 2011;27(21):2957-63. doi:10.1093/bioinformatics/btr507.

57. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7(5):335-6. doi:10.1038/nmeth.f.303.

58. Kuczynski J, Stombaugh J, Walters WA, Gonzalez A, Caporaso JG, Knight R. Using QIIME to analyze 16S rRNA gene sequences from microbial communities. Curr Protoc Bioinformatics. 2011;Chapter 10:Unit 10 7. doi:10.1002/0471250953.bi1007s36.

59. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A et al. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. 2017;5(1):27. doi:10.1186/s40168-017-0237-y.

60. Richardson MT, Leon AS, Jacobs DR, Jr., Ainsworth BE, Serfass R. Comprehensive evaluation of the Minnesota Leisure Time Physical Activity Questionnaire. J Clin Epidemiol. 1994;47(3):271-81.

61. Eric V. Regression Methods in Biostatictics. Second ed. Springer.

62. Brunkwall L, Chen Y, Hindy G, Rukh G, Ericson U, Barroso I et al. Sugar-sweetened beverage consumption and genetic predisposition to obesity in 2 Swedish cohorts. Am J Clin Nutr. 2016;104(3):809-15. doi:10.3945/ajcn.115.126052.

63. Brunkwall L, Almgren P, Hellstrand S, Orho-Melander M, Ericson U. Commonly consumed beverages associate with different lifestyle and dietary intakes. Int J Food Sci Nutr. 2018:1-10. doi:10.1080/09637486.2018.1466272.

64. Brunkwall L, Orho-Melander M. The gut microbiome as a target for prevention and treatment of hyperglycaemia in type 2 diabetes: from current human evidence to future possibilities. Diabetologia. 2017. doi:10.1007/s00125-017-4278-3.

65. Ottosson F, Brunkwall L, Ericson U, Nilsson PM, Almgren P, Fernandez C et al. Connection Between BMI-Related Plasma Metabolite Profile and Gut Microbiota. J Clin Endocrinol Metab. 2018;103(4):1491-501. doi:10.1210/jc.2017-02114.

66. Ericson U, Brunkwall L, Alves Dias J, Drake I, Hellstrand S, Gullberg B et al. Food patterns in relation to weight change and incidence of type 2 diabetes, coronary events and stroke in the Malmo Diet and Cancer cohort. Eur J Nutr. 2018. doi:10.1007/s00394-018-1727-9.

67. Tyrrell J, Wood AR, Ames RM, Yaghootkar H, Beaumont RN, Jones SE et al. Gene-obesogenic environment interactions in the UK Biobank study. International journal of epidemiology. 2017;46(2):559-75. doi:10.1093/ije/dyw337.

68. Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S, Brody L et al. The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. Nat Commun. 2014;5:3611. doi:10.1038/ncomms4611.

69. Gunter MJ, Murphy N, Cross AJ, Dossus L, Dartois L, Fagherazzi G et al. Coffee Drinking and Mortality in 10 European Countries: A Multinational Cohort Study. Ann Intern Med. 2017;167(4):236-47. doi:10.7326/M16-2945.

70. Ioannidis JPA. The Challenge of Reforming Nutritional Epidemiologic Research. JAMA. 2018;320(10):969-70. doi:10.1001/jama.2018.11025.

71. Ling C, Ronn T. Epigenetics in Human Obesity and Type 2 Diabetes. Cell Metab. 2019. doi:10.1016/j.cmet.2019.03.009.

72. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C et al. Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe. 2016;19(5):731-43. doi:10.1016/j.chom.2016.04.017.

73. Collaborators GBDRF. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1923-94. doi:10.1016/S0140-6736(18)32225-6.