New pathways to obesity prevention and metabolic health: The relationship between diet and the gut microbiome

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Nutritional Science

at Massey University, Tāmaki Makaurau,

Aotearoa New Zealand

Nikki Rochelle Renall

2020

Abstract

Background

Diet is one of the key drivers of the global obesity epidemic. Based on the results of rodent experiments, the gut microbiota may play an important role in this multifaceted disease. Additionally, the microbiota is known to be influenced by the habitual diets consumed by humans.

Aims and objectives

The aim of this PhD research was to characterise the habitual dietary intake of two New Zealand populations (Pacific and New Zealand European (NZE) women) with different metabolic disease risk and body fat profiles (lean and obese). The first objective of the research was to explore the relationship between habitual macronutrient intake in relation to body fat content and metabolic health markers. The second objective was to characterise *a posteriori* dietary patterns (derived from multiple days of dietary assessment) and to explore the association with body fat content and metabolic health markers. The third objective was to explore the characteristics of microbiota composition in relation to habitual diet (dietary patterns, foods, and nutrients), body fat content and metabolic health markers.

Methods

Between July 2016 and September 2017, Pacific (n=126) and NZE (n=161) women, aged 18-45 years, living in Auckland, New Zealand, were recruited to a cross-sectional study, based on their body mass index (lean and obese) and stratified as having low (<35 % body fat) or high (≥35 % body fat) body fat percentage (BF%). Dietary intake was assessed using a 5-day estimated, non-consecutive, food record and a validated semi-quantitative food frequency questionnaire, which were used to calculate habitual dietary intake using the National Cancer Institute (NCI) method. Body composition and BF% were assessed by dual-energy x-ray absorptiometry. Fasting blood samples were analysed for metabolic biomarkers (lipid and glucose profiles). Bulk DNA was extracted from faecal samples and the metagenomic sequences associated with the microbiota were analysed using MetaPhlAN and QIIME2 software. Enterotypes characterising the microbiotas of the participants were predicted in R and the species that defined enterotypes were determined using STAMP software. *A posteriori*

dietary patterns were identified using principal component analysis. Adjusted multivariate regression models were conducted to explore the association between BF% and habitual macronutrient intake and adherence to dietary patterns, as well as the association between microbiota composition and habitual diet.

Results

There were no significant differences in BF% between Pacific and NZE women (p=0.498). Higher energy adjusted habitual dietary fibre (DF) intake was associated with lower BF% (β= -0.35, $p \leq 0.001$) for both Pacific and NZE women, and this relationship became stronger after further adjustments for protein (g/day), total carbohydrate (g/day), and total fat (g/day) intake (β= -0.47, p≤ 0.001). Women in the highest tertile of DF intake were older, had lower concentrations of fasting plasma insulin, and lower socioeconomic deprivation levels. Four dietary patterns that explained 30.9 % of the observed variance in habitual diet were identified. Higher adherence to dietary patterns characterised by core foods (the "colourful vegetable, plant protein, and dairy" and "fruit, starchy vegetables, and nuts" patterns) were inversely associated with BF%. In contrast, patterns characterised by more 'discretionary' foods ("sweet and fat rich carbohydrate") and less diversity of core foods ("animal meat and fat") were positively associated with BF% for both Pacific and NZE women. Three enterotypes were identified by higher relative abundance of specific bacterial species: enterotype 1 was characterised by Pacific and NZE women (n=146) and the abundances of *Faecalibacterium prausnitzii* and *Eubacterium rectale*. Enterotype 2 (n=70) was characterised by Pacific women, *Bifidobacterium adolescentis*, *Bifidobacterium bifidum*, and *Lactobacillus ruminis*; and by higher BF%, visceral adipose tissue, and concentrations of fasting insulin. Enterotype 3 (n=70) was predominately found in older NZE women with lower deprivation, and characterised by *Akkermansia muciniphila, Ruminococcus bromii, Subdoligranulum species*, and *Methanobrevibacter smithii*. Adherence to the "colourful vegetables, plant protein, and dairy" dietary pattern was positively associated with enterotypes 1 and 3 and negatively with enterotype 2.

Conclusion

Consuming more core foods rich in dietary fibre was associated with enterotypes 1 and 3, including lower adiposity and metabolic disease risks. In contrast, consuming more discretionary foods was associated with enterotype 2, higher adiposity and metabolic disease risks. This PhD research highlights habitual diet-microbiota-host associations, which are similar for a population of women with different metabolic disease risk, body fat profiles, and deprivation levels. Whether the microbiota is a cause or consequence of metabolic health has yet to be elucidated. However, habitually consuming more core foods rich in dietary fibre is associated with microbiota composition, and lower metabolic disease risks.

Acknowledgements

He aha te mea nui o te ao. He tāngata, he tāngata, he tāngata.
What is the most important thing in the world? It is people, it is people, it is people.

This PhD research would not have been possible without the support and input of several people. Firstly, I would like to sincerely thank my supervisors Professor Bernhard Breier, Professor Gerald Tannock, and Associate Professor Rozanne Kruger for providing me the opportunity to be a part of the PROMISE team, as well as the wider PROMISE team for supporting me to bring this PhD to fruition.

To the Health Research Council of New Zealand thank you for funding the PROMISE study, and to the participants, thank you for your willingness to engage and participate in the study, it would not have been possible without you all.

I would like to extend my gratitude to my supervisor Professor Gerald Tannock, thank you for hosting me at your lab at the University of Otago, and for your patience and guidance while I explored the microbiota data. A special thank you to Dr Blair Lawley, for all the time you invested in teaching me how to use the bioinformatics software, and for your detailed answers to all my questions. Having the opportunity to work with you both, and your willingness to share your expertise has been rewarding, thank you.

I would also like to acknowledge the Max Rubner-Institut in Karlsruhe, Germany, for providing me with the opportunity to spend time at the institute, which enabled me to learn the NCI method from my dietary mentor Dr Benedikt Merz. Thank you Benedikt, for your support, guidance, knowledge, and friendship; for always answering my never-ending questions so warmly, and for all that you have taught me, thank you.

I am also grateful for the statistical expertise of Professor Jeroen Douwes and Dr Marine Corbin, thank you for welcoming me to the Centre of Public Health Research in Wellington, and sharing your knowledge. Your willingness to help, and passion for statistics is infectious, and thank you for guiding me to understand the story of the diet-microbiota data.

I could not have pursued this PhD without the support of the Riddet Institute. The team at the Riddet Institute work hard to create a supportive and growth orientated environment for PhD students. Thank you for all the opportunities and experiences offered to me; I am grateful for all the people I have met and the friends that I have made.

I would also like to thank the PROMISE girls – Jo Slater, Sophie Kindleysides and Niamh Brennan, the hours, weeks, and months we invested in this project will never be forgotten, thank you for walking alongside me on this journey. Your support, friendship and laughter uplifted me and made this experience all the better for that I got to meet you all.

Lastly, I would like to say a huge thank you to my whānau and friends, your unwavering support and belief has been second to none. Thank you for accepting my mahi related absences over the past years and welcoming me back as if it was just yesterday when we last saw each other. To those who I cannot share this achievement with, I dedicate this to your memory, and thank you for being a part of my journey to get here.

And finally, to my partner Shaun, thank you. You have sacrificed so much for me to pursue this PhD; no words can express how grateful I am. Thank you for your belief and standing beside me.

Thank you to all the people who contributed to my PhD journey.

Table of contents

List of Tables

Supplementary Tables

List of Figures

Supplementary Figures

Abbreviations

%EI	Percentage of energy intake
5DFR	5-day non-consecutive estimated food record
AF%	Android fat percentage (measured with DXA)
AHEI	Alternative Healthy Eating Index
AMDRs	Acceptable Macronutrient Distribution Ranges
AusBrands	Xyris brand name database based on Australian food composition database (2017)
AusFoods	Xyris brand name database based on Australian food composition database (2017)
AUSNUT	Australian food composition database (2011-13)
BF%	Total body fat percentage (measured with DXA)
BMI	Body mass index (kg/m^2)
BMR	Basal metabolic rate
CCK	Cholecystokinin
CHD	Chronic heart disease
CHO	Total carbohydrate
CNS	Central nervous system
CVD	Cardiovascular disease
DASH	Dietary Approaches to Stop Hypertension
DBP	Diastolic blood pressure
DPMP	Dietary Patterns Method Project
DXA	Dual-energy X-ray Absorptiometry
EDTA	Ethylene diamine tetraacetic acid
EI	Energy intake
F:B ratio	Firmicutes/Bacteroidetes ratio
FFQ	Food frequency questionnaire
FOODFiles 2016	New Zealand's food composition database
GBD	Global burden of disease study
GF%	Gynoid fat percentage (measured with DXA)
GLP-1	Glucagon-like peptide-1
HbA1c	Glycosylated haemoglobin
HDL-C	High density lipoprotein cholesterol
HEI	Healthy Eating Index
HFD	Healthy Food Diversity Index
High-BF%	Total body fat percentage over 35 %
HOMA-IR	Homeostasis model assessment index for insulin resistance
HPFS	The Health Professionals Follow-Up Study
HNU	Human Nutrition Unit
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
ISAK	International Society for the Advancement of Kinanthropometry
ISUF	Iowa State University Foods method
LDL-C	Low density lipoprotein cholesterol,
Low-BF%	Total body fat percentage under 35 %

MD	Mediterranean diet
MUFA	Monounsaturated fat
NCD	Non-communicable disease
NHS	The Nurse's Health Study
NHS2	The Nurse's Health Study II
NCI	National Cancer Institute
NPY	Neuropeptide Y
NZ	New Zealand
NZDep2013	New Zealand's socioeconomic deprivation index
NZE	New Zealand European
NZRD	New Zealand Registered Dietitian
NZ NNS	New Zealand National Nutrition Survey
NZWFFQ	New Zealand Women's Food Frequency Questionnaire
OECD	The Organization for Economic Cooperation and Development
OTU	Operational taxonomic unit
OR	Odds ratio
PAL	Physical activity level
PCA	Principal component analysis
PROMISE	The PRedictors linking Obesity and gut MIcrobiomE study
PUFA	Polyunsaturated fat
SBP	Systolic blood pressure
SCFA	Short chain fatty acids
SES	Socioeconomic status
SFA	Saturated fat
SOP	Standard operating procedures
SSB	Sugar sweetened beverages
STAMP	Statistical analysis of metagenomic profiles
T1	Tertile 1
T2	Tertile 2
T2DM	Type 2 diabetes mellitus
T3	Tertile 3
TGS	Triglycerides
TC	Total cholesterol
TC:HDL	Total cholesterol/ high density lipoprotein cholesterol
TMA	Trimethylamine
TMAO	Trimethylamine oxide
USDA	The United States Department of Agriculture
VAT	Visceral adipose tissue percentage (measured with DXA)
WCRF	The World Cancer Research Fund
WHO	World Health Organization

1. Chapter 1. Introduction

Worldwide, the prevalence of obesity has tripled since 1975 (1); and obesity and associated non communicable diseases (NCDs) are now considered global epidemics (1). Obesity is defined as an excess accumulation of adipose tissue which can impair health, and is classified as a body mass index (BMI) ≥30 kg/m^2 (1). In New Zealand (NZ) one in three people over the age of 15 years are obese (2); consequently, it ranks as the third highest among the Organisation for Economic Co-Operation and Development (OECD) countries (3). Disparities in the prevalence of obesity exist in NZ, related to ethnicity and deprivation. For example, 29 % of NZ Europeans (NZE), 48 % of indigenous Māori, and 67 % of Pacific people were classified as obese in the 2018/2019 annual NZ health survey (2). Similar to global trends, in NZ, living in an area with higher levels of deprivation is associated with an increased vulnerability to developing obesity, and women are more likely to be obese than men (2,4,5).

The causes of obesity are multifaceted and complex. However, in its simplest form, obesity arises as a disruption to a complex energy regulation system (6). Overnutrition and sedentary behaviours are considered principle factors driving this disruption (7), where energy consumed exceeds energy expended for a prolonged period of time, leading to the accumulation of excess adipose tissue. It is improbable that obesity is related to an excessive or inadequate intake of a single nutrient or food; however, the combination of nutrients and foods an individual usually consumes (i.e., one's habitual diet) is considered a modifiable risk factor for NCDs such as obesity and type 2 diabetes (T2DM) (8–10). In contrast to other risk factors, diet affects all people regardless of ethnicity, social economic status (SES), gender, or age (9). Habitual diets characterised by core or 'healthy' foods and nutrients (e.g., vegetables, fruit, wholegrains, omega-3, dietary fibre, calcium), are consistently associated with a reduced risk of poor health outcomes such as cardiovascular disease (CVD) and other NCDs (9–13). In contrast, overconsumption of non-essential 'discretionary' foods and nutrients (e.g., refined grains, processed meats, sugar, sodium, saturated fat) likely displaces the intake of core foods. Thus, suboptimal diets, specifically inadequate intake of essential and excessive intake of non-essential foods and nutrients, are a key driver in the development of NCDs, in particular T2DM and obesity (8,9,14).

The gut microbiota, the bacteria and their associated genomes residing in the large bowel, is of interest due to its metabolic functions which can influence host physiology (15). Animal

models have demonstrated the causal role of the microbiota in the pathogenesis of obesity (16,17). For example, germ free mice (mice without a microbiota) are protected from diet-induced obesity (18), and the obese phenotype (from either mice or humans) can be transferred to a lean host via a microbiota (faecal) transplant (16,19). Microbial signatures, such as higher relative abundance of the phyla *Firmicutes* to *Bacteroides* (20,21) or a state of microbiota 'dysbiosis' (variation or imbalance of microbiota (22)), have been proposed to be a characteristic of obesity in animal models. However, the translation of these observations to humans is less clear. Several attempts have been made to characterise the obese and lean microbiota in humans; however, the same trend of the two main phyla in relation to adiposity has not been consistently observed (23–26). Contradicting results have been proposed to be due to the heterogeneity across studies (e.g., methodology and populations investigated) and are further confounded through failure to account for diet (25–27). Despite this, reduced microbiota diversity (and diversity of bacterial genes) has been associated with obesity in both animal and human studies (22,28,29). Higher microbiota diversity has further been proposed to be a marker of health status in humans, because reduced diversity has been associated with inflammatory bowel disease (30), T2DM (31) and obesity (22,28,29) when compared to healthy subjects. Thus, the diversity of microbiota composition could be important for the health of the host, as increased diversity may translate to increased capacity to positively influence human physiology.

The composition of an adult's microbiota is considered relatively stable following dynamic changes during the early years of life (32). Throughout the lifespan, several factors can shape an individual's microbiota composition, such as: mode of birth delivery, infant feeding practices including habitual diet, antibiotic use, and geographical location (33,34). Consequently, the unique characteristics and inter-individual differences in microbiota composition observed between individuals hinders the ability to characterise a 'normal' or 'healthy' microbiota. Subsequently, there is increasing scepticism if a normal microbiota even exists (35). Despite these unresolved questions, the habitual diet is considered to exert the strongest selective pressure on the microbiota (34,36). The microbiota are driven by indigestible polysaccharides (cellulose, hemicellulose, and resistant starch) found in plant based foods (vegetables, fruit, and wholegrains), which are collectively known as dietary fibre (37). Plant polysaccharides are largely indigestible to the host and thus pass through to the

large colon where they are fermented by the resident microbiota (38). The microbiota are equipped with carbohydrate degrading enzymes required to metabolise the complex carbohydrate structures (39), and short chain fatty acids (SCFAs: acetate, propionate, and butyrate) are characteristic of major metabolites produced by carbohydrate fermentation (39,40). These SCFAs provide additional energy sources for the human host and, therefore, influence energy homeostasis by contributing to appetite regulation, blood glucose, and lipid control (15,39). The composition of an individual's microbiota determines the capacity for fermentation (15); however, the proportion of SCFAs produced is largely dependent on the availability of substrates. Therefore, habitual diet has the potential to influence the functions of the microbiota.

Habitually consuming a diet rich in plant-based products (e.g., unrefined and complex carbohydrates rich in dietary fibre) is associated with lower long-term weight gain and reduced risk of obesity, NCDs, and mortality (9–11,41–43). The beneficial associations attributed to habitually consuming such diets have yet to be elucidated; however, have been proposed to be due to the functions of the microbiota (39). For example, higher adherence to a Mediterranean dietary pattern (characterised by fruit, vegetables, legumes, nuts, olive oil, and fish with minimal intake of processed foods, meat, and dairy) is associated with a reduced risk of CVD (10,44) as well as higher fibre degrading microbiota and increased production of SCFAs (45). In addition, higher habitual dietary fibre intake is associated with increased microbiota diversity and lower long term weight gain (46). Therefore, it appears that dietary diversity (e.g., consuming a variety of core and less non-essential foods and nutrients) and microbiota diversity complement each other for better health outcomes. Thus, there is increasing interest in the prospect of manipulating the microbiota with dietary interventions to promote better health outcomes for the host. For example, control and/or reverse obesity via promoting appetite regulation, blood glucose and lipid normalisation, and weight loss. Animal models have demonstrated the potential benefit of such interventions (47,48); although more studies are required to determine if modulation of the microbiota translates to health benefits in humans.

However, before pursuing individualised nutrition to modulate the microbiota, there is currently no clear consensus about the association between diet, microbiota, and health

outcomes in humans. It is also possible that modulation of the microbiota may not translate to health benefits for the host (i.e., dysbiosis). Early diet intervention studies (e.g., delivering specific nutrients or foods) cited rapid changes in the microbiota (e.g., 24-hours (36)) while others had mixed results (49). Later studies emphasised that the habitual diet was the main driver of the microbiota composition (34,36,50–53). Arumugam et al. (54) proposed that robust clusters of taxa or 'enterotypes' can be recognised, and these were independent of age, BMI, cultural background, and gender. Habitual diet was later proposed to drive these enterotypes (36). For example, the *Bacteroides* enterotype was associated with a diet rich in animal protein and fat, the *Prevotella* enterotype was associated with a carbohydrate rich diet, while the third enterotype *Ruminococcus* was not distinguishable by dietary intake (36). Enterotypes are debated to be more of a continuum rather than discrete clusters (55,56), because enterotypes are characterised based on the relative abundance of how common one genus is to another. However, subsequent studies have since characterised two (36,50,57,58) or three (34,59,60) of these discrete enterotypes recognised by Arumugam et al., in different populations (some related and unrelated to host characteristics). Enterotypes have also been suggested to influence responsiveness to dietary interventions. For example, individuals with a *Prevotella* enterotype lost more weight on a higher fibre diet in comparison to participants with a *Bacteroides* enterotype (61,62). Yet, across other studies, the association of enterotypes and diet is less clear, primarily due to the quality of dietary results and, in some cases, the choice of dietary assessment. For example, two studies stratified the population by self-reported adherence to diet type (e.g., vegan, vegetarian or omnivore) and reported no association with diet and enterotypes (59,60). As discussed by Losasso et al. (59) such qualitative categories do not accurately capture the integral details of the composition of the diet, which could give rise to simplified conclusions. Therefore, broad, self-determined categories of dietary intake should be coupled with more thorough dietary assessment methods to accurately assess diet composition in relation to the microbiota.

It is widely appreciated that capturing accurate dietary data from free living humans is inherently complex; however, there are methods to ensure that self-reported dietary data are collected, analysed, and interpreted correctly to minimise measurement error (63–66). Self-reported dietary data provide rich insight into the complexities of what and how individuals choose to eat, provided the appropriate adjustments and interpretations of the data are made

(65). However, across most of the diet and microbiota studies, thorough diet methodology (e.g., adequate description of dietary assessment tool used, processing of dietary data, and handling of energy misreporters) is rarely reported. In addition, only a few studies adjust their analysis for factors such as age, gender, BMI, SES (50,51,57,67,68) and energy intake (45,46,50,69,70) all of which can effect dietary intake and microbiota composition (32–34,63,64). The difficulty in interpreting clear associations between diet and microbiota appears to be secondary to the numerous approaches employed to assess and interpret dietary data, including how the methodology is reported across studies.

Similar to dietary data, there are also considerations for the analyses and interpretation of the microbiota data. The composition of the faecal microbiota is assessed using DNA based methods, and the sequencing based techniques have rapidly developed over the last two decades (71,72). Earlier studies measured the microbiota using 16S ribosomal RNA genes sequences and later methods have employed whole genome shotgun sequencing, which has enabled greater analytical depth (72). However, whole genome sequencing is more costly so is not widely used. Most investigations into the composition of the faecal microbiota are by analysis of 16S rRNA sequences. Utilising the 16S rRNA sequencing is relatively accurate down to the family level; however, greater analytical depth is required for more accuracy to explore bacterial species abundances. Therefore, when interpreting results reporting the composition of the microbiota, consideration of the sequencing technique is required.

When exploring the reported associations between diet-microbiota and health outcomes, intervention studies do show causality although they are usually focused on the influence of a single nutrient or food, without consideration of the whole diet (which would also be providing substrate to the microbiota). In addition, the inter-individual differences in microbiota composition can influence responsiveness to dietary interventions (73–75). Therefore, intervention studies are not the most effective means to explore the association between habitual diet and the microbiota.

Large observational studies have the power to detect diet, microbiota, and health associations, but such studies tend to focus on adherence to a specific diet (e.g., the Mediterranean diet (45,76,77) or vegetarian vs. omnivore (59,60,69,78)), as well as being

conducted within specific populations such as elderly (57,67,77,79), or subsamples from population-based studies (22,36,67,80). Furthermore, some observation studies have relatively small sample sizes (e.g., n=31-98 (36,68,69,77,81)). Only a few studies have been conducted in larger samples of healthy, free-living populations (34,51,52). In these observation studies, dietary assessment is limited in that habitual diet is typically only assessed with a retrospective food frequency questionnaire (FFQ).

A FFQ requires participants to recall specific foods consumed over a given timeframe; hence rely on an individual's memory and portion size estimation of food intake. FFQs impose low participant burden and are more cost and time effective to administer, thus, are frequently used in large observational cross-sectional studies. However, FFQs are prone to overestimation of intake (63,64,82). Prospective food records (weighed or estimated) are considered the gold standard of dietary assessment (63,64), but they are infrequently used due to the increased participant and researcher burden. In addition, food records are often short in duration (2-5 days) so do not efficiently capture episodically consumed foods, and thus habitual diet.

Most studies exploring diet in relation to the human gut microbiota focus on individual nutrients (36,46,69,70), or specific foods (34,51,52,68). However, people do not typically consume nutrients or foods in isolation (83,84). Dietary pattern analysis enables the assessment of the broader diet by considering the combination of foods and nutrients consumed together, which can then be explored in relation to nutrition related health outcomes (84). Dietary patterns can be derived either statistically (*a posteriori* approach) or by scoring the quality of a diet based on adherence to an established criterion (*a priori* approach) (85). However, there are limited studies which explore population specific dietary patterns (statistically derived) in relation to the microbiota. In a recent study, Shikany et al. (79) identified four dietary patterns in a population of older American men (n=517). They found that higher adherence to the 'Western' pattern was positively associated with BMI, and conversely the 'prudent' pattern was negatively associated with BMI; however, neither pattern was associated with microbiota (alpha) diversity. Alpha diversity is a measure of microbiota diversity (number and distribution of species present) within an individual. In another recent study with older German adults (n=225), Oluwagbemigun et al. (67) identified

and characterised five dietary patterns none of which were associated with the microbiota composition of the participants. Furthermore, three dietary patterns were identified in a small sample (n=45) of healthy overweight and obese French adults (68). The healthier pattern (characterised by fruit, yoghurt, soup, and lower intake of sugar sweetened foods and beverages) was associated with higher microbiota gene richness and diversity when compared to the two less healthy patterns (68). Overall there is a paucity of studies exploring dietary patterns in relation to the microbiota of healthy, free-living younger populations. In addition, the quality and detail of the reported dietary assessment tends to decline in larger studies, likely secondary to methodological limitations considering the scale of such studies. Thus, despite the growing interest in exploring the habitual diet and microbiota in humans, the current evidence is limited, extremely difficult to interpret or translate into practical insights. Diet is a key modifiable risk factor for NCDs such as obesity, and evidence from animal models have established a causal role of microbiota in diet-induced obesity. Therefore, to advance understanding of the role of the habitual diet, and the link between microbiota and health, further research is required to establish whether diet is associated with the microbiota and health outcomes in humans too.

Employing multiple methods of dietary assessment improves the validity and accuracy of dietary data. Multiple food records or 24-hour recalls are considered more accurate than a FFQ (which is prone to higher rates of misreporting) (82). Statistical methods have been derived using multiple dietary assessments (e.g., 24-hour recalls) to accurately and cost effectively capture habitual dietary data on a large-scale. Such methods are frequently used in large population studies to assess adherence to dietary guidelines and health outcomes (86). However, there are numerous statistical challenges which must be addressed to accurately assess habitual dietary intake: day-to-day variation of intake, episodically consumed foods, allowing for high intake of foods (positively skewed), and the inclusion of person specific covariates which can influence dietary intake (e.g., age, sex, BMI) (87). Statistical methods have been successively refined to overcome such challenges and limitations identified in earlier methods (87,88). Tooze et al. (88) developed the National Cancer Institute (NCI) method to effectively estimate the probability of consumption of a nutrient, food, or food group, and the respective amount consumed, whilst considering covariates which could either effect the amount consumed or probability of consumption (88–

90). The modelling process is usually based on one or more 24-hour recalls; and Kipnis et al. (86) suggested that the inclusion of a FFQ can improve the precision in the estimation of episodically consumed foods.

Therefore, statistical modelling of habitual dietary data based on multiple methods of dietary assessment, needs to be explored in the context of diet and microbiota research. Dietary pattern analyses would overcome the limitations of focusing on a single nutrient or food by enabling the assessment of the interrelatedness of foods and nutrients consumed together and thus the whole diet in relation to the microbiota. Higher adiposity in women of childbearing age increases the risk of obesity for the next generation, highlighting the importance of exploring diet-microbiota-health associations in this population. Exploring the association of habitual dietary intake in a population of free-living healthy women whom have different metabolic disease risks would further advance the understanding of diet, microbiota, and health associations; and open new avenues for therapeutic targets (i.e., personalised nutrition based on microbiota composition). By focusing on two different body weight profiles (normal and obese BMI) it will cover the greatest difference in body weight and may increase the likelihood of identifying microbiota health associations. In NZ, Pacific women have a high risk of obesity (70 % obesity) and New Zealand European (NZE) women have a moderate risk of obesity (30 % obesity); therefore, comparisons between Pacific and NZE women, who differ in terms of ethnic-cultural, physical, and socio-economic characteristics, allows the effect modification of these factors to be assessed. Furthermore, assessing the composition of the microbiota with metagenomic shotgun sequencing would enable a more in-depth assessment of the microbiota. To address the deficit in robust dietary assessment in relation to microbiota in a large healthy free-living population, this research project was designed to thoroughly characterise habitual dietary intake using robust and reliable methodological approaches. The purpose of this PhD research was to improve understanding of the association between habitual diet and the composition of the microbiota, in relation to body composition and metabolic health markers, in a large healthy free-living population with different metabolic disease risk and body fat profiles (lean and obese).

1.1. Study aims and objectives

The aim of this PhD research, which utilised the PROMISE cross-sectional study, was to characterise the habitual dietary intake (using the National Cancer Institute method) of two populations, Pacific and NZE women, aged 18-45 years, with different metabolic disease risk (Pacific women have a high risk of obesity (70 % obesity) and NZE women have a moderate risk of obesity (30 % obesity)), and body fat profiles (lean and obese), and to further explore the associations with body fat content, metabolic health markers, and gut microbiota composition (assessed with metagenomic shotgun sequencing).

The objectives of this PhD research were to:

- Explore habitual dietary intake in a population of healthy NZ women (Pacific and NZE), aged 18-45 years, with different metabolic disease risks and body fat profiles (lean and obese), and to investigate the relationship between habitual macronutrient intake and body fat content.
- Characterise *a posteriori* dietary patterns, derived from multiple days of dietary assessment, and to explore the association with body fat content and metabolic health markers in a population of NZ women (Pacific and NZE), who have different metabolic disease risk and different body fat profiles (lean and obese).
- Explore the characteristics of microbiota composition in relation to habitual diet (dietary patterns, foods, and nutrients), body fat content, and metabolic health markers, in a population of healthy NZ women (Pacific and NZE) with different metabolic disease risk and body fat profiles (lean and obese).

1.1.1. Structure of the thesis

This PhD thesis begins with a review of the literature (chapter two) which explores the public health issue of obesity and the association with dietary intake, along with the important considerations for dietary assessment. The review then explores the gut microbiota and its association with health outcomes in relation to diet, before concluding with a review of the analyses of habitual diet in relation to the gut microbiota composition.

This PhD thesis is written in the style of a PhD thesis by publication of which the results chapters four-six, are written in manuscript form for submission. Each chapter individually

addresses the objectives mentioned above and because they are written in the form of manuscripts there may be repetition throughout the thesis. Chapter four explores the relationship between habitual macronutrient intake, body fat content, and metabolic health markers in healthy premenopausal Pacific and NZE women. Chapter five explores the relationship between habitual dietary patterns, body fat content, and metabolic health markers of Pacific and NZE women. Chapter six explores the relationship between habitual dietary intake (dietary patterns, foods, and nutrients) of Pacific and NZE women, in relation to the composition of their gut microbiota, body fat content, and metabolic health markers. The thesis concludes with a discussion that brings together the main results of this PhD research, highlighting their significance and relevance, as well as methodological strengths and limitations of the research. Final conclusions and consideration of future research bring the thesis to a close.

1.2. References

1. World Health Organization. Obesity and overweight [Internet]. 2018 [cited 2019 Apr 12].
2. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].
3. Organization for Economic Co-operation and Development (OECD). OECD Obesity update 2017. Paris, France; 2017.
4. Zhang Q, Wang Y. Socioeconomic inequality of obesity in the United States: Do gender, age, and ethnicity matter? Soc Sci Med. 2004 Mar;58(6):1171–80. doi:10.1016/S0277-9536(03)00288-0
5. The National Advisory Committee on Health and Disability. The social, cultural and economic determinants of health in New Zealand: Action to improve health a report from the National Advisory Committee on Health and Disability. Wellington; 1998.
6. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, et al. Obesity pathogenesis: An endocrine society scientific statement. Endocr Rev. 2017;38(4):267–96. doi:10.1210/er.2017-00111
7. Swinburn B, Sacks G, Vandevijvere S, Kumanyika S, Lobstein T, Neal B, et al. INFORMAS (International Network for Food and Obesity/non-communicable diseases Research, Monitoring and Action Support): Overview and key principles. Obes Rev. 2013;14(October):1–12. doi:10.1111/obr.12087
8. Ludwig DS, Hu FB, Tappy L, Brand-Miller J. Dietary carbohydrates: Role of quality and quantity in chronic disease. BMJ. 2018;361:k2340:1–6. doi:10.1136/bmj.k2340
9. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8
10. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.
11. World Cancer Research Fund/American Institute for Cancer Research. Diet, nutrition, physical activity and cancer: A global perspective [Internet]. Continuous update project expert report 2018. 2018.
12. Yip CSC, Chan W, Fielding R. The associations of fruit and vegetable intakes with burden of diseases: A systematic review of meta-analyses. J Acad Nutr Diet [Internet]. 2019;119(3):464–81. doi:10.1016/j.jand.2018.11.007
13. Schlesinger S, Neuenschwander M, Schwedhelm C, Hoffmann G, Bechthold A, Boeing H, et al. Food groups and risk of overweight, obesity, and weight gain: A systematic review and dose-response meta-analysis of prospective studies. Adv Nutr. 2019;10(2):205–18. doi:10.1093/advances/nmy092
14. Dahl WJ, Stewart ML. Position of the Academy of Nutrition and Dietetics: Health

implications of dietary fiber. J Acad Nutr Diet. 2015;115(11):1861–70. doi:10.1016/j.jand.2015.09.003

15. Cani PD, Van Hul M, Lefort C, Depommier C, Rastelli M, Everard A. Microbial regulation of organismal energy homeostasis. Nat Metab [Internet]. 2019;1(1):34–46. doi:10.1038/s42255-018-0017-4

16. 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

17. Bäckhed F, Ding H, Wang T, Hooper L V., Gou YK, Nagy A, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A. 2004;101(44):15718–23. doi:10.1073/pnas.0407076101

18. Bäckhed F, Manchester JK, Semenkovich CF, Gordan JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. Proc Natl Acad Sci. 2007;104(3):979–84. doi:10.1007/BF00540675

19. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science (80-) [Internet]. 2013 Sep 6;341(6150):1241214–1241214. doi:10.1126/science.1241214

20. Ley RE, Turnbaugh PJ, Klein S, Gordan JI, Gordon JI. Human gut microbes associated with obesity. Nature [Internet]. 2006;444(7122):1022–3. doi:10.1038/nature4441021a

21. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci. 2005;102(31):11070–5. doi:10.1073/pnas.0504978102

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

23. Schwiertz A, Taras D, Schäfer K, Beijer S, Bos NA, Donus C, et al. Microbiota and SCFA in lean and overweight healthy subjects. Obesity [Internet]. 2010;18(1):190–5. doi:10.1038/oby.2009.167

24. Duncan SH, Lobley GE, Holtrop G, Ince J, Johnstone AM, Louis P, et al. Human colonic microbiota associated with diet, obesity and weight loss. Int J Obes. 2008;32(11):1720–4. doi:10.1038/ijo.2008.155

25. Walters WA, Xu Z, Knight R. Meta-analyses of human gut microbes associated with obesity and IBD. FEBS Lett. 2014;588(22):4223–33. doi:10.1016/j.febslet.2014.09.039

26. Finucane MM, Sharpton TJ, Laurent TJ, Pollard KS. A taxonomic signature of obesity in the microbiome? Getting to the guts of the matter. PLoS One. 2014;9(1):1–5. doi:10.1371/journal.pone.0084689

27. Sze MA, Schloss PD. Looking for a signal in the noise: Revisiting obesity and the microbiome. MBio [Internet]. 2016 Sep 7;7(4):1–10. doi:10.1128/mBio.01018-16

28. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. Nature [Internet]. 2013;500(7464):541–6. doi:10.1038/nature12506

29. Beaumont M, Goodrich JK, Jackson MA, Yet I, Davenport ER, Vieira-Silva S, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol [Internet]. 2016;17(1). doi:10.1186/s13059-016-1052-7

30. Manichanh C. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut [Internet]. 2006;55(2):205–11. doi:10.1136/gut.2005.073817

31. Lambeth SM, Carson T, Lowe J, Ramaraj T, Leff JW, Luo L, et al. Composition diversity and abundance of gut microbiome in prediabetes and type 2 diabetes. J Diabetes Obes. 2015 Dec 26;2(2):108–14. doi:10.15436/2376-0949.15.031

32. O'Toole PW, Claesson MJ. Gut microbiota: Changes throughout the lifespan from infancy to elderly. Int Dairy J [Internet]. 2010;20(4):281–91. doi:10.1016/j.idairyj.2009.11.010

33. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220–30. doi:10.1038/nature11550

34. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science (80-) [Internet]. 2016 Apr 29;352(6285):560–4. doi:10.1126/science.aad3503

35. McBurney MI, Davis C, Fraser CM, Schneeman BO, Huttenhower C, Verbeke K, et al. Establishing what constitutes a healthy human gut microbiome: State of the science, regulatory considerations, and future directions. J Nutr [Internet]. 2019;1–14. doi:10.1093/jn/nxz154

36. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science (80-) [Internet]. 2011 Oct 7;334(6052):105–8. doi:10.1126/science.1208344

37. Mann J, Cummings JH, Englyst HN, Key T, Liu S, Riccardi G, et al. FAO/WHO Scientific update on carbohydrates in human nutrition: Conclusions. Eur J Clin Nutr. 2007;61(S1):S132–7. doi:10.1038/sj.ejcn.1602943

38. Cummings JH, Stephen AM. The role of the dietary fibre in the human colon. Can Med Assoc J. 1980;123(11):1109–14.

39. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. Cell. 2016;165(6):1332–45. doi:10.1016/j.cell.2016.05.041

40. Cummings JH, Macfarlane GT. The control and consequences of bacterial fermentation in the human colon. J Appl Bacteriol [Internet]. 1991 Jun;70(6):443–59. doi:10.1111/j.1365-2672.1991.tb02739.x

41. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

42. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality

and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9

43. Wang DD, Li Y, Afshin A, Springmann M, Mozaffarian D, Stampfer MJ, et al. Global improvement in dietary quality could lead to substantial reduction in premature death. J Nutr [Internet]. 2019;149(6):1065–74. doi:10.1093/jn/nxz010

44. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts. N Engl J Med [Internet]. 2018;378(25):e34. doi:10.1056/NEJMoa1800389

45. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut. 2016;65(11):1812–21. doi:10.1136/gutjnl-2015-309957

46. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int J Obes. 2017;41(7):1099–105. doi:10.1038/ijo.2017.66

47. Griffin NW, Ahern PP, Cheng J, Heath AC, Ilkayeva O, Newgard CB, et al. Prior dietary practices and connections to a human gut microbial metacommunity alter responses to diet interventions. Cell Host Microbe [Internet]. 2017 Jan;21(1):84–96. doi:10.1016/j.chom.2016.12.006

48. Ravussin Y, Koren O, Spor A, LeDuc C, Gutman R, Stombaugh J, et al. Responses of gut microbiota to diet composition and weight loss in lean and obese mice. Obesity [Internet]. 2012 Apr 19;20(4):738–47. doi:10.1038/oby.2011.111

49. So D, Whelan K, Rossi M, Morrison M, Holtmann G, Kelly JT, et al. Dietary fiber intervention on gut microbiota composition in healthy adults: a systematic review and meta-analysis. Am J Clin Nutr. 2018;965–83. doi:10.1093/ajcn/nqy041

50. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, et al. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. Front Genet [Internet]. 2019;10:454. doi:10.3389/fgene.2019.00454

51. Partula V, Mondot S, Torres MJ, Kesse-Guyot E, Deschasaux M, Assmann K, et al. Associations between usual diet and gut microbiota composition: Results from the Milieu Intérieur cross-sectional study. Am J Clin Nutr. 2019;109(5):1472–83. doi:10.1093/ajcn/nqz029

52. 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 (80-) [Internet]. 2016 Apr 29;352(6285):565–9. doi:10.1126/science.aad3369

53. Johnson AJ, Vangay P, Al-Ghalith GA, Hillmann BM, Ward TL, Shields-Cutler RR, et al. Daily sampling reveals personalized diet-microbiome associations in humans. Cell Host Microbe. 2019;25(6):789-802.e5. doi:10.1016/j.chom.2019.05.005

54. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature [Internet]. 2011;473(7346):174–80.

doi:10.1038/nature09944

55. Knights D, Ward TL, McKinlay CE, Miller H, Gonzalez A, McDonald D, et al. Rethinking enterotypes. Cell Host Microbe [Internet]. 2014;16(4):433–7. doi:10.1016/j.chom.2014.09.013

56. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol [Internet]. 2018 Jan 18;3(1):8–16. doi:10.1038/s41564-017-0072-8

57. Claesson M, Jeffery IB, Conde S, Power S, O'Connor E, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;(488):178–84. doi:10.1038/nature11319

58. Dhakan DB, Maji A, Sharma AK, Saxena R, Pulikkan J, Grace T, et al. The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. Gigascience. 2019;8(3):1–20. doi:10.1093/gigascience/giz004

59. Losasso C, Eckert EM, Mastrorilli E, Villiger J, Barrucci F, Pernthaler J, et al. Assessing the influence of vegan, vegetarian and omnivore oriented westernized dietary styles on human gut microbiota: A cross sectional study. Front Microbiol. 2018;9(March):1–12. doi:10.3389/fmicb.2018.00317

60. de Moraes ACF, Fernandes GR, da Silva IT, Almeida-Pititto B, Gomes EP, Pereira A da C, et al. Enterotype may drive the dietary-associated cardiometabolic risk factors. Front Cell Infect Microbiol. 2017;7(FEB):1–9. doi:10.3389/fcimb.2017.00047

61. Hjorth MF, Roager HM, Larsen TM, Poulsen SK, Licht TR, Bahl MI, et al. Pre-treatment microbial Prevotella-to-Bacteroides ratio, determines body fat loss success during a 6-month randomized controlled diet intervention. Int J Obes [Internet]. 2018;42:580–3. doi:10.1038/ijo.2017.220

62. Hjorth MF, Blædel T, Bendtsen LQ, Lorenzen JK, Holm JB, Kiilerich P, et al. Prevotella-to-Bacteroides ratio predicts body weight and fat loss success on 24-week diets varying in macronutrient composition and dietary fi ber: Results from a post-hoc analysis. Int J Obes. 2019;43:149–57. doi:10.1038/s41366-018-0093-2

63. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.

64. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001

65. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45. doi:10.3945/jn.115.219634

66. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, et al. DIET@NET: Best practice guidelines for dietary assessment in health research. BMC Med [Internet]. 2017 Dec 15;15(1):202. doi:10.1186/s12916-017-0962-x

67. Oluwagbemigun K, Foerster J, Watkins C, Fouhy F, Stanton C, Bergmann MM, et al. Dietary patterns are associated with serum metabolite patterns and their association is

influenced by gut bacteria among older german adults. J Nutr [Internet]. 2019 Sep 2;1–10. doi:10.1093/jn/nxz194

68. Kong LC, Holmes BA, Cotillard A, Habi-Rachedi F, Brazeilles R, Gougis S, et al. Dietary patterns differently associate with inflammation and gut microbiota in overweight and obese subjects. PLoS One. 2014;9(10). doi:10.1371/journal.pone.0109434

69. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. BMJ. 2016;65(1):63–72. doi:10.1136/gutjnl-2014-308209

70. Lin D, Peters BA, Friedlander C, Freiman HJ, Goedert JJ, Sinha R, et al. Association of dietary fibre intake and gut microbiota in adults. Br J Nutr [Internet]. 2018;120(9):1014–22. doi:10.1017/s0007114518002465

71. Sekirov I, Russell SL, Antunes LCM, Finlay BB. Gut microbiota in health and disease. Physiol Rev. 2010;90:859–904. doi:10.1152/physrev.00045.2009.

72. Tannock GW. Understanding the gut microbiota. Hoboken (NJ): John Wiley & Sons; 2017. 1–174 p.

73. Kovatcheva-Datchary P, Nilsson A, Akrami R, Lee YS, De Vadder F, Arora T, et al. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. Cell Metab. 2015;22(6):971–82. doi:10.1016/j.cmet.2015.10.001

74. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. Cell. 2015;163:1079–94. doi:10.1016/j.cell.2015.11.001

75. Korem T, Zeevi D, Zmora N, Weissbrod O, Bar N, Lotan-Pompan M, et al. Bread affects clinical parameters and induces gut microbiome-associated personal glycemic responses. Cell Metab [Internet]. 2017;25:1243–53. doi:10.1016/j.cmet.2017.05.002

76. Mitsou EK, Kakali A, Antonopoulou S, Mountzouris KC, Yannakoulia M, Panagiotakos DB, et al. Adherence to the Mediterranean diet is associated with the gut microbiota pattern and gastrointestinal characteristics in an adult population. Br J Nutr [Internet]. 2017;117(12):1645–55. doi:10.1017/s0007114517001593

77. Gutiérrez-Díaz I, Fernández-Navarro T, Salazar N, Bartolomé B, Moreno-Arribas MV, de Andres-Galiana EJ, et al. Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish cohort of middle-age and older people. J Agric Food Chem [Internet]. 2017 Jan 25;65(3):586–95. doi:10.1021/acs.jafc.6b04408

78. Liang C, Tseng H-C, Chen H-M, Wang W-C, Chiu C-M, Chang J-Y, et al. Diversity and enterotype in gut bacterial community of adults in Taiwan. BMC Genomics [Internet]. 2017;18(S1). doi:10.1186/s12864-016-3261-6

79. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. Am J Clin Nutr [Internet]. 2019 Oct 1;110(4):1003–14. doi:10.1093/ajcn/nqz174

80. Bowyer RCE, Jackson MA, Pallister T, Skinner J, Spector TD, Welch AA, et al. Use of dietary indices to control for diet in human gut microbiota studies. Microbiome [Internet]. 2018 Dec 25;6(1):77. doi:10.1186/s40168-018-0455-y

81. Koo SH, Chu CW, Khoo JJC, Cheong M, Soon GH, Ho EXP, et al. A pilot study to examine the association between human gut microbiota and the host's central obesity. JGH Open [Internet]. 2019 Dec 16;3(6):480–7. doi:10.1002/jgh3.12184

82. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

83. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr. 2016;7(3):445–54. doi:10.3945/an.115.011718

84. Hu FB. Dietary pattern analysis: A new direction in nutritional epidemiology. Curr Opin Lipidol [Internet]. 2002 Feb;13(1):3–9. doi:10.1097/00041433-200202000-00002

85. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. Nutr Rev. 2004;62(5):177–203. doi:10.1301/nr.2004.may.177

86. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009;65(4):1003–10. doi:10.1111/j.1541-0420.2009.01223.x

87. Dodd KW, Guenther PM, Freedman LS, Subar AMYF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. J Am Diet Assoc. 2006;106(10):1640–50. doi:10.1016/j.jada.2006.07.011

88. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

89. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

90. Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. Stat Med [Internet]. 2010;29(27):2857–68. doi:10.1002/sim.4063

2. Chapter 2. Literature review

2.1. Introduction

Obesity is a global problem of epidemic proportions and has serious public health implications. Diet is considered one of the key drivers of the obesity epidemic, and a modifiable risk factor for NCDs such as obesity. Based on the results from animal studies, the gut microbiota (the community of microorganisms in the large intestine) may play a role in the pathogenesis of obesity. Therefore, the gut microbiota is of interest due to its metabolic functions which can influence human physiology. Additionally, what an individual usually eats (i.e., one's habitual diet) has the potential to influence the composition of the microbiota, which, in turn, could influence the functions of the microbiota and indirectly disease risk. However, the association of habitual diet and the human gut microbiota remains unclear, notably within the context of health outcomes.

This literature review opens with a discussion of obesity in the context of public health and explores the causes in relation to dietary intake. Understanding the determinants of obesity may help explain the observed disparities in the prevalence of obesity between ethnic groups. The cost and causes of obesity will be explored within an international and a NZ setting. Following on from this, discussion moves onto exploring the composition of diet, dietary assessment methods, and considerations for exploring diet and health related outcomes in the context of obesity. Finally, the review will transition to discussing the gut microbiota in relation to health outcomes and will conclude by reviewing previous work exploring the assessment of habitual diet and the association with the gut microbiota in healthy adults.

2.2. Obesity

2.2.1. Obesity, a major public health issue

Globally, the prevalence of obesity has tripled since 1975 (1). Worldwide the industrialisation and urbanisation of environments has changed the way people live and interact with their physical and social environments. Subsequently, the incidence of obesity has prolifically increased in developed and developing countries, such that we are now facing an era of the double burden of disease (attributable to under and overnutrition) (1). In NZ, one in three people over the age of 15 years are obese (2), and NZ is ranked the third highest among the

Organization for Economic Co-Operation and Development (OECD) countries, with Mexico and the United States of America taking second and first place respectively (3). Of concern the 2017 OECD obesity update projected that the global prevalence of obesity would increase further by 2030. Higher levels of deprivation and lower socioeconomic status (SES) are associated with increased vulnerability to developing obesity, notably within minority groups and women (4,5). However, the burden of obesity is shared across all social classes (6) demonstrated by SES inequality in obesity prevalence, which varies across ethnicity, sex, and age (4). Obesity and associated NCDs have now outpaced infectious diseases as the major global health concern (1). The risk of developing NCDs increases almost parallel to increases in BMI (7). The growing burden of these diseases, and the direct and indirect costs associated with excess body weight, is felt worldwide (8).

2.2.2. Prevalence and cost of obesity

The rising prevalence and incidence of obesity is a major cause for concern from the direct impact on an individual's physical and mental wellbeing (including increased risk of premature mortality), and cumulatively, the indirect economic and societal costs. In 2006 Lal et al. (9) estimated the combined cost in NZ, attributable to overweight and obesity (for the loss of productivity and health care expenditure), to be between NZ$722-849 million (9). Since 2006, the prevalence of adult obesity in NZ has increased from 26.5 % to 30.9 % in 2018/19 (2). In parallel, the societal and economic costs would have likely increased; however, there are no recent data to support this suggestion.

2.2.3. Health inequity of obesity

Socioeconomic, environmental, and cultural factors are the main determinants of health. Similar to global trends, significant health disparities exist in NZ, notably between indigenous and non-indigenous people (10). Health inequalities have a significant cost to both individuals and economies. Although multifaceted and complex, there is global appreciation that these inequalities arise due to the interaction between socioeconomic and cultural factors stemming from historical influences such as colonisation (5,10–13). Comparable to global trends, NZ has a SES gradient for mortality, where low income groups are at a higher risk of

premature death at every age in comparison to high income groups (14). In NZ, Pacific people are disproportionally represented in lower SES areas (5,14); and, driven by socioeconomic factors, indigenous Māori and Pacific people are disproportionally affected by health inequalities in comparison to Europeans (5). Worldwide, NCDs, such as CVD and T2DM, are leading causes of preventable deaths associated with higher BMIs (7). Within NZ, CVD accounts for 40 % of deaths annually, and Māori have higher rates of CVD and mortality than non-Māori (13). Māori are three times more likely to have T2DM than non-Māori, and like Pacific people, are more likely to develop T2DM earlier and experience complications (15,16). In NZ, Pacific people have higher rates of T2DM than Māori (16), and T2DM and associated complications account for the greatest difference in mortality rates between Pacific and non-Pacific people (14,17).

The influence of socioeconomic determinants of obesity (and higher BMIs) is reflected in the prevalence of obesity within NZ. For example, individuals living within deprived areas are 1.62 times (adjusted for age, sex, and ethnic group) more likely to be obese than those who live in the least deprived areas (2). Certain ethnic groups are disproportionately affected by a higher prevalence of obesity. Across the main ethnic groups within NZ, Pacific people were 2.46 times (adjusted for age and sex) more likely to be obese in comparison to non-Pacific people. Māori were 1.76 times more likely, and Asian were less than half as likely, to be obese compared to the rest of NZ population (2). NZ women are also more likely than men to be obese. These findings raise questions as to whether the drivers of obesity are the same for different ethnicities in NZ. What is apparent, is that NZ's rate of obesity will directly cost the country's economy.

2.2.4. Aetiology and biology of obesity

Obesity manifests as a complex interaction between metabolic, genetic, environmental, behavioural, and cultural factors, which interplay across the lifespan leading to the excess accumulation of fat mass (18). Obesity is defined by the World Health Organization (WHO) as an excess of adipose tissue which has the potential to impair health and is classified by a BMI ≥30 kg/m^2 (1). The American College of Endocrinology described obesity as an "adiposity-based chronic disease" (19), to foster an appreciation for how adipose tissue, as a

metabolically active organ, secretes hormones that influence energy intake and many metabolic pathways (18). Thus, the amount and distribution of adipose tissue influences metabolic function (18–21); therefore, the health risks associated with having a higher BMI and/or increased central adiposity are extensively documented, notably for increasing the risk of developing NCDs (7,18,21–23).

BMI is a weight to height ratio utilised to indicate relative body size; however, individuals with the same BMI can have different metabolic disease risks (20,24). Therefore, assessment of an individual's BMI needs to be interpreted with caution and not used in isolation as BMI does not accurately reflect the distribution of the adipose tissue, differentiate fat from fat free mass, or account for ethnic differences (18,25). In comparison to BMI, assessment of total body fat percentage (BF%) could be considered a more objective measure of body fat mass, as it is differentiated from fat free mass (20,22,26). Perhaps more importantly, is the assessment of the distribution of excess fat; for example, higher central obesity (android and visceral fat mass) which surrounds vital organs is associated with increased metabolic disease risks (21,23). Thus, utilising a gold standard such as a Dual-energy X-ray Absorptiometry (DXA) to assess BF% and distribution enables a more accurate assessment of obesity and the evaluation of metabolic disease risk (27). However, the cost and practicalities of utilising DXA is often a barrier for large epidemiological studies. Therefore, BMI, including waist circumference (and waist to hip ratio) as a measure of central adiposity, is often utilised in large studies as it is a cost-effective anthropometric measure.

Obesity arises as a consequence of the disruption to a complex energy regulation system (18). Over-nutrition and sedentary behaviours stemming from the industrialisation of environments are principle factors driving this disruption (28,29). Conventionally, obesity was purely considered a state of energy imbalance, if energy consumed exceeded energy expenditure for a prolonged period of time, the outcome was the accumulation of excess adipose tissue. The pathogenesis of obesity is, however, far more complex. Neural pathways and physiological networks of homeostatic and non-homeostatic controls influence the regulation of appetite and energy metabolism. Hunger is the physiological response to a need for food triggered by chemical messages from the hypothalamus and the gastrointestinal system, whereas appetite is the psychological desire to eat, sometimes triggered by hunger

(30). The Satiety Cascade eloquently describes the process of how appetite is regulated by a close interplay between the digestive tract, adipose tissue, and the brain (30). Collectively hunger and appetite interact to determine when, what, and how much is consumed, with the potential to disrupt energy balance. Hunger is stimulated by an empty stomach, changes in circulating hormone levels (increases in glucagon and ghrelin, and decreases in insulin), and decreases in blood glucose levels, which collectively trigger the physiological need for fuel (30). Appetite can be stimulated by a myriad of independent factors, such as the perceived reward of palatable foods driven by the sensory desire to eat, learned behaviours, emotional cues, or social, cognitive, and external factors (30–32). After ingesting food, the inherent characteristics of food, gastric distension, and variations in the circulating hormones (increased insulin, leptin, cholecystokinin (CCK), and decreased glucagon etc.) signal a state of satiety to the brain, encouraging the cessation of eating. This highlights the complexity of eating; for instance, an individual could easily override physiological cues as they do not have to be hungry to eat.

Insulin is the main endocrine regulator of adipose tissue mass and distribution. For example, following the ingestion of carbohydrates, plasma glucose levels would rise and insulin levels would mirror glucose. Rising insulin levels would stimulate glucose oxidation and glycogen storage in the target tissues (skeletal muscle, liver, and adipose tissue) (33). The regulation and secretion of insulin and glucagon in response to eating is complex. Insulin and glucagon exert opposing actions on nutrient metabolism and are secreted in a reciprocal fashion. There are multiple regulatory systems to ensure a balance of these two hormones to maintain plasma and glucose levels within a tight range (33). However, the accumulation of excess adipose tissue exacerbates weight gain, due in part to leptin and insulin resistance, which subsequently impedes appetite regulation (18,33,34). Insulin resistance is defined as an inadequate response of target tissues to the physiological effects of circulating insulin. There can be a genetic component increasing an individual's predisposition to develop insulin resistance. However, in 'Western' cultures, lifestyle factors, such as sedentary lifestyles and obesity, are considered key drivers of insulin resistance; and insulin resistance is considered a major risk factor for the development of T2DM and components of the metabolic syndrome.

The hormone leptin, produced by and in proportion to adipose tissue stores, has a causal influence in obesity development. Leptin is a hormone that affects behaviour. The discovery of leptin dispelled previously held beliefs that obesity arises due to lack of willpower, and cemented obesity as an endocrine disorder (18). Leptin acts on the hypothalamus in the central nervous system (CNS) to inhibit neuropeptide Y (NPY), a neurotransmitter that, stimulates eating. Leptin, therefore, signals to the brain that enough energy has been consumed. Hence, it has been coined the 'satiety hormone' (18,34). In contrast, the 'hunger hormone' ghrelin has the opposing effect. Fasting and low calorie diets stimulate the release of ghrelin from the endocrine cells in the stomach which, in turn, stimulates NPY, triggering the physiological feeling of hunger and, if food is available, food intake (34).

Overweight and obese individuals have higher circulating levels of leptin due to their higher fat mass; however, the higher circulating levels do not result in an effective feedback system (34). The appetite signalling effect of leptin is weaker in the obese state, which is thought to fuel overconsumption (18,30,34). Leptin resistance can occur peripherally and centrally, failing to reduce insulin secretion or feedback to the hypothalamus, respectively. Leptin resistance, in the obese state, gives rise to a false sense of leptin deficiency, where the satiety signalling effect and the feedback signals to the hypothalamus about adequate energy stores are impaired. Additionally, adipose tissue lipolysis is difficult to stimulate in a fasting state, because leptin inhibits lipolysis when insulin levels are low (see review (35)). Thereby, perpetuating the obese state as the surplus stores cannot be broken down for fuel, and muscle mass is favoured in this state. Furthermore, excess adipose tissue stores promote chronic low-grade inflammation and increased metabolic risks (e.g., developing insulin resistance and T2DM) (18,19).

Obesity is multifaceted in its origins, and the obese state perpetuates ongoing metabolic and endocrine issues fuelling the obese state in an obesogenic environment (an environment which promotes weight gain) (18,30). As mentioned above, there are multiple avenues where appetite, or the conscious (and/or subconscious) decision to eat, can override the physiological need for food and potential satiety cues; for example, the perceived pleasure of eating a tempting dessert after a satisfying meal. In turn, the obese state can disturb appetite regulation via changes in the action of endocrine regulators including insulin, leptin, ghrelin,

and glucagon-like peptide-1 (GLP-1). In such instances, disruption to homeostatic feedback systems results in appetite dysregulation promoting overconsumption, energy imbalance, and subsequent weight gain. Essentially, the social complexities of eating behaviours within an obesogenic environment further perpetuate the accumulation of excess adipose tissue.

Is the environment playing on our genetics?

It is widely accepted that genetics are influential in the development of obesity, notably when interacting with an obesogenic environment. Rare single gene mutations have been attributed to several genetic syndromes, where obesity is a characteristic hallmark. Additionally, single gene mutations may predispose an individual to develop obesity, such as the variation observed in BMI attributable to the FTO gene and its influence on appetite regulation in children (36–39). However, rather than an influence of a single gene for most individuals', genetic influences for 'common' obesity appears to be an interplay between many genes (polygenic obesity). The complexities of how these genes function to increase susceptibility are just starting to be unravelled with genome wide association studies; for example, genes expressed in the hypothalamus and adipose tissue influence energy intake and energy expenditure (40). However, Locke et al. (40) results suggest that the greatest risk to developing obesity and metabolic disorders is seen not from the presence of these genes, or any gene acting in isolation; rather, through the interplay of multiple genes interacting with their environment throughout the life cycle. For instance, lifestyle factors which could influence either energy intake or energy expenditure such as, diet and physical activity.

Twin and adoption studies provide convincing evidence that there is a genetic component to weight gain and the development of obesity (39,41,42). However further studies emphasised how healthy lifestyle behaviours have the potential to ameliorate such genetic predispositions (38,41–44). For example, in genetically predisposed individuals, consuming more fruit and vegetables (43) or improving habitual diet quality are associated with lower BMIs (44). In contrast, frequently consuming fried foods (45) or not being regularly physically active (46) have been associated with higher BMIs. It is well established that the environmental factors can have a considerable role in influencing gene expression through epigenetic changes, where the genes' activity can be altered without a change in DNA sequence (47). For example, for some individuals, weight gain can facilitate a change in gene expression which increases

the risk of developing T2DM, and changes in gene expression that occur in utero can increase the risk of the offspring developing obesity (47). Thus, epigenetic changes could be involved in the pathogenesis of obesity for some individuals. Unfortunately, it appears that many environmental factors which could contribute to epigenetic changes and obesity (i.e., influencing lifestyle factors) are often out of the individual's control.

2.2.5. Obesogenic environment

An environment which promotes weight gain whilst inhibiting weight loss has been termed 'obesogenic' (28). Such an environment is characterised by an abundance of inexpensive energy dense and nutrient poor 'discretionary' foods, accompanied with infrastructure and lifestyle conveniences which enable sedentary behaviours and leisure time (28,48,49). The industrialisation and urbanisation of Western environments has promoted systemic sedentary behaviours. Workforces are more automated, and transportation systems and numerous other daily activities of living require less physical exertion than ever before. Therefore, lack of physical activity is a major contributor to the development of obesity, as is excessive food (and energy) intake. In 2009, the American College of Sports Medicine recommended that 150-250 minutes per week of moderate physical activity (i.e., brisk walking) was sufficient to avoid weight gain (50). In corroboration, the WHO recommends engaging in at least 150 minutes of moderate physical activity per week for adults, with additional health benefits conferred when achieving up to 300 minutes (51). Higher levels of habitual physical activity are associated with lower BMIs, improved cardiometabolic health, and mental wellbeing (52). However, unless an individual makes a conscious effort to be regularly physical active, it is relatively easy to live a sedentary lifestyle within an obesogenic environment. Thus, westernised environmental pressures are considered a major influence on positive energy balance, by reducing incidental activity and providing ample opportunities to satisfy dietary overindulgence (e.g., increased marketing and ease of access to cheap energy dense foods) (49,53).

2.2.6. The modern food environment

Modern lifestyle changes are fuelling weight gain, and the food environment and systems are considered key drivers of the global obesity epidemic (29,53). The food environment influences food choice. The way we eat and what we choose to eat is intertwined with how we live and engage with our physical and social environments, which subsequently has the power to impact on nutritional status (49,53). A healthy food environment is characterised by the availability of, and access to, affordable foods that enable the population to adhere to dietary guidelines, with widely publicised promotion of such guidelines (49,53).

A shift in the last several decades has seen the current food environment saturated with an abundance of processed, 'ultra-processed', ready to consume, and 'discretionary' foods (e.g., confectionary, potato chips, biscuits, cereal bars, sugary drinks) that are easily accessible, cheap, and highly marketed (54). We are living in a profit driven environment where products are developed to drive a margin. These profit driven foods are characterised by their energy dense, yet nutrient poor, and highly palatable composition, which is achieved by utilising highly refined and processed products, derived from whole foods, to formulate these products (55). Price is a major determinant of food choice, and unfortunately in low-income countries healthy fresh foods (e.g., fruit and vegetables) are often more expensive in comparison to such ultra-processed foods (e.g., sugary and salty snacks) (56). However, high-and middle-income countries have increased their consumption of such convenience foods, snacks, and sugar sweetened beverages (SSB) (55), thus driving the availability of, and ease of accessibility to, inexpensive ultra-processed foods (54). Higher consumption of such 'discretionary' refined and processed products likely displaces the intake of core nutrient-dense foods (e.g., fruit, vegetables, wholegrains, lean proteins), which is transforming dietary patterns at the individual and population level, and is subsequently driving the burden of diet related NCDs (e.g., T2DM and obesity) (57). Globally, sales of ultra-processed foods are increasing, which is positively associated with BMI (58). Thus, despite higher quality diets being associated with higher SES (59), the relative cheapness of such 'discretionary' foods is proposed to be directly attributable to rising obesity prevalence in high-income countries due to the ease of accessibility and palatability (56). Many countries are now facing a double burden of disease, where undernutrition and overnutrition are co-existing (1), due to the food environment (49).

Foods deliver energy and nutrients. Thus, food-based guidelines are perceived to be more tangible for the public to translate to nutrition practices (60,61). There has been a global movement towards developing and implementing sustainable food-based dietary guidelines (60–62). Such guidelines are based on evidence derived from diet and health research. They are adapted to a country's nutrition practices (62) and reinforce the ideal patterns of food consumption that provide an optimal balance of nutrients to promote health and prevent chronic diseases (60,62,63). Adoption of these recommendations is likely aligned with higher disposable income (59), personal values, and potentially higher educational achievement. For example, such individuals could financially afford to eat core foods often and may actively implement these recommendations to guide their dietary habits due to being health conscious. However, the way people eat must make sense in the context of their socioeconomic and cultural setting. To save costs, consumers with limited financial resources often choose inexpensive energy dense refined foods (53,64). Thus, the cost and convenience of food will always play a role in consumer choice, as will palatability, which can easily override the best intentions to adhere to dietary guidelines.

The quality of the food sources consumed directly impact health outcomes, and the current food environment is driving problematic dietary changes at a global level. A previous assessment of diet quality in 187 countries highlighted a global trend where consumption of 'discretionary' and processed foods outpaced that of core foods (65). Core or 'healthy' foods are considered nutrient-rich foods from the four core food groups (63,66). In contrast, 'discretionary' foods are do not fit into the core food groups and are characteristically energy dense and nutrient poor processed foods (61,66). The most recent 2017 global burden of disease (GBD) study reiterated such trends, while emphasising the impact of food choice on health outcomes at a global scale (57). Suboptimal diet, defined as higher intake of 'discretionary' foods and nutrients (e.g., SSB, processed meats, sodium, and trans-fat) and inadequate intake of core foods and nutrients (e.g., vegetables, fruit, nuts and seeds, wholegrains, dietary fibre, calcium), was associated with increased risk of premature death (57). At this global scale, the authors observed a SES gradient of effect where high-income countries had a lower burden of exposure to such dietary risks in comparison to low- and middle-income countries (57).

Unsurprisingly, higher quality diets are consistently associated with higher SES, likely due to higher disposable income; and are often observed in older women who are more likely to attempt to manage their weight (59). However, food choice being dictated by cost is nothing new in NZ. In 1943 Gregory et al. (67) reported that the higher cost of fresh fruit and vegetables was a barrier to consumption; and recently, Amoah et al. (68) commented on the social inequities of food cost for those living in the most deprived households in NZ. Amoah et al. highlighted the ever-increasing cost of fruit and vegetables, coupled with wages that have remained relatively constant for these individuals, subsequently decreasing the affordability and thus increasing barriers to accessing the recommended servings of fruit and vegetables a day (63,68). Furthermore in NZ's last national nutrition survey (NNS, conducted in 2008/09), individuals living in more deprived areas reported that, within their household, a lack of money restricted the variety in their diet and the amount of food they were able to purchase and consume (69).

A survey conducted in five different European countries observed similar associations; individuals from lower SES neighbourhoods consumed less fruit and vegetables, more SSB, and had higher BMIs in comparison to higher SES neighbourhoods (70). It has been suggested these food consumption behaviours manifest because of perceptions held by lower SES families; for example, perceptions that healthy foods such as fruit, vegetables, and whole grains are more expensive in comparison to convenience foods (59) and more time consuming to prepare (71). However, contextual factors affecting dietary intake for an individual (and population), such as affordability and accessibility, must be considered. For instance, if a food costs more, has a reduced shelf life, and requires preparation, then it may not be economically viable for the individuals' current situation (64). Thus, nutrition decisions must make sense in the context of people's lives. NZ is in need of an updated NNS to assess the current dietary trends of the population, notably in the face of changing food systems (58). However, Amoah et al. (68) contended the increasing cost of core foods could easily fuel the same trends observed within the last NNS.

2.3. Dietary assessment

It is well recognised that the association between diet and health is complex and multifaceted to investigate, especially in free-living humans. Despite the striking observations of the 2017 GBD study (57), the authors highlighted practical and methodological issues in diet health studies, which hinder definitive conclusions on dose response relationships of nutrients and foods being identified. Self-reported dietary data is the basis of most dietary data in health research, and it is widely appreciated that there are numerous factors which can increase the likelihood of misreporting, thereby reducing the overall reliability of results. Additionally, deaths due to some risk factors may not be mutually exclusive; thus, the burden of disease attributable to diet may be overestimated in the report (57). However self-reported dietary data offer rich insight into the complexities of what and how individuals eat that no current set of reliable biomarkers are able to provide (72). Further, the nature of the research question and study design will influence the combinations of the dietary components and health outcome investigated.

Large epidemiological and observational studies have the power to detect relationships but cannot infer causality. In contrast, intervention studies directly manipulate food and nutrient intake and can infer causality; however, long-term interventions are rare due to compliance and cost. Thus, such studies often have reduced power and the inability to assess long-term impact. In addition, careful consideration needs to be given to the control or reference group employed to assess the effect/s of the diet. Despite not being able to infer causality, large epidemiological and observational dietary research has provided rich insight into foods and patterns associated with positive health outcomes. For example, the intake of whole grains, fruit, vegetables, and fish have all been associated with a decreased risk of CVD, stroke, and obesity; opposed to higher intake of red meat, processed meat, and SSB which increases risks (73–75). Thus, identifying diet and health associations from large epidemiologic studies can inform intervention trials to explore causality. Furthermore, replication of findings in interventions studies is important to support the outcomes and the narrative and generalisability of the findings of observational studies.

2.3.1. Dietary assessment methods

There are two main approaches to capture dietary data: prospective and retrospective recording of dietary intake to assess current or habitual intake. When designing a study, the appropriate dietary assessment tool depends on the aim, design, and study population itself (e.g., age, literacy), whilst considering the potential participant and researcher burden (as summarised in Table 2.1). Common assessment methods include FFQs, food records (weighed or estimated), 24-hour recalls, and diet histories. The gold standard method for capturing dietary data is a weighed prospective food record, however, more accurate methods are typically more expensive and have a higher participant burden (76). In addition, the techniques used to assess dietary intake need to be relevant in the population and era in which they are used (e.g., a household survey may not be relevant today as many individuals eat more foods and meals out of the home (77)).

2.3.2. Errors associated with dietary assessment

Poor validity in dietary data arises from errors in dietary assessment methodology, and there is the potential for errors to arise throughout the collection, processing, analyses, and interpretation of the dietary data (Table 2.1). For example, choosing an inappropriate dietary assessment method, interviewers may introduce bias, many tools rely on an individual's memory, and computation issues can arise when estimating portion sizes (76). All these factors can introduce measurement error in dietary assessment, with the potential to influence the dietary data captured, which, in turn, can lead to false associations during data analysis. Individuals may consciously or subconsciously under- or over-report their energy intakes, due to the social desirability of expected responses, such as underreporting foods perceived as 'unhealthy' and over reporting those considered 'healthy' (27,76,78). Misreporting of energy intake is a widespread issue within the nutrition literature and is commonly associated with obesity and dieting behaviours (27,76,79). There are methods to identify misreporters such as the Goldberg method which determines misreporting based on a cut off ratio (80). The method assumes the individuals are in energy balance (not dieting), and the ratio is based on reported energy intake, estimated basal metabolic rate and estimated physical activity levels (*reported energy intake/basal metabolic rate = physical*

activity level, EI:BMR= PAL). Once misreporters are identified they are recommended to be excluded

Due to the widespread misinterpretation of the application of the Goldberg method, Black (78) (one of the original authors) later clarified the principles and relevant use of the Goldberg method, further encouraging critical evaluation of potential influences on reported nutrient intakes before exclusion. Black (78) further recommended, if feasible secondary to cost, the use of triaxial accelerometers to measure actual energy expenditure, which can then be directly compared to reported energy intakes, thereby making the use of the Goldberg equation irrelevant (78, 80).

Arbitrary epidemiological cut offs of >500kcal/day and <3500kcal/day for women and >800kcal/day and <4000kcal/day for men have also been proposed (27). Studies with ethnic minority groups have employed higher cut offs (81,82). However, there is also widespread discussion on the potential bias introduced by the inclusion and/or exclusion of individuals identified as misreporters (79,83). The DIET@NET guidelines were established as 'best practice' for dietary assessment, providing guidelines to adequately design studies to reduce misreporting (79). It is important to explore the plausibility of the data before excluding participants based on cut-offs (79), as very low and high energy intakes may be correct in relation to individuals' physical activity levels, energy expenditure and the intra-individual variation inherent in day to day dietary intake (27,76,79).

Table 2.1. Dietary assessment tools: Advantages and disadvantages

Tool	Description	Advantages	Disadvantages
		Prospective	
Food record (weighed)	Participant required to weigh and record all food and beverages consumed over a specific time period (e.g., 3-5 non- or consecutive days) Assess actual intake of individual	• Considered the gold standard of dietary assessment • Used for validation of other dietary biomarkers or assessment tools to see how they are performing • Less reliance on memory • More accurate than an estimated food record • Provides highly detailed dietary information which is quantifiable • Higher reproducibility than estimated food record • May capture more cultural foods and cooking practices • Can capture weekday and weekend intake	• High participant burden which may impact food choice and eating behaviours (e.g., simplify their eating behaviours due to the weighing process) • Requires more detail (e.g., brand names of foods consumed) which may lead to low compliance • Significant under reporting may occur due to high burden and social desirability in reporting foods consumed "good" or "bad" foods. • Accuracy of record relies on participant motivation • The number of days recorded is prone to day to day variation in food intake • Need to ensure all days of the week are covered/represented for the population
Food record (estimated)	Participant estimates portion sizes consumed or/and in household measures (cups, spoons) for a specified time period (e.g., 3-5 non- or consecutive days) Assess actual intake of individual	• Less reliance on memory • Less participant burden in comparison to weighed food record • May capture more cultural foods and cooking practices • Intake is quantifiable • Can capture weekday and weekend intake	• Participants may not estimate portion sizes correct leading to errors in reporting • Requires more detail (e.g., brand names of foods consumed) which may lead to low compliance rates as it relies on participants' motivation to complete it accurately • Potential for social desirability bias in reporting "good" or "bad" foods • The number of days recorded is prone to day to day variation in food intake • Participants may change behaviours due to being conscious of food choice • Need to make sure all days of the week are covered/represented for the population

Tool	Description	Advantages	Disadvantages
		Retrospective	
24-hour recall	Trained interviewer assesses daily intake of a participant for the 24-hour period prior Typically used for determining intake for large populations (e.g., national nutrition surveys) If portion sizes are collected specific nutrient intakes can be calculated based on food composition data	• Low participant burden, so compliance is high • Can be completed face to face or over the phone • Relatively cost effective and time efficient • Participants are less likely to change eating behaviours • Intake is quantifiable • Can be a valid measure of usual intake if multiple single day recalls are collected • Repeated 24-hour recalls can be carried out on a subset of the population and extrapolated	• Retrospectively completed and relies on memory therefore some foods maybe forgotten • The recall day needs to be repeated on several days (including weekends) to be reflective of usual intake • Requires a skilled interviewer and a motivated participant who is honest • Under-reporting of energy intake • Estimation of food portions • Potential for social desirability in reporting "good or bad" foods
Diet history	Participants questioned by trained interviewer To assess estimated usual food intake and meal patterns for a longer period of time (1 month)	• Estimates food intake and meal patterns • Does not affect eating behaviours as participants are not recording it themselves • No limit to responses (like an FFQ); therefore, can overcome the limitations of an FFQ • Can assess usual intake of a long period of time and takes seasonal variation in dietary intake into account	• Relies on memory • Labour intensive and requires a skilled interviewer with nutrition knowledge • Mis reporting common. Can overestimate nutrition intake in comparison to food records • Foods may not be reported due to social desirability
Food frequency questionnaire (FFQ) Quantitative/Semi-quantitative	Uses frequency of use responses to assess which foods are consumed during a specific time period (e.g., daily, weekly, monthly or yearly)	• Low - moderate participant burden – depending on the length of the FFQ • Participant can complete independently either on paper or online • Low researcher burden because the results are easy to collect and process • Represents participants intake over an extended period of time • Suitable for large groups/studies – widely used • Can be used to assess patterns of intake	• Retrospectively completed • Relies on memory and can be cognitively challenging trying to estimate amounts and frequencies of foods consumed • Lower accuracy compared to other methods • May not be representative of usual intake (e.g., foods or portion sizes typically selected by participants)

Tool	Description	Advantages	Disadvantages
		Retrospective	
FFQ Quantitative/Semi-quantitative (*continued*)	Quantitative require participants to record portion sizes Semi-quantitative provides standard portion sizes Assesses habitual/ long term dietary intake	• Quantitative require participants to record portion sizes (small, medium, or large) – can increase participant burden • Semi-quantitative FFQs assess the frequency of intake, and provides standard portion sizes • Does not affect dietary behaviours as participants are not prospectively recording their intake • FFQ data can be used to rank individual's intake (e.g., low, medium, high) of foods or nutrients.	• Prone to overestimation as participants estimate portion sizes • Data can be invalid if multiple food items are contained in the same listing (e.g., when foods are reported separately and as mixed dishes)
FFQ Qualitative	Uses frequency of use responses to assess which foods are consumed during a specific time period (e.g., daily, weekly, monthly or yearly) Assesses habitual/ long term dietary intake	• Low - moderate participant burden – depending on the length of the FFQ • Participant can complete independently either on paper or online • Low researcher burden because the results are easy to collect and process • Represents participants intake over an extended period of time • Qualitative FFQ assess frequency of consumption of foods • Can be used to assess patterns of intake • Suitable for large groups/studies – widely used • Does not affect dietary behaviours as participants are not prospectively recording their intake	• Retrospectively completed • Relies on memory and can be cognitively challenging trying to estimate amounts and frequencies of foods consumed • Lower accuracy compared to other methods • Typically, only assesses food items that are included in the questions (e.g., reduced ability to capture the intake of cultural or less frequently consumed foods) • Prone to overestimation as participants estimate portion sizes • Data can be invalid if multiple food items are contained in the same listing (e.g., when foods are reported separately and as mixed dishes)

Table adapted from Gibson (76) and Willet (27)

When investigating diet-health associations it is important to adjust dietary data for energy intake (27). Higher energy intakes may contribute to disease risk; however, the intake of most nutrients is correlated with energy intake, as individuals who consume more energy will consume more nutrients. Consequently, the health associations identified (e.g., with specific nutrients) may be distorted due to total energy intake, as the body size (including physical activity and basal metabolic rate) of an individual is a major determinant of energy intake. Therefore, during analysis of dietary data, adjusting for energy intake is usually appropriate as absolute amounts of nutrients may have less of an effect on a larger bodied higher energy consuming individual, in comparison to a smaller bodied lower energy consuming individual (27,84). Thus, nutrient intake in relation to the total energy intake of the diet is of most interest, as this can infer the composition of the diet and effect of the nutrient/s per se.

2.3.3. Habitual dietary data

When exploring the association between diet exposure and health outcomes, habitual diet is most often of interest to researchers (especially in longitudinal observation studies). As highlighted in Table 2.1, dietary assessment tools, such as FFQs can capture habitual dietary intake; however, despite reducing participant and researcher burden, these tools are prone to overestimation (27,76,85). Further, one of the biggest challenges is accounting for the inherent day-to-day variation, including, episodically consumed dietary components which contribute to overall diet. Day-to-day variation refers to the fluctuation in dietary intake that occurs from one day to the next, and episodically consumed dietary components are foods and nutrients that are not consumed daily (86).

2.3.4. Statistical methods to estimate habitual diet

Due to the cost and impracticalities of capturing habitual dietary data on a large scale, statistical methods have been derived to enable the use of cost-effective dietary assessment tools (e.g., 24-hour recalls), and to reduce measurement error (i.e., misreporting) (86). Statistical methods can therefore be used to estimate a population's adherence to food based dietary guidelines, which can then be related to health outcomes. However, there are some statistical challenges when estimating habitual intake of foods, because unlike nutrients, they are often not consumed daily. Dodd et al. (86) discussed the challenges as follows: 1)

accounting for episodically consumed foods, 2) allowing for high intake (positively skewed) of foods, 3) distinguishing day-to-day variation of intake within individuals from variation observed between individuals, 4) allowing for the probability of consuming a food (and how much consumed), 5) including how the estimated usual intake relates to covariate information (e.g., age, sex, ethnicity) (86). In 2006 Dodd et al. explored four different statistical methods which were available, concluding that the Iowa State University Foods (ISUF) method addressed most of these statistical challenges. However, the ISUF method does not allow for the probability of consumption (e.g., episodically consumed food/s), nor can the influence of covariates (as these can influence the amount consumed) be incorporated into both parts of the model, and thus within the estimation of habitual intake (86). Dodd et al. recommended that the ISUF method be extended to meet these challenges.

Tooze et al. (87), proposed a new statistical method to address all of the challenges in estimating habitual dietary intake mentioned above, and specifically to account for episodically consumed foods. The NCI method was developed at the United States NCI by Tooze et al. (87) to estimate the probability of consumption of a nutrient, food or food group, and the respective consumed amount, considering covariates which affect either the probability of consumption or the amount consumed. The NCI method uses a two-part modelling approach and is based on two or more 24-hour recalls. The basic assumption of the NCI method is that the 24-hour recall is an unbiased measure of dietary intake for a consumption day. For episodically consumed dietary components (e.g., foods) the first part of the model estimates the probability of consumption using a logistic regression model allowing for a residual between-person variation (covariates) to be included. The second part of the model uses linear regression on a transformed scale (as the NCI method requires near normal distribution of the residuals) to estimate the amount consumed and, like the first model, considers covariates. The inclusion of covariates (e.g., age, sex, ethnicity) within the model accounts for the intra individual day-to-day variation caused by these variables. The individual habitual intake is then defined as the product of probability of consumption multiplied by the consumed amount (*Usual intake = Probability x Amount*). Estimation of a dietary component consumed daily (e.g., nutrients) requires a one-part model, where the process is the same as mentioned above. However, there is no need to model probability of consumption, so this step is omitted. The dietary data from the 24-hour recalls are the dependent variable, and the

average daily intake of dietary components from a FFQ can also be used as a covariate within the model. Kipnis et al. (88) described the benefits of including covariates within the modelling process, especially how the inclusion of FFQ information improves the precision in estimation of episodically consumed foods or food groups.

2.4. Dietary patterns

Historically, large epidemiological trials focused on inadequate nutrient intake and their association with health status (89). However, in the modern era we are now faced with the double burden of disease. Thus, there has been a shift to focusing on the whole diet and exploring optimal combinations of foods and nutrients consumed together and subsequent associations with health outcomes (61,89).

2.4.1. Dietary pattern analysis

Analysis of dietary patterns enables the assessment of the interrelationship between foods and nutrients consumed together (e.g., as meals), and how they relate to disease and health outcomes. There are two main approaches to assess dietary patterns in self-reported dietary data, *a priori* or *a posteriori* approaches. *A priori* approach utilises a predefined criterion and scores how a diet aligns to this criteria (e.g., Mediterranean Diet Score, Healthy Eating Index (HEI), Alternative Healthy Eating Index (AHEI)). The HEI is used to assess dietary quality by scoring (out of 100) how close an individual's reported dietary intake aligns with the Dietary Guidelines for Americans (90). A higher score (e.g., 90/100) reflects a diet that aligns to the dietary guidelines. *A priori* dietary pattern tools are often updated to align with new knowledge (i.e., updated guidelines), thus numerous editions evolve over time. In contrast, *a posteriori* approaches are data driven using statistical methods where dietary data are reduced to a smaller number of variables (e.g., principal component and cluster analysis). Principal component analysis (PCA) reduces dietary data into patterns based on dietary components which are correlated with each other (e.g., consumed together) (91); whereas cluster analysis reduces dietary data in to distinct clusters, based on differences in mean intake of dietary components between individuals (91). Cluster analysis, therefore, assigns an individual to a mutually exclusive group and, PCA assigns an individual a score of adherence

to the pattern. Clusters can reduce the statistical power to identify associations between dietary intake and disease, thus the continuous nature of PCA is seen as more advantageous as it may increase the power to detect associations (91). *A priori* dietary pattern also assigns an individual a score of adherence (to the predefined criteria), where the scores reflect their intake and, similar to PCA, these scores can be used to rank individuals and differentiate between high and low intakes (91). Thus, the decision to use either approach depends on the study design and the population of interest. *A priori* criterion must be relevant to the dietary practices of the population to whom it is applied, and *a posteriori* method can identify unique eating patterns of a study population, however, these are not generalisable. In addition, the quality of the dietary assessment initially conducted will determine the quality of the dietary data for either *a priori* or *a posteriori* approaches. Further, subjectivity can be introduced during dietary pattern analysis; for example, the creation of food groups, metric of input variables (e.g., g/day, percentage of energy), number of patterns to retain and naming of the patterns (92), which can all influence the patterns identified.

2.4.2. Dietary patterns and health outcomes

Dietary patterns assess the whole diet and for this reason are more likely to be associated with health outcomes rather than a single nutrient or food. It is highly unlikely that a single nutrient or food will cause or prevent chronic diseases, like NCDs; thus, focusing on the patterns of foods consumed together enables a more consistent (and tangible) message to promote healthy eating across numerous chronic and complex conditions (61,93). During analysis, the dietary patterns identified will be influenced by gender, ethnicity, SES, location, and age (73,91); yet, similar patterns of foods have been identified and associated with health outcomes across populations. For example, 'unhealthy' or 'Western' patterns characterised by high intake of animal protein and fat, refined, and sugar sweetened products (e.g., refined grains, processed meat and foods, and SSB) are associated with an increased risk of weight gain, obesity, and chronic diseases (73,74). In contrast, 'healthy' or 'prudent' patterns characterised by high intake of unrefined carbohydrates, more vegetables than animal protein, and unsaturated fats (e.g., fruit, vegetables, wholegrains, nuts, and lean proteins) are associated with decreasing these risks (73).

Across studies, higher quality diets, defined by higher scores on dietary pattern indexes such as the HEI, AHEI, and the Dietary Approaches to Stop Hypertension (DASH), have all been associated with a reduced risk of chronic disease (74). Further, despite some controversies, the Mediterranean diet is the only dietary pattern to be associated with decreased risk of CVD in both observational and intervention studies (73,94). The Mediterranean diet is characterised by high intake of vegetables, fruit, nuts and seeds, legumes, wholegrains, seafood and olive oil, with minimal red meat, dairy, and limited intake of refined, processed and sugar sweetened foods.

The dietary pattern indexes HEI, AHEI, DASH and Mediterranean diet all prioritise different combinations of foods; yet emphasise similar foods as markers of diet quality. However, the heterogeneity across studies has hindered the ability to make definitive conclusions to inform public health policy. In accordance Liese et al. (95) developed The Dietary Patterns Methods Project (DPMP), with the objective to conduct standardised and comparable analysis across studies, by applying four commonly used dietary indexes (HEI-2010, AHEI-2010, DASH and alternate Mediterranean diet) to three prospective cohorts, to improve the consistency and reliability of results. Across studies these indexes could consistently identify individuals with higher quality diets, which were associated with a reduced risk of all-cause mortality. Diet quality was characterised by higher intake of wholegrains, fruit, vegetables and plant-based proteins across study populations.

In a recent study, Wang et al. (96) applied an adjusted AHEI to the same dietary data from the 2017 GBD study (within a sub sample of 190 countries) to assess global diet quality. Wang et al. focused on the combination of 10 dietary components consumed together and compared this to a reference dietary pattern (i.e., high intake of, fruit, vegetables, wholegrains, nuts and legumes), which was a modified version of the EAT-Lancet Commission's sustainable plant-based diet (96). Essentially, core dietary components were assigned positive scores where discretionary components were assigned negative scores. Overall dietary quality improved slightly from 1990 to 2017, despite this, it was far from optimal and quality varied largely around the world. At this global scale, lower quality diets were associated with industrialised food systems and characterised by higher processed food intake and thus higher intake of added sugar or trans-fats. Ultimately, they concluded that over 11 million premature deaths

could be prevented annually with improvements in dietary quality (e.g., consumption patterns similar to the healthy reference diet rich in core foods). The authors further stated that "the best time to adopt a healthier diet is now" (97 p1072) to reduce the global burden of NCDs attributable to poor dietary intakes. Taken together with the DPMP, diet quality appears to be consistently characterised by key drivers which impact chronic disease risk.

The evidence suggests there are key drivers and many ways to consume a healthful diet, which are relevant to a population's socioeconomic position, including sociocultural and nutrition practices (60,62). The United States Department of Agriculture (USDA) conducted a series of systematic reviews exploring the relationship between dietary patterns and health outcomes worldwide (73). They concluded that there was not one dietary pattern more favourably associated with positive health outcomes (due to the heterogeneity in the use and naming of patterns, observed variation in diets across study populations etc.); however, specific dietary components were consistently observed across the patterns. Dietary patterns characterised by higher intake of vegetables, fruit, plant-based proteins, wholegrains, unsaturated fats, fish, and low fat dairy were consistently associated with a reduced risk of weight gain, obesity, T2DM, and CVD (73). Therefore, globally, diet quality appears to be determined by a high intake of core foods and minimal intake of discretionary foods.

Food source

As discussed above, the quality and quantity of the food source consumed can have implications for health outcomes. Reference guidelines such as the acceptable macronutrient distribution ranges (AMDRs) are recommended ranges of energy intakes to consume from macronutrients (expressed as a percentage) and can differ between countries (98,99). These AMDRs complement food based dietary guidelines and aim to reduce the risk of chronic illness and disease, while ensuring adequate micronutrient intake for healthy living (99). A recent meta-analysis by Seidelmann et al. (100) explored the influence of different food sources contributing to total energy intake, and subsequent risk of mortality. Consuming either a low (≤40 %) or high (≥70 %) energy intake from carbohydrates was associated with an increased risk of all-cause mortality, whereas consuming between 50 and 55 % of total energy intake from carbohydrates was associated with a decreased risk. Despite this, risk of mortality increased if carbohydrate intake was replaced with animal proteins and fats, and decreased if

replaced with plant-based proteins and fats (100). Therefore, the quality of the food source that contributes to overall energy intake is perhaps more important to focus on. For example, in NZ the AMDR for percentage of total energy to consume from total fat is 20-35 %; however the Mediterranean diet which is associated with a decreased risk of CVD (73), is characterised by up to 40 % energy from total fat intake. The important differentiation is the source of the fat. For the Mediterranean diet, fat source is primarily unsaturated, as higher saturated fat intake is associated with an increased risk of CVD (101,102).

Further highlighting the importance of the quality of the food source, Satija et al. (103) explored the effect of consuming either a healthy or unhealthy plant-based diet on the incidence of CHD events in three large longitudinal cohorts: The Nurse's Health Study (NHS), the Nurse's Health Study II (NHS2), and Health Professionals Follow-Up Study (HPFS) (103). Semi-quantitative FFQs, administered every 2-4 years, were used to create three plant-based diet indexes (overall, healthy, and unhealthy plant-based diets). Overall, higher adherence to a healthier plant-based diet rich in core foods (i.e., fruit, vegetables, wholegrains, nuts and legumes) was associated with a decreased risk of CHD; in contrast, higher adherence to a plant-based diet with higher discretionary food (i.e., SSB, refined grains, fried potatoes) intake was associated with an increased CHD risk (103). Individuals who scored higher on healthier plant-based diets were leaner, older, and more active, which highlights that the health associations observed may not solely be attributable to diet. However, taken together with the recent meta-analysis by Seidelmann et al. (100) and the systematic reviews from USDA, the key to diet quality is focusing on the quality of the food source, by consuming more core and less discretionary foods (73).

2.4.3. Healthy dietary pattern and the beneficial role of dietary fibre

A healthy dietary pattern is characterised by consuming more quality core foods. The World Cancer Research Fund (WCRF) recommends consuming a diet rich in wholegrains, fruit, vegetables and beans, limiting red and processed meat intake; and maintaining a healthy weight, to reduce the risk of developing cancer (104). Diets rich in unrefined complex carbohydrates such as wholegrains are protective against, and inversely associated with, components of the metabolic syndrome (e.g., abdominal obesity, high blood pressure, and

insulin resistance) (75,105–107). Higher intake of wholegrains, including fruit and vegetables decreases the risk of weight gain and obesity (43,75), and higher intake of fruit and vegetables is associated with a decreased risk of developing breast cancer (108). Further, a meta-analysis by the WCRF (104) found that for every 10 g/day increase of dietary fibre intake (from foods with naturally occurring or added fibre) the risk of colorectal cancer decreased by 9 %. A recent meta-analysis by Reynolds et al. (109) further corroborated the beneficial and health protective association of higher fibre and wholegrain intake. The authors recommended that adults aspire to consume at least 25g of dietary fibre a day from quality carbohydrate foods, to reduce the risk of developing NCDs, cardiometabolic diseases, and all-cause mortality (109).

The numerous health benefits associated with consuming quality core foods rich in dietary fibre has been attributed to the functionality of the dietary fibre (110,111). Dietary fibre is defined as the indigestible carbohydrate plant polymers found intrinsically within foods, including isolated or chemically synthesised fibres (111,112). Fruit, vegetables and wholegrains are rich sources of indigestible polysaccharides (cellulose, hemicellulose, and resistant starch), which are collectively described as dietary fibre (111). Dietary fibre can also be classified as insoluble and soluble fibre; however, this distinction is considered to be less meaningful, because the action and health effect of dietary fibre differs depending on its physiochemical properties (e.g., solubility, viscosity, and fermentability) (110,111,113). Higher fibre intake has been shown to improve satiety and thus reduce overall energy intake, as well as improving insulin sensitivity (114), which aligns with higher habitual fibre intake being associated with lower weight gain and BF% overtime (43,73,107,115). Historically defining and translating a clear health role of dietary fibre has been challenging, in part due to the heterozygous nature and functionality of fibre, coupled with the different definitions of dietary fibre and methodologies employed across studies (e.g., how fibre is analysed, type of fibre used, and populations investigated). Thus, there is appreciation for the importance of defining a dietary component by its physiological function (i.e., decreases colonic transit time) rather their biochemical composition (113,116). Chen et al. (105) reiterated such issues within the studies they explored in their meta-analysis, highlighting that the source and functionality of different types of fibres (with the potential to exert different preventative effects) was often unaccounted for. In addition, despite finding an inverse association with dietary fibre intake and risk of metabolic syndrome, they concluded insufficient evidence to draw definitive

conclusions (105). Therefore, ensuring the type of fibre is clearly reported in future studies will improve the clarity of results and enable informed conclusions to be drawn.

Despite the difficulties in defining or classifying the food source of dietary fibre, dietary fibre is largely indigestible to the human host. Indigestible polysaccharides reach the large colon where they are fermented by the gut microbiota (the community of bacteria residing in the human gut) (117). The principle role of the gut microbiota is to degrade and ferment otherwise indigestible substrates which escape human digestion in the small intestine (118). The microbiota are equipped with more carbohydrate degrading enzymes than the host, thus, 80 % of the dietary fibre consumed is degraded and fermented by the microbiota (113). A product of fermentation are SCFAs (mainly acetate, propionate, and butyrate) and gases. These SCFAs influence host physiology via a wide variety of effects such as promoting satiety (via interactions with GLP-1 and peptide YY), and improving glucose and lipid control (117,119).

It is important to consider the amount and physiochemical properties of the dietary fibre consumed (i.e., the composition of the diet), as this will directly influence the available substrates for the microbiota. For example, a low fibre diet has been associated with lower levels of butyrate, because there are less available substrates for the microbiota to ferment to produce SCFAs (120). In contrast, higher habitual fibre intake has been associated with increased microbiota diversity (number and distribution of species present) (121), including a decreased risk of developing colon cancer (104), which has been proposed to be associated with the anti-inflammatory effects of SCFAs (122). Dietary phytochemicals (consumed in plant-based foods) also have anti-inflammatory properties, and are released from plant cell walls by the microbiota (122). In addition, dietary fibre intake can regulate transit time through the colon by contributing to faecal bulk and water retention (113,116). Transit time exerts a selective pressure on the microbiota, as rapid transit would favour bacteria with shorter doubling times (123), whereas longer transit time could increase the time for fermentation in the large colon (and thus production of SCFAs). Therefore, the composition of the diet will influence the available substrates including the environmental conditions for the microbiota. Further, the beneficial effects of consuming a plant-based diet rich in core foods (and host indigestible substrates) may be mediated by the functions of the microbiota.

2.5. The gut microbiota

A growing body of evidence suggests the gut microbiota may have a profound impact on our health. The gut microbiota (referred to as 'microbiota' henceforth) is considered a functional organ of the body (124). The microbiota exist as complex community of microorganisms, where competition for dietary substrates and cross feeding on bacterial metabolites (e.g., SCFAs) occurs (125). The microbiota interact with the host's immune system to reduce the risk of some infections, and microbial metabolites affect numerous metabolic functions, which can influence the biological function and health of the host (117,125–127). For example, butyrate supports the integrity and function of intestinal epithelial barriers which are part of the innate immune system (128). Thus, the microbiota is of interest due to its metabolic influences on human health.

Over the last two decades, high-throughput next-generation sequencing techniques have evolved and become readily available and cheaper, revolutionising how the composition of the faecal microbiota is analysed. Most studies use 16S rRNA gene sequencing where variable (e.g., V3-V4) regions of a single gene (16S rRNA gene) are sequenced (125). Next generation sequencing has contributed to the early pioneering work by many microbiologists (125), and there is now certainty of the major bacterial phyla present within the human faeces. The two main phyla are *Firmicutes* and *Bacteroidetes* (125). Examples of families within the *Firmicutes* phyla include *Eubacteriaceae*, *Lachnospiraceae*, *and Ruminococcaceae*; and within the *Bacteroidetes* phyla the *Bacteroidaceae*, *Rikenellaceae* and *Prevotellaceae* families (125). Other phyla (and associated families) are present in lower relative abundance such as *Actinobacteria* (*Bifidobacteriaceae*), *Proteobacteria* (*Desulfovibrionaceae*) and *Verrucomicrobia* (*Verrucomicrobiaceae*). Next-generation sequencing provides reliable taxonomic information of the types of bacteria present to the bacterial family level. However, differentiation at the species or subspecies is not as reliable. In contrast, shotgun metagenomic sequencing sequences the whole 16S rRNA gene and bacterial genome, therefore greater analytical depth is achieved (125). In addition, metagenomic analysis enables the functional capacity of the microbiota to be assessed (e.g., identify genes involved in carbohydrate fermentation). However, the cost of shotgun sequencing is considered a

barrier to widespread use. Thus, the sequencing technique needs to be considered when interpreting results regarding the reported relative abundances of bacterial species.

Regarding the association between diet, microbiota, and health outcomes, animal models have enabled mechanisms to be identified. However, the translation of these observations to humans is currently difficult to establish. Observational and cross-sectional studies enable associations to be identified in humans although, the causal mechanisms cannot be elucidated. Alternatively, intervention studies in humans enable the identification of causal mechanism; however, emerging evidence is suggesting that the inter-individual differences in the composition of microbiota between individuals is influencing the responsiveness to dietary interventions (129–131). The variability in responsiveness is proposed to be driven by both the microbiota composition and the substrates provided. For example, the responsiveness to the delivery of any substrate will depend on the species' capability to utilise the substrate (determined by the species' genome). Specialised species may only respond to the delivery of specific substrates (e.g., *Rumniococcus bromii* and resistant starch (132,133)) in contrast to more dominant generalist species, which may have a greater capacity to switch between energy sources depending on the availability of substrates (and thus dietary intervention). Further, different species may have different tolerance to specific features of the gut environment (e.g., transit time) which can be influenced by diet. Therefore, the capabilities of the present species and the substrates provided will influence the observed responsiveness of the microbiota (i.e., changes in composition) to dietary interventions. In addition, the sequencing technique needs to be considered when interpreting results, notably in relation to diet and bacterial species abundance and health outcomes. The microbiota will now be discussed in the context of health outcomes and obesity in relation to dietary intake.

2.5.1. What shapes our microbiota?

Factors such as mode of birth, infant feeding practices, genetics, lifestyle, geographic location and habitual diet, all shape an individual's microbiota (127,134). Twin studies have indicated that although there is a heritable component to the microbiota, environmental influences are more influential in determining the composition of the microbiota (135). Therefore, inter-individual differences in microbiota composition further exacerbate the complexity of

understanding what determines a 'healthy' microbiota and raise the question, does one even exist? A 'healthy' microbiota may be contextual, and whether 'dysbiosis' (deviations from normal microbiota patterns) is a cause or consequence of disease is unknown. The early (before the age of two) and late years of life are characterised by compositional shifts and reduced microbial diversity (136). However, the composition of an adult microbiota is thought to be relatively 'stable' or persistent, despite initial changes observed following dietary interventions (137,138). The use of antibiotics can disrupt the microbiota, and across individuals the resilience of their microbiotas to antibiotic treatments differ for some, changes to the microbiota (induced by antibiotics) can persist for years (127). The use of antibiotics in early life has been proposed to increase the risk of obesity in children (139). However, a retrospective analysis exploring antibiotic use and weight gain in a cohort of twins, revealed administration of antibiotics in the first six months of life did not increase the risk of weight gain (140), highlighting that the microbiota is part of a complex issue.

2.5.2. The gut microbiota and obesity

Animal models have established a causal role of microbiota in obesity. Studies have highlighted the functional capacity of microbiota to regulate energy utilisation from food and influence the adiposity (and distribution) of the host (141,142). For example, germ free mice (mice with no microbiota) consuming a Western diet appear to be protected from diet-induced obesity (143), and the obese phenotype (from mice and humans) can be transferred to lean hosts via faecal transplants (141,144). However, the observation of the microbiota increasing energy harvest from food has yet to be replicated in other studies (145). Microbial signatures of obesity have also been proposed, such as an altered ratio of the two main phyla *Firmicutes* to *Bacteroidetes* (with a higher ratio being attributed to the obese phenotype (146,147)). In addition, reduced microbiota diversity (148), low gene richness (the number of genes associated with the microbiota present in the gut (149)) and absence of specific species (e.g., *Akkermansia muciniphila* (150)) have all been proposed to be markers of obesity. However, our understanding of the microbiota-obesity relationship in humans is only starting to be unravelled. It is further complicated because the translation of insights from animal studies to humans has yet to be elucidated and the complexity and contradictions in the scientific literature do not support clear conclusions. For instance, the heterogeneity between

sample collection and analysis techniques, including the populations investigated (e.g. age, sex, health status, dietary intake etc (127)). In addition, microbial signatures are considered too simplistic to reflect obesity and methodological limitations hinder the ability to identify causal mechanisms in humans, subsequently driving the conflicting observations across studies (151,152).

The pathogenesis of obesity is multifactorial and complex, and the microbiota is posed to be another host risk factor, as are inactivity, suboptimal dietary intake, and genetic predispositions. Sze et al. (153) reported reduced diversity, richness (number of species present), and evenness (distribution of species) were weakly associated with an increased risk of being obese, and further questioned the biological significance of these findings. Despite the heterogeneity of reports in the literature, experimental findings in animals and humans appear to be in agreement with the obese state being characterised by a state of 'dysbiosis' (147,148,154) in comparison to lean individuals who have a richer and diverse microbiota (154). However, Sze et al. (153) also proposed that it is highly likely each individual has their own microbial signature of obesity which, in theory, would mean inter-individual responsiveness to lifestyle factors such as diet. This may be why we have yet to identify a single microbiota pattern (and/or signature) which reflects obesity across populations.

2.5.3. Diversity of the microbiota and health

Advances in sequencing techniques and bioinformatics tools have enabled the characterisation of microbiota profiles and the association of these with different health states. Subsequently, the diversity of the microbiota has been posed as a marker of health status in humans, as lower diversity has been associated with obesity (148,154,155), inflammatory bowel disease (IBD) (156), irritable bowel syndrome (IBS), and T2DM (157), in comparison to healthy subjects. However, these observed associations might be contextual.

Diversity as a metric may be too crude to accurately reflect the health status of the host, as it fails to appreciate inter-individual differences in microbiota composition (e.g. differentiate between the presence or absence of specific microbiota), and thus the functions of the microbiota. For example, two individuals may have the same diversity but vary greatly

regarding the taxonomic structure of their microbiota. Insights suggest the functions of the microbiota are somewhat constant despite the inter-individual differences in microbiota composition, because different species can have the same metabolic capacity (158). Higher microbiota diversity is thought to reflect a more stable community structure (i.e., more resilient to external influences such as limited resources), because the microbiota functions as an ecosystem which interacts to utilise substrates (e.g., cross feeding between primary and secondary degraders) for mutual benefit. However, the intricacies of how the functions of the microbiota can influence the host are still being unravelled. Hindered by the challenges in identifying and characterising microbiota profiles, as the best proxy currently available, and the basis of most microbiota studies, is the faecal sample which indicates what has happened, rather than what is happening. Currently, the consensus appears to be that microbial diversity is more important than the presence and/or absence of a specific species and low dietary fibre intake does not support a diverse microbiota. Disruption or 'dysbiosis' of the microbiota appears to be associated with diseases where inflammation is a key characteristic (e.g., obesity, IBD). Although there is no definition of a 'healthy' microbiota, and likely that one does not exist, it appears that the modern lifestyle may disrupt the assembly of the microbiota which could be contributing to the increase in diseases characterised by inflammation.

2.5.4. Diet and the gut microbiota

The evidence suggests an individual's habitual diet can exert the strongest selective pressure on the microbiota (134,137), which, in turn, can influence the functionality of the ecosystem and subsequent disease risk. Early diet intervention studies cited rapid changes in the microbiota in response to dietary interventions (137), while later studies emphasised habitual dietary intake was a key determinant of the microbiota composition (134,137,138,159–161). It has been proposed that increased diversity of the microbiota enables the microbiota ecosystem to exert homeostatic regulation in response to such short term dietary interventions (127). However, as discussed earlier, responsiveness will depend on the composition of the microbiota and the substrates provided.

The human colon is densely populated with microbiota and there has been considerable interest in characterising and associating the microbiota with dietary intake. In 2011 Arumugam et al. (162) classified the microbiota, characterised in two human studies by recognising three taxonomic clusters called 'enterotypes', hypothesising that the members of each enterotype co-exist for functionality (162). These enterotypes were unrelated to nationality or host characteristics such as BMI, gender, and age (162); remained unchanged after dietary intervention; and were later associated with habitual dietary intake (137). The *Prevotella* enterotype was associated with a carbohydrate rich diet and the *Bacteroides* enterotype was characterised by an animal protein and fat diet. The third enterotype, *Ruminococcus*, was reportedly less evident and not as distinguishable by dietary intake (137). Two or three of these discrete enterotypes have been identified in different populations around the world (134,159,163–165), and some have been related to host characteristics. However, the concept of enterotypes has been hypothesised to be more of a spectrum rather than a discrete cluster, if not at least a useful tool for stratification (163).

Later intervention studies have since proposed that these enterotypes influence responsiveness to dietary interventions. In a population of Danish subjects (n=181), individuals with the *Prevotella* enterotype lost more weight and BF% when consuming a high fibre diet rich in vegetables, fruits, and wholegrains; in contrast to individuals with a *Bacteroides* enterotype (166,167). Further, self-reported dietary fibre intake at the one year follow up was strongly associated with weight loss for the *Prevotella* enterotype but not the *Bacteroides* enterotype (167). These observations appear to reflect the specialised functions of the bacteria associated with the enterotypes. Clusters of microbiota co-existing together for functionality purposes are concordant when viewing the microbiota as an ecosystem. Further, considering habitual diet drives the composition of the microbiota, in theory, an individual's microbiota could be shaped by their physical and social environments as these influence dietary intake (accessibility and affordability of food). Thus, the discrepancies between studies where differences have been observed in the composition of microbiota between ethnicities (168), geographical locations (169), and enterotypes (163) could be reflective of sociocultural and nutrition practices, including local food systems and lifestyle behaviours. Considering diet is a key modifiable risk factor in the development of obesity, and

habitual diet drives the composition of the microbiota, is the microbiota mediating the relationship between what we eat, our body weight, and subsequent disease risk?

2.5.5. Diet diversity and the microbiota

Habitually consuming a diverse range of core foods is associated with better health outcomes and appears to be positively associated with microbiota diversity (164). The nutrition literature emphasises the impact of the composition, quality, and quantity of food sources consumed, and the beneficial influence of higher habitual dietary fibre intake on health outcomes (57,73,97,100,104,109). Dietary patterns which are rich in core and plant-based foods are consistently associated with better health outcomes (73,95,100,103,109), which might be mediated by the functions of the microbiota (117). For example, higher adherence to a Mediterranean diet is associated with higher fibre degrading microbiota and increased production of SCFA (170). In contrast, following a 'Western' pattern has been associated with increased weight gain, BMI, risk of obesity and CVD across populations (73), including inflammation in the gut (171) and reduced microbiota diversity (172).

It is unclear whether dysbiosis is a cause or consequence of disease; however, habitual diet and dietary fibre intake drive microbiota composition, which in turn influences health outcomes. Combined, this suggests potential avenues for external manipulation of the microbiota with diet interventions to promote better health outcomes. However, before we can develop therapeutic targets we need to unravel what the current data is suggesting; that is, how translatable is the evidence? Most of the evidence linking diet-mediated microbiota health associations is based on animal models that, need to be interpreted carefully. More importantly, what is the research in human subjects suggesting? Can we translate any findings to practical nutrition recommendations? The complexity of capturing accurate dietary data, including the lack of standardised diet assessment across microbiota studies, has further obscured the clarity of results.

2.6. Habitual diet and the gut microbiota

Intervention and controlled feeding studies (e.g., delivering specific nutrients, or foods) cite having a rapid impact on the microbiota composition (137,173,174). However, such studies tend to focus on one specific nutritional component at a time, which fail to appreciate the synergistic effect of foods and nutrients consumed together. Intervention studies are also often of short duration (<12 weeks), likely secondary to feasibility, cost and compliance, and there are limitations of what dietary components can be explored (ethically only beneficial ones). In addition, due to inter-individual differences in the responsiveness to dietary interventions (130,131) and the potential for associations to be contextual in the population explored (e.g., taxa abundance associated with disease), exploring the analysis of habitual dietary intake in healthy adult populations would offer the best insight into the current knowledge of diet, microbiota, and health in humans.

This literature will now transition to exploring observational studies in relation to habitual diet, microbiota, and health outcomes in humans. The review will focus on exploring the dietary assessment and analyses employed within these studies and has been divided into three themes: 1) adherence to self-reported diet category, 2) nutrients and foods, and 3) dietary patterns.

2.6.1. Adherence to self-reported diet category

As summarised in Table 2.2, observational studies often focus on specific hypotheses within distinct populations such as comparison between self-reported diet categories (e.g., vegetarian, vegan, or omnivore (175–177)), number of plants consumed (178), and geographic location (165). In some studies, differences in the composition of the microbiota were observed when participants were stratified by self-reported classifications (165,178), and not in others (170,175–177). Capturing such qualitative data does not enable clear associations or relationships to be drawn, as the composition of the actual diet can vary greatly between individuals within a self-reported diet category (e.g., vegetarian), due to factors such as age, sex, BMI, SES, energy intake etc.

Losasso et al. (177) concluded that despite observing that vegetarians had higher richness and overall diversity of their microbiota in comparison to omnivores, using such broad categories to analyse the dietary data could lead to the oversimplification of results, due to failing to accurately capture the composition of the diet (e.g. nutrient intake) (177). In a recent study, Dhakan et al. (165) also employed broad dietary categories (e.g., vegetarian and omnivore) to explore the microbiota of healthy Indian adults (n=110) from two distinct locations. Dietary intake explained the largest variance in the taxonomic diversity of the participants; however, no dietary assessment was conducted (165); the foods which contributed to these diet types were based on the foods habitually consumed for that region. Therefore, despite diet explaining the largest amount of variance in the diversity of the microbiota, it is not explicitly clear what the participants actually consumed and how diet is associated with microbiota diversity or taxon abundance. Overall, such broad self-determined categories are not a useful measure to analyse diet composition in relation to microbiota composition, and should be coupled with more thorough dietary assessment methods and analyses.

Table 2.2. Observational studies exploring the association of self-reported diet category and gut microbiota composition

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Self-reported adherence to diet type in cross-sectional studies						
Losasso et al. 2018 (177)	•Healthy Italian adults (n=101) •males (32 %) •females (68 %)	•16S rRNA sequencing	•Semi quantitative FFQ covering a period of 14 days •24hour recall to assess reliability of self-reported adherence to diet type •PCA used to assess diet clusters •Self-reported adherence to diet – required to have followed for 12 months	•32 vegetarians (V: 70 % females, age 42 ± 13 years, BMI 23.8 ± 9 kg/m^2) •26 vegans (VG: 65 % females, aged 39 ± 11 years, BMI 23.7 ± 3.4 kg/m^2) •43 omnivores (OM: 73 % females, aged 45 ± 14 years, BMI 23.8 ± 4.7 kg/m^2) •BF% assessed with BIA	•Number of Bacteroidetes was higher in VG and V •No difference in Firmicutes/Bacteroidetes ratio between the diet groups •OM had lower alpha diversity compared to V (OTU and Chao1). V had higher richness •No difference between the groups for Simpson or Shannon alpha diversity •Three enterotypes were identified: dominated by: *Bacteroidaceae*, *Prevotellaceae*, and *Ruminococcaceae* but no association between enterotypes and diet group was observed	•Stratification by diet type did not result in differential clustering of the microbiota (alpha and beta diversity)
McDonald et al. 2018 (178)	•Citizen Science project n=10,000 •~79 % of the participants were from America, 25 % were from the UK and the remainder was from 43 other territories around the world	•16S rRNA sequencing •Samples shipped through the mail at room temperature (analysis adjusted for blooms)	•FFQ was administered through - Vioscreen •Self-reported diet type was used to stratify the population •Number of plants consumed was calculated to assess dietary plant diversity – limited information on methodology	•Focused on healthy adults (n=3,942) aged 20-69 years and with a BMI of 18.5–30 kg/m^2 •Self-reported meta data (e.g., weight and height) •n=1762 who completed the FFQ	•Faiths phylogenetic diversity was higher in UK samples •Multiple assessments of n=565 participants showed inter-individual variability even after 1 year •*Faecalibacterium prausnitzii* and *Oscillospira* were associated with eating more than 30 types of plants •Consuming higher diversity of plant intake was associated with higher molecular and 16S alpha diversity	•Number of plant species consumed was associated with microbiota composition more than self-reported diet type “vegan” and “omnivore”
Wu et al. 2016 (176)	•Healthy American vegans (n=15), omnivores (n=16)	•16S rRNA-tagged sequencing	•3x 24h recalls during the week before faecal, plasma and urine collection •One participant excluded based on under reporting •Nutrient intake was adjusted for energy intake	•Metabolic markers (CRP, lipid profile, fasting glucose/insulin assessed). Only total cholesterol and LDL was higher in omnivores •No difference in BMI between diet groups.	Between the 2 diet groups: •After multiple corrections there was no difference in the presence of taxa at the genus level. •There was no difference in evenness or diversity Colonic transit time was longer in the individuals with lower fibre intakes	•Vegans consumed more carbohydrate, but less protein and fat in comparison to the omnivores •Vegans consumed a mean fibre intake of 35 g/day of and omnivores 17.5 g/day •Differences in dietary intake between the diet types did not translate to differences in the microbiota

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
de Moraes et al. 2017 (175)	•Brazilian adults (n=268)	•16S rRNA sequencing •Enterotypes identified •Faecal consistency was not assessed/controlled for	•Self-reported dietary adherence to diet type: either vegetarian, lacto-ovo vegetarians or omnivore •No other dietary assessment conducted	•Age: 49 ± 8 years •54 % were women •41 % had BMI >25 kg/m^2 •Vegetarians n=66: BMI: 23.1 ± 4.1 kg/m^2 •Lacto-ovo vegetarians n=102: BMI: 24.4 ± 3.9 kg/m^2 •Omnivores n=100: BMI: 26.4 ± 4.7 kg/m^2 •Strict vegetarians had the lowest BMI	•Identified 3 enterotypes, there was no difference in sex, age or BMI between enterotypes •Enterotype 1 *Bacteroides* (n=111) •Enterotype 2 *Prevotella* (n=55) •Enterotype 3 *Ruminococcaceae* (n=102) •There were more vegans in the than the other 2, but there was no difference with the lacto-ovo and omnivores. •In *Bacteroides* enterotype: •The abundance of *Desulfovibrio* was negatively associated with BMI •The abundance of *Streptococcus* was associated with BMI •In *Prevotella* enterotype •Abundance of *Roseburia* was negatively associated with BMI • *Faecalibacterium* was negatively associated with fasting insulin *and Akkermansia* was negatively associated with fasting glucose •In Ruminococcaceae enterotype •*Roseburia* positively associated with lipid profile. •*Eubacterium hallii* was positively associated with BMI *Bifidobacterium* was associated with total cholesterol	•There were more vegetarians in Prevotella enterotype •No difference between the prevalence of lacto-ovo vegetarians and omnivores across enterotypes •Vegetarian diet was associated with lower LDL cholesterol
Other diet category in cross-sectional studies						
Dhakan, et al. 2019, (165).	•Healthy Indian adults (n=110) from either Northern-Central or Southern region	•16S rRNA gene and shotgun metagenome sequencing •Core microbiome of Indian population compared to samples from	•Diet defined by regional cultural practices (e.g., omnivore or vegetarian) •No dietary assessment conducted	•Females n=62 •Males n=58 •BMI: 21.2 ± 5.2 kg/m^2 •Age: 30 ± 17	•*Prevotella, Mitsuokella, Dialister, Megasphaera*, and *Lactobacillus* were associated with Indian population •*Bacteroides, Ruminococcus and Alistipes, Clostridium* were found in lower relative abundance in the Indian population •Indian population had highest abundance of *Prevotellaceae*	•Diet types associated with location were found to explain the largest variation in the taxonomic diversity between the participants • Northern Central Indians typically consume more carbohydrate fat rich diets (vegetable-based foods, grains, dairy products, trans fat foods (frequent frying of food)

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Dhakan, et al. 2019, (165).					•In comparison to other cohort Indian microbiome was less diverse than other populations •Indian microbiome was enriched with carbohydrate and energy metabolism •Two enterotypes were identified •Enterotype 1: *Prevotella* (n=53: Northern central Indians) •Enterotype 2: *Bacteroides* (n=57: Southern Indians) •Variation was not attributable to BMI •Enterotype 1: higher in *Prevotella* and Megasphaera including *P. copri and P. stercore* •Enterotype 2: *Ruminococcus* and *Faecalibacterium* higher relative abundance including *F.prausnitzii* and *Ruminococcus bromii, Akkermansia muciniphila, Eubacterium siraeum*, and *Roseburia hominis*	•Southern Indians consume an omnivorous diet. Meals consist mainly of rice, and animal products (sea food, red and white meat)
Turnbaugh et al. 2009 (148)	•American adults (n=154) •n=31 monozygotic twin pairs • n=23 dizygotic twin pairs •n=46 mothers •Participants part of longitudinal Missouri Adolescent •Female Twin Study (MOAFTS) cohort	•16S rRNA sequencing	•Each participant reportedly completed a dietary and lifestyle questionnaire	•All twins were 25-32years •Ethnicity: European or African •Twins were concordant for obesity or leanness	•Individuals from the same family had a more similar microbiota in comparison to unrelated individuals – this was not related to BMI •Obesity was associated with a decrease in diversity (Shannon) •Higher *Actinobacteria* and lower proportion in *Bacteroidetes* in obese individuals compared to lean individuals •Higher diversity was associated with higher relative abundance of *Bacteroidetes* •Microbiotas with higher relative abundance of *Firmicutes/ Actinobacteria* had lower diversity	•Not discussed

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Turnbaugh et al. 2009 (148)					•Presence of 'core' functional groups (involved in transcription and translation) were consistent across samples (e.g., pathways linked to carbohydrate and amino acid metabolism) •The obese groups microbiota were enriched for functional gene groups involved in processing of carbohydrates •75 % of obesity enriched genes were from *Actinobacteria* (lean had 0 %) and 25 % from *Firmicutes* •42 % of lean enriched genes were from *Bacteroidetes* (obese had 0 %)	

2.6.2. Nutrients and foods

Dietary fibre is a key driver of the microbiota, and numerous studies have explored the association of dietary fibre intake and the microbiota. In a population of women (n=1632) from the twins UK cohort, Menni et al. (121) assessed dietary fibre intake with a validated FFQ administered at baseline and at the nine year follow up. Higher dietary fibre intake was positively associated with microbiota diversity, and inversely with weight gain over the follow up period (Table 2.3). However, the microbiota was only assessed at follow up which does not enable conclusions to be drawn as to whether diversity is a cause or consequence of weight gain (121). In two cohorts of older American adults (n=151) Lin et al. (179) observed that dietary fibre from fruit and vegetables was associated with an increased relative abundance of the genera *Clostridia*, and fibre intake from beans (not further defined) was associated with a higher abundance of phyla *Actinobacteria*. Participants in both cohorts were control subjects from hospital-based studies which employed different dietary assessments, used different sequencing platforms, and collected data in different decades (see Table 2.3). Thus, although there is agreement with dietary fibre intake and taxa abundance across these cohorts, results may be contextual within these convenience samples who may not necessarily be reflective of the general population or 'healthy' per se and need to be interpreted carefully.

Table 2.3. Observational studies exploring the association of nutrient and food intake and gut microbiota composition

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Nutrient intake in cross-sectional studies						
Wu et al. 2011 (137)	•Healthy males and females •COMBO study: n=98 (2y-50y) •Males (n=43), females (n=55)	•16S rDNA sequencing	•3x 24h recalls during one week before (1x one day before) stool collection and FFQ (timeframe 1year) •Measured: recent (24hour recalls), usual (FFQ)	•BMI between 18.5 -35kg/m2	•BMI was the most strongly associated with microbiota composition •Identified 2 enterotypes *Bacteroides* and *Prevotella* dominate •Boundaries between enterotypes was not sharply defined •Vegetarians showed enrichment with *Prevotella*	•Long-term dietary intake was associated with enterotype composition •*Bacteroides* was associated with animal protein & Saturated fat (SFA) intake • *Prevotella* was associated with carbohydrate & simple sugar intake but was inversely associated with animal protein and SFA
Menni et al. 2017 (121).	•Healthy participants •Caucasian females (n=1632)	•16S rRNA sequencing •Microbiota analysis done at follow up (no baseline microbiota data)	•Validated 131 item FFQ used to estimate energy; and dietary fibre intake •Dietary fibre and saturated fat intake adjusted for energy intake •Dietary fibre assessed as total NSP (Englyst method) •Dietary assessment completed at baseline	•Twins UK cohort •Weight and BMI assessed 9 years apart •Baseline: Age: Mean: ~50years (±8.9) •Baseline: BMI: 25 kg/m2 (±4.2) •Follow up Age: Mean: ~59years (±9.2) •Follow up: BMI: 26 kg/m2 (±4.6) •Fibre intake was 20.4g/day	•59 % of the weight gain over time was not explained by genetic heritability in this population •Alpha diversity (Shannon Index, Simpson and OTU's) was lower in the weight gain group, despite having similar BMI's at baseline •Dietary fibre intake was positively associated with diversity •Higher diversity was associated with lower weight gain over time •No association between protein intake and microbiome diversity •The family *Ruminococcaceae* was positively associated with diversity and lower risk of weight gain • *Bacteroides* was negatively associated with diversity and higher risk of weight gain. •*Faecalibacterium prausnitzii* was positively correlated with lower BMI	•Dietary fibre intake was negatively associated with weight gain (even after adjusting for saturated fat intake) •For individuals with higher diversity, dietary fibre intake was associated with lower weight gain

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Lin et al. 2018 (179)	•Two cohorts of American adults (n=151) •NCI: n=75 (73.3 % male, and 26.7 % female) •New York University (NYU): n=76 (51.3 % male and 48.7 % female)	•16S rRNA sequencing •Microbiota data analysed separately as 2 different sequencing platforms were used. •NCI: 454 Roche FLX titanium prosequencing system •NYU: used Illumina MiSeq platform	•Self-administered semi-quantitative FFQ •Different FFQs used within each cohort •Energy intakes ≤2092 kJ and ≥16736 kJ were excluded from data analysis (n=151 within these cut offs) •Dietary data were adjusted for energy intake •Source of dietary fibre classified into 4 food groups	• In NCI & NYU: >80 % of participants identified as "white" •NCI: enrolled 1985-89. •NYU: Enrolled 2012-14	•In NCI & NYU: higher dietary fibre intake was associated with lower abundance of *Actinomyces, Odoribacter, Oscillispira. Eubacterium dolichum* and *Bacteroides uniformis* •Higher dietary fibre intake was associated with higher abundance of *Clostridia: SMB53, Lachnospira* and *Faecalibacterium prausnitzii.* •In NCI & NYU: Higher dietary intake of fibre from fruit and vegetables was associated with lower abundance of *Actinomyces, Odoribacter* and *Oscillispira.* •In NCI & NYU: Higher intake of 'beans' was associated with higher abundance of *Faecalibacterium prausnitzii* and lower abundance of *Bacteroides uniformis*	•Study specific quartiles of dietary fibre intake were used for NCI and NYU studies, and for food group sources of dietary fibre • Fibre intake from grains was not associated with taxon abundance •Higher fibre intake was associated with increased abundance of *Proteobacteria* in NCI (not in NYU) •Fibre intakes varied between the cohorts •NCI: Q1: <11·3, Q2: 11·3–14·3, Q3: 14·4–16·6, Q4: ≥16·7 g/day •NYU: Q1: <19·8, Q2: 19·8–23·7, Q3: 23·8–31·6, Q4: ≥31·7 g/day).
Koo et al. 2019 (180).	•Chinese (n=14), Malay (n=10) and Indian (n=11) adults living in Singapore	•16S rRNA sequencing •Participants were stratified by central obesity vs no central obesity	•Advised by dietitian how to complete 3-day food record (included 1 weekend day) •Mean daily macronutrient intakes were calculated from food record	•Males n=23, females n=12 •Age: median 39 (range: 22-70 years) •BF% assessed with Tanita Body Fat Composition Analyser •63 % had central obesity: defined by WHO waist circumference cut offs for Asian males (>90 cm) and females (>80 cm) •Central obesity: BMI: 29.2 kg/m^2, BF%: 29.9 ± 11.3 % •No central obesity: BMI: 21.6 kg/m^2, BF% 23.7 ±7 •Inflammatory biomarkers were assessed – no relation to central obesity	•No significant differences in Shannon diversity when stratifying the population by ethnicity or central obesity groups •Relative abundances of *Anaerofilum, Gemellaceae, Streptococcaceae,* and *Rikenellaceae* were significantly lower in the central obesity group •No difference in *Prevotella* to *Bacteroidetes* ratio between central obesity vs. non central obesity group •No significant differences between BMI groups as assessed by Bray Curtis at the genus level	•Participants from both groups were all consuming high fat >30 % (high SFA>10 %) and low dietary fibre <10 g/day intake •Central obesity group consumed more sugar, there was no other differences in diet intake between the groups

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Food intake						
Falony et al. 2016 (134).	•Healthy males and females •Total n= 3948 •Belgian Flemish Gut Flora Project (FGFP n=1106) •LifeLines-DEEP study (n=1135) (161)	•2 independent cohorts •The FGFP was analysed and the lifelines-DEEP used to validate findings. •16S rRNA gene sequencing •FGFP: GP collected anthropometric, health and lifestyle information Stool samples were encouraged to be the day of GP visit	•FGFP – limited information reported regarding how diet information was collected	•FGFP: •Age-median 53 years (range 19-85). •BMI-median: 24 kg/m^2 (range: 16-52) •CRP, eGFR and HOMA-IR assessed	•FGFP identified 3 enterotypes characterised by *Ruminococcaceae, Bacteroides* and *Prevotella* •Medication was found to have the largest effect of explaining the microbiota community composition •Stool consistency (assessed with the Bristol stool chart score) showed the largest effect size in FGFP •Independent of gender gene richness correlated positively with age •BMI showed a small effect on microbiome composition •*Methanobrevibacter* populated a 'cluster' which were predominately female, with lower weight and whom had longer transit time. – This cluster had higher gene richness and was overrepresented by *Ruminococcaceae* enterotype •The second cluster was dominated by *Bacteroides* enterotype and was characterized by lower microbial diversity. •*Prevotella* enterotype was associated with looser stools •In FGFP mode of delivery and infant nutrition was not associated with adult microbiome profiles in this cohort	•Dietary intake reported in frequencies "fruits frequency past week" "bread type preference" "magnesium supplement" and "coffee days since consumed" were all significant effect sizes •Enterotype 2/*Bacteroides:* had a preference for low fibre white bread

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Partula et al. 2019 (160)	•Healthy French adults (n=862) •Males n=445 •Females n=417	•16S rRNA sequencing •Faecal samples were collected 24 hours prior to study visit.	•Habitual diet was assessed with a 19 item FFQ •Researchers administered the FFQ •Diet and microbiota diversity analysis were adjusted for covariates: BMI, age, sex, smoking status, and physical activity levels.	•Age: 20-69 years •BMI: 24.3 kg/m^2 (± 3.3) •Physical activity: 5.5 (± 6.2) hours/week	•Microbiota composition demonstrated the largest difference at the phylum level •Firmicutes/Bacteroides ratio ranged from 0.23 to 78.6 (3.45 + 4.29, mean + SD)	•Cheese was negatively correlated with *Verrucomicrobia* and *Akkermansia muciniphila* •Sweet products were negatively correlated *Proteobacteria* •Fried products, sodas and SSB were all negatively associated with alpha diversity (Simpson, OTUs, and Chao1) •Fish and raw fruit were positively associated with alpha diversity •Beta diversity (Bray-Curtis) was only associated with raw fruit •Jaccard was associated with cheese, ready cooked meals, cooked fruit, raw fruits and fried products
Johnson et al. 2019 (138)	•Healthy American adults (n=34) •Males: n=14 •Females: n=20	•Metagenomic shotgun sequencing •Participants provided daily faecal samples for the 17-day period	•Participants self-recorded dietary intake for 17 consecutive days online with the automated self-administered 24-hour (ASA24) tool– thus they were collected as “diet records” •Dietitians instructed participants how to complete diet records •Dietary outliers (n=12 diet records) were excluded based on ASA24 recommendations of comparing energy intake to TEE	•Age: 31 ± 10 •Metabolic biomarkers captured and reported (e.g. lipid profile, insulin and glucose) •Weight, height and waist circumference reported. •Most participants identified as non-Hispanic Whites, Asian (n=5), African (n=1) •n=1 participant had T2DM (taking metformin) •TEE was estimated with Miffin-St. Jeor equation and self-reported energy intake and physical activity levels.	•Microbiota composition was more variable in some subjects than others across the study •Authors reported no significant differences (for all tests) between intervention arms and microbiota •Gender, BMI and age explained 34 % of differences in community structure •The daily variation of the microbiota depends on multiple days of dietary intake •Recent diet history and current microbiota can be used to predict the microbiota state on the next day •Day to day changes in the microbiota composition was similar between some subjects. Which led to the hypothesis that stability of the microbiota is shaped by community membership, rather than the stability of the diet	• Two subjects consumed meal replacement shakes throughout the study •Compliance to the intervention was assessed with self-report •Using the unweighted Unifrac food-based distances (from trees) average food intake corresponded with microbiota composition •When comparing ‘average’ microbiota and averages of dietary information (nutrients) they did not find any association. The authors stated this provided support for using tree-based approaches in dietary analysis with the microbiota •Using unweighted tree-based approach diet accounted for 44 % of the variation in microbiota composition •Individuals whom consumed fibre from similar fruit and grains tended to have similar microbiota composition

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Johnson et al. 2019 (138)			Diet was explored in terms of reported food choices to account for episodically unique foods. •Explanation of food categorization •They applied a tree based alpha and beta diversity methodology to the dietary information. Per participant they calculated a tree-based diversity of total foods consumed and food-based sources of nutrients		•*Dakarella massiliensis, Prevotella copri*, and *Dorea longicatena* were negatively associated with microbiota stability •*Bacteroides uniformis Alistipes sp.CHKCI003, Clostridium phoceensis, Alistipes onderdonkii*, and unclassified genus *Alistipes* were positively associated with microbiota stability •Dietary diversity did not affect microbiota diversity (when assessed with traditional methods). Diet diversity, assessed with faiths phylogenetic diversity, was associated with more stable microbiota	•Grain based foods were positively associated with *Lachnospiraceae* for multiple participants •Meat based foods were negatively associated with *Bacteroidaceae* for multiple participants •For one participant high intake of vegetables was associated with multiple species in the family of *Bifidobacteriaceae* •Food and microbiota interactions were highly personalized
Food and nutrient intake in cross-sectional studies						
Zhernakova et al. 2016 (161).	•LifeLines-DEEP study: n=1135 (males n=474, females n=661)	•16S rRNA sequencing results compared to Metagenomic shotgun sequencing results from same participants •Stool frequency measured with Bristol stool chart	•Validated 183 item FFQ which covered a period of 4 weeks. •Estimated nutrient intakes from FFQ	•Dutch adults – both parents are born in the Netherlands •50 % of participants never smoked •Age: Men 44 (± 19) years, Women: 43 (± 14) years •BMI: Men 25 (± 3.5) kg/m^2, Women: 25 (± 4.7) kg/m^2 •Molecular measurements assessed from multiple sites (e.g., plasma, faecal, whole blood)	•Chromogranin A (CgA) showed the strongest association with distance (Bray Curtis), diversity (Shannon) and gene richness and with the abundance •High levels of faecal CgA was associated with high calprotectin, TGS, stool frequency, IBS symptoms •Lower levels of CgA were associated with higher diversity, functional richness, higher HDL and fruit and vegetable intake. •Fecal calprotectin was positively correlated with age, BMI, and HBA1c, and negatively correlated with consumption of vegetables, plant proteins, chocolate and bread •Higher BMI was associated with lower abundance of *Alistipes finegoldii* and *Alistipes senegalensis*	•Consuming buttermilk was associated with higher diversity and *Leuconostoc mesenteroides* and *Lactococus lactis* •Consuming SSB was negatively associated with alpha diversity •Consuming coffee tea and wine was associated with higher alpha diversity. •Red wine consumption correlated with *Faecalibacterium prausnitzii* abundance •Higher energy (kcal/day), full fat milk and total carbohydrate (CHO) intake, including snacking behaviour were associated with lower alpha diversity •Consuming fruit and fish was associated with higher alpha diversity •Total CHO intake was positively associated with *Bifidobacteria,* but negatively with *Lactobacillis, Streptococcus* and *Roseburia* species. A low carbohydrate diet showed an opposite association for these species

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Zhernakova et al. 2016, (161)					•Age and gender were correlated with distance, diversity and richness. Women had higher richness and richness increased with age •*Methanobrevibacter smithii* was significantly associated with CgA •Self-reported IBS was associated with lower microbial diversity •*Alistipes shahii* was associated with lower TG levels, and higher species richness	•No association was identified with CHO intake and *Prevotella* species •Higher abundance of *Alistipes shahii* was associated with higher fruit intake
Tang et al. 2019 (159)	•Healthy adults: n=136 males (n=55) and females (n=95) aged 18-50years.	•16S rRNA oral and gut microbiome •Metabolomic profiling of plasma and stool samples (n=75) •Fecal samples were collected the 24 hours prior to the study visit	•Short term: Validated 3-day food record was completed prior to the study. Including 1 weekend day and the day before the study •Long term: 134 item semi quantitative FFQ covered a period of 1 year (NCI DHQ I) •Dietary data final data set consisted of 91 long term and 82 short term dietary variables •Dietary variables were adjusted for energy intake and gender	•Recruited from Pennsylvania in 2012-2014 •Age: ~20 (± 8) years •BMI: 25.4 kg/m^2 (± 5.4) •Ethnicity: White (54 %), Black (25 %), Asian (13 %), Other (9 %)	•812 metabolites were identified in the plasma and 770 in stool samples •Microbiome classified into 11 phyla, 20 classes, 21 orders, 32 families and 130 genera •No global associations were identified between diet, the gut microbiome, or metabolome and demographic data (age, sex, ethnicity, or BMI) •Two enterotypes were identified; Enterotype 1 (n=54), and enterotype 2 (n= 82). There was no difference in age or ethnicity across enterotypes. Enterotype 2 had lower BMI and significantly higher abundance of *Ruminococcaceae* •112 plasma and 122 stool metabolites were significantly different between the enterotypes •Plasma metabolites that differed by enterotype were significantly enriched for amino acids •Stool metabolites that were different between the enterotypes were enriched in vitamin B3 metabolism	•Long term diet was more strongly associated with the gut microbiota compared to short term diet •Short term diet was more strongly associated with the gut and plasma metabolome. •61 long term nutrients were associated with at least one genus •Several nutrients from plant based and dairy products were associated with three or more genera •Enterotype 1 consumed more alcohol and total cholesterol. There were no other dietary differences, between enterotype 1 and 2. •Dietary variables derived from plant-based foods were significantly associated with microbiome composition (enterotype) and exhibited microbiome mediated relationships with metabolites •Higher consumption of processed foods and lower intake of plant-based foods influenced the gut microbiota and metabolite products

In a small pilot sample (n=35) of adults living in Singapore, Koo et al. (180) explored the association between dietary intake and central obesity in relation to the microbiota. Three-day food records were completed by participants, and the only difference in dietary intake between participants (without and with central obesity), was for the central obesity group to consume more total sugar. Adding to this, there was no difference in microbiota diversity between the groups or when stratifying the population by ethnicity (180). Food records are considered the gold standard assessment method for dietary intake, and the reason no association was observed between diet and the microbiota may be related to the composition of the diet rather than the assessment itself. All participants were reportedly consuming a high fat (>30 % of total energy intake) and low fibre diet (<10 g/day); therefore, the low fibre intake and similar diet composition, including the small sample size, could collectively reduce the likelihood of identifying any associations.

Large scale observational studies have the power to detect associations but, have rarely been conducted in healthy populations. Two recent and large population studies (134,161) provided insight into the relationship between dietary intake, microbiota composition, and health outcomes. Stool consistency (assessed with the Bristol Stool chart) had the largest effect size on the microbiota in Belgian participants (n=1106) (134) which is unsurprising considering that colonic transit time affects time for bacterial replication (e.g., shorter times favouring bacteria with longer doubling times and vice versa), and diet composition directly influences stool consistency and frequency. Diet was proposed to explain ~6 % of the variation in the microbiota. However, despite reporting some broad diet information (e.g., fruits frequency past week, bread type preference), the assessment methodology and details (e.g., serving sizes, composition of food groups) were not reported, making it challenging to interpret the diet-microbiota results reported (134) (Table 2.3). Zhernakova et al. (161) did report that the intake of specific nutrients and food groups was associated with diversity and specific taxa abundance (summarised in Table 2.3). For example, higher energy intake, SSB, and full fat milk consumption were associated with lower diversity in Dutch adults (n= 1135). Again, these associations need to be interpreted carefully, as there is limited detail reported regarding the dietary assessment and analysis methodology (161).

In a large population of healthy French adults (n = 862), Partula et al. (160), administered a 19-item FFQ to assess habitual intake in relation to the microbiota, and analysis was adjusted for covariates which could either affect diet or microbiota composition (e.g., age, sex, BMI, physical activity levels). Partula et al. reported that the intake of healthy foods was positively associated with alpha diversity whereas unhealthy foods were negatively associated (160). However, the qualitative FFQ was short in duration, and consisted of a small number of broad food groups which resulted in high variability within groups. Thus, again these results need to be interpreted accordingly.

Multiple 24-hour recalls and food records are considered more accurate (in capturing dietary data) in comparison to FFQs, which are prone to higher rates of misreporting (85). A recent intensive study by Johnson et al. (138) assessed the dietary intake of 34 individuals consecutively over a 17-day period (coupled with daily metagenomic shotgun sequencing). Participants self-recorded their daily intake (as prospective food records) online with the automated self-administered 24-hour (ASA24) assessment tool developed by the NCI. Utilising the food records, researchers explored the association of habitual diet and microbiota using novel approaches to handle dietary data (similar to phylogenetic trees used in microbiota analyses), enabling focus on episodic and individually consumed food choices. The authors observed individualised responsiveness of the microbiota to dietary intake (138) as summarised in Table 2.3 and proposing that individualised responsiveness to dietary components may hinder cross-sectional studies capacity to detect diet and microbiota relationships. Reportedly, conventional approaches to dietary analysis in their study (e.g., averages of macro- and micronutrient intakes) were inadequate for exploring and identifying diet and microbiota relationships. Johnson et al. (138) concluded that multiple days of recent diet intake had the strongest influence on the composition of the microbiota and, impressively, the dietary assessment methodology is thoroughly explained (e.g., exclusion of misreporters, collapsing of food groups), improving the transparency of results. However, the novel dietary analysis method needs to be validated, and the intensive nature of this study makes it challenging (e.g., cost, participant/researcher burden) to reproduce at a large scale (to increase the power of observed associations).

2.6.3. Dietary patterns

The analyses of dietary patterns enable the assessment of the broader diet, and how foods and nutrients are consumed together. As mentioned above, employing multiple dietary assessment measures can improve the quality and validity of dietary data captured. However, only a few studies use multiple dietary assessment methods, such as food records, FFQs and 24hour recalls (137,159,177,181). combining FFQs and food records to assess *a priori* dietary patterns has enabled the clear association of diversity in the diet and subsequent diversity of the microbiota to be observed in elderly Irish (164) and Italian adults (170) (Table 2.4).

2.6.4. *A priori* dietary patterns

Due to the interest in exploring the association between dietary intake and microbiota composition, Bowyer et al. (182) validated and explored which dietary pattern indices (HEI, Healthy Food Diversity Index (HFD) or the Mediterranean diet) would explain the most variance in microbiota composition in a subsample of the UK twins cohort (n=2070). A validated FFQ was used as the basis of the dietary information, and details were provided with how they scored the dietary data from the FFQ against these indices. The HEI was positively associated with alpha diversity, whereas the HFD was negatively associated with alpha diversity. Bowyer et al., emphasised the importance of using a validated dietary assessment tool in future diet-microbiota studies and to ensure that any index used is relevant to the dietary practices of the population it is applied to (182). Thus, although the HEI explained the most variance in microbiota composition for this population of older European females, the HEI may not be generalisable to other global populations with different dietary habits.

Table 2.4. Observational studies exploring the association of *a priori* and *a posteriori* dietary patterns and gut microbiota composition

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
***A priori* dietary patterns in cross-sectional studies**						
De Filippis et al. 2016 (170)	•Healthy Italian male and female adults (n=153) •Vegetarian (VG: n=51), vegan (V: n=51) and omnivore (O: n=51)	•16S rRNA gene sequencing •3 faecal and urine samples were collected on the same day across 3 weeks	•7 day weighed food record •*A priori* dietary pattern: Adherence to Italian Mediterranean diet index (IMD) •*A priori* dietary pattern: Adherence to Healthy Food Diversity (HFD) index •Self-reported adherence to diet pattern (V, VG and O) for at least 1 year •Nutrient intakes normalized /1000kcal	•Vegetarians (33 females, 18 males and, age 39 ± 9 years, BMI 21.9 ± 2.5 kg/m^2) •Vegans (28 females, 23 males, aged 37 ± 10 years, BMI 21.3 ± 2.2 kg/m^2) •Omnivores (28 females and 23 males, aged 37 ± 9 years, BMI 22.1 ± 2.0 kg/m^2)	•Vegans and vegetarians had higher production of SCFA, and animal-based diet (omnivore) had lower SCFA production •Bacteroidetes phylum was more abundant in vegetarians & vegans, in comparison to omnivores •Higher *Firmicutes/Bacteroidetes* ratio associated with omnivores •*Lachnospira* was significantly linked to highest, and *Prevotella* the medium quartile of HFD index, and were associated with plant-based diets •Omnivore diet was positively associated with *L-Ruminococcus* •Stratification by self-reported diet did not result in clear differentiation in microbiota taxa	•Adherence to the MD and HFD were correlated (r=0.61) •Consumption of fibre, fruit, vegetables and legumes, increased the abundance of faecal SCFA •Median fibre intakes were 25 g/day for omnivores, 35 g/day for vegetarians and 48 g/day for vegans •Population had higher median intake of starch and fibre in comparison to US cohort of vegans and omnivores •Habitual vegetarian and vegan diets promote enrichment of fibre degrading microbiota •Lower adherence to IMD associated with higher urinary TMAO levels
Claesson et al. 2012 (164)	•Elderly Irish adults (n=178) •Males and females	•16S rRNA gene sequencing	•Validated semi quantitative 147 item FFQ. Adapted from European Prospective Investigation into Cancer (EPIC) study and validated to use in Irish population. •Intake of foods weighted by 10 consumption frequency categories •Diversity of the diet calculated for each participant using FFQ information and HFD index	•Elderly adults Age: 78 ± 8 years (64-102 years) •Community-dwelling, *n* = 83 (43 % male, BMI: 27.5 (± 5 kg/m^2) •Out-patient day hospital, *n* = 20 (65 % male, BMI: 28.6 (± 6 kg/m^2); •Short-term hospital care (rehabilitation < 6 weeks), *n* = 15 (53 % male, BMI: 30.6 (± 5 kg/m^2)	•Weighted and unweighted Unifrac PCoA of OTU's showed clear separation of community dwelling and long stay adults •Microbiota from younger adults clustered with those community dwelling older adults •Microbiota was associated with duration of long stay care. Longer duration of stay further reduced the similarity of microbiota from community dwelling subjects •No relation between *Firmicutes/Bacteridetes* ratio and BMI •Microbiota did not differ between males' females after adjustments for location and age	•HFD index correlated with alpha diversity •Higher dietary diversity was associated with higher alpha diversity (OTU's, Shannon Index and Phylogenetic diversity) and vice versa. •Identified 4 dietary groups DG1: low fat/high fibre, DG2: moderate fat/high fibre, DG3: moderate fat/low fibre and DG4: high fat/low fibre •DG1 and DG2 predominately community dwelling subjects and scored higher on HFD •DG3 and DG4 predominately long stay subjects and had least diverse diet and microbiota

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Claesson et al. 2012 (164)			HFD index	•Long-term residential care *n* = 60 (23 % male, BMI: 23.5 (± 5 kg/m^2). • n=13 young adults were included (36 ± 6 years) •Non fasted blood samples collected •Long stay adults scored worse on physiological and psychological assessments	•Microbiota – health associations were more evident in long stay subjects •Increased fragility and poorer health (adjusted for gender, age, and location) was the major determinates which differentiated the microbiota from community versus long stay patients (as location determines diet) •Identified 2 enterotypes: enterotype 1: *Bacteroides* and E2: *Prevotella* •Community dwelling subjects were more likely to be in *Prevotella* enterotype	•DG1: had more diverse diet (assessed with HFD) and microbiota (alpha) •Longer stay patients whom had been in care longer diets differed from those recently admitted •Diet changed quicker than microbiota composition. Diets changed to long stay diet after 1 month of care, but microbiota took 1 year to change to long stay microbiota.
Mitsou et al. 2017 (183)	•Healthy participants •n=116 (male: n=61, and female: n=55)	•16S rRNA sequencing	•Adherence to Mediterranean diet assessed with 11 item MedDietScore • Validated semi-quantitative FFQ conducted with a researcher •13 food groups created •Low energy reports = 13.8 % of the population (based on Schofield and Goldberg EI:BMR) n=100 with valid dietary data.	•Recruited between 2011- 2015 from Athens, Greece. •Age: ~41 (± 13) years •BMI: 27.3 kg/m^2 (± 4.5) •Body Fat % assessed with bioelectrical impedance (BIA) •Physical activity assessed with validated questionnaires	• MedDietScore was positively associated with faecal moisture, *Bifodobacteria:eColi* ratio and *Bacteroides spp* • Consumption of snacks and junk food was negatively associated with the abundance of *Faecalibacterium prausnitzii*, *Lactobacillis* group and Bacteroides •Consumption of sodas was negatively associated with *Akkermansia muciniphila*. •Higher adherence to MedDiet was associated with detecting higher amount of SCFA and acetate •Higher adherence to Mediterranean diet associated with lower *Escherichia coli* and higher faecal water content and stool frequency	•Higher adherence to MedDietScore was associated with higher intake of starchy products, fruit, vegetables, fish and eggs. With lower intake of meat, snacks and coffee/tea/SSB •Study specific cut offs were used for high vs low MedDietScore (which could differ to other populations) •Consumption of snacks and junk food correlated negatively with MedDietScore •Fibre intake: Tertile 1: ~17 g/day (±6), Tertile 2: ~24 g/day (±8), Tertile 1: ~30 g/day (±10)
Gutiérrez-Díaz et al. 2017. (181)	•Healthy Spanish Adults (n=74)	•Faecal sample was provided after dietary assessment	•160 item semi quantitative FFQ covering a 1 year period of designed as *a priori* for this project, and was validated with a 24-hour recall	•Age: n=37 50-65years, n=37 >65years •BMI and physical activity levels reported	• Higher adherence (score) to the Mediterranean diet was associated with higher relative abundance of *Faecalibacterium prausnitzii* •Higher Mediterranean diet score was associated with higher faecal phenolic metabolites	•Detailed information on the flavonoid, phenolic acid and dietary fibre content reported of the foods assessed in the cohort

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Gutiérrez-Díaz et al. 2017. (181)		•Microbiota assessed with quantitative PCR and plate-count techniques	•Med Diet Score created for this project •Nutrient intake estimated from FFQ and 24H recall •The phenolic compound content of diet was assessed •Dietary fibre content broken down to insoluble and insoluble.		•No association with transit time and abundance of phenolic compounds in the faecal samples observed.	
Bowyer et al. 2018 (182)	•Older European Twins (n=2070) from Twins UK cohort	•16s rRNA (V4) sequencing •Illumina MiSeq platform used. •Models adjusted for age, BMI, Twin zygosity, sex and OTU count per sample	•Validated 131 item FFQ collapsed into 152 food groups. •Diet was assessed at different time points secondary to cohort design (~ 5years between each assessment) •HEI, Med Diet Score(MDS) and HFD index were validated and compared to each other •Details provided on how FFQ data was used to create indexes	•Age, BMI, smoking status and frailty assessed •For the participants with microbiota data (n=2070) •90 % were female •99 % were European •BMI: 25.9 ± 4.7 •Age: 60.5 ± 11.5	•n=2070 sub sample had microbiota data •Shannon diversity, chao1, Simpson and OTUs were assessed •Beta diversity showed that the HEI showed the largest variance in the microbiota composition •HEI and MDS were significantly associated with alpha diversity (Shannon, Simpson and OTUs) •Only the HFD was associated with Chao1 index •HFD was inversely associated with all alpha diversity metrics •HEI and MDS were negatively associated with the relative abundance of Ruminococcus, *Lachnospira* and *Actinomyces*	•Dietary indexes were validated against larger population (n=4428) then compared to microbiota composition •HEI was significantly associated with alpha diversity •HEI and MDS were inversely associated with BMI and frailty •HFD was positively associated with frailty index

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
***A posteriori* Dietary Patterns in cross-sectional studies**						
Shikany et al. 2019 (184)	American older men (n=517)	•16S rRNA sequencing •Stool sample completed 4 ± 12 days after completing FFQ	•Semiquantitative FFQ covering period of 1 year •FFQ completed at home – asked about 69 foods and 13 questions about food preparation •*A posteriori* dietary patterns derived with factor analysis •Dietary patterns divided into quartiles	•Community dwelling men >65 years •Age: 84 ± 4 years •Ethnicity: 88 % White (Non-Hispanic) •BMI 26.9 ± 3.7 kg/m^2	•Adherence to Western and prudent dietary patterns not associated with alpha diversity (Shannon Index and inverse Simpsons index) •Non-Hispanic White participants had higher alpha diversity compared to other ethnic groups. •No association with BMI and alpha diversity •Alpha diversity was different between study centers •After adjustment for age, ethnicity, energy intake, BMI, (study location and library size), adherence to the Western pattern was positively associated with the relative abundance of the genera *Eubacterium, Alistipes, Anaerotruncus, Collinsella, Coprobacillus, Desulfovibrio, Dorea*, and *Ruminococcus.* Relative abundance of *Coprococcus, Prevotella Haemophilus, Faecalibacterium, Lachnospira,* and *Paraprevotella*, were negatively associated with adherence to the Western pattern. •After the same adjustments (as above) adherence to the Prudent pattern was positively associated with the relative abundance of the genera *Veillonella Faecalibacterium, Paraprevotella* and *Lachnospira*, •Relative abundance of *Ruminococcus, Desulfovibrio, Dorea, Cloacibacillus, Collinsella*, and *Coprobacillus* were negatively associated with adherence to the Prudent pattern	•Identified 4 dietary patterns. •DP1: “Western” pattern (processed meats, refined grains, potatoes, eggs, sweets and salty snacks) •DP2: “prudent” pattern (vegetables, fruit, nuts, fish, skinless chicken and turkey) •Higher adherence to Western pattern was associated with higher BMI •Higher adherence to prudent pattern associated with lower BMI

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Oluwagbemigun. et al, 2019 (185)	•German adults (n=225) •Participants part of European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort	•16S rRNA sequencing •Serum samples were assessed with untargeted metabolomics	• Three 24-hour recalls (first collected at study visit and remaining two randomly over the phone) •Food intake collapsed into 39 food groups. •Treelet transformation analysis used for dietary pattern analysis (combines PCA and hierarchical clustering analysis) •Food groups standardised before dietary pattern analysis •Named the dietary patterns based on loadings >0.4 •Multivariate models were adjusted for sex, age, BMI, smoking status, education level, energy intake, occupation and relevant health issues	•53 % Women •Age: Median 63years [15] •BMI: 26.7 kg/m^2 [5] •~1/10 were smokers •~2/5 had university qualifications •Non fasted serum samples provided for untargeted metabolomics analysis	•Identified 7 bacterial patterns (BCP) with Treelet transformation analysis. Which explained 19.9 % overall variance in the 317 OTU's identified •BCP1: named "*Veillonellaceae, Comamonadaceae,* and Family XI-dominated" •BCP2: "*Erysipelotrichaceae, Coriobacteriaceae,* and *Lachnospiraceae*" •BCP3: "*Ruminococcaceae* dominated" •BCP4: "*Anaerovibrio,* Uncultured genus in family *Rhodospirillaceae*, and *Brachyspira* genera" •BCP5: "Prevotella 6, Ezakiella, and Porphyromonas genera" •BCP6: "*Butyrivibrio*, Unidentified genus in family uncultured organism, order NB1-n, and *Victivallis* genera" •BCP7: "*Enterobacteriaceae*." The "*Anaerovibrio*, Uncultured genus in family *Rhodospirillacea*e, and *Brachyspira* genera" and "*Prevotella 6, Ezakiella*, and *Porphyromonas* genera" •After adjustments for covariates, 1 SD increase in DP2 was associated with increase (β= 0.35) in amino acid serum metabolites •1 SD increase in DP3 was associated with increase (β= 0.45) in fatty acids serum metabolites •1 SD increase in DP4 was associated with decrease (β= -0.28) in amino acids serum metabolites	•Identified 5 dietary patterns which explained 17.7 % of the variance in food intake. •DP1: "alcohol and red meat" (spirits, beer, red meat and sugar sweetened beverages) •DP2: "Bread margarine and processed meat" (bread, margarine, confectionary and processed meats) •DP3: "fruit vegetables and vegetable oils" (Fruits, vegetables and vegetable oils) •DP4: "tea and miscellaneous" (Tea and condiments e.g. herbs, spices, artificial sweeteners) •DP5: "pasta rice and sauce" (pasta rice and sauces) •No dietary pattern significantly predicted a bacterial pattern •Dietary patterns are reflected in the patterns of serum metabolites stable

Ref.	Participants	Microbiota analysis	Dietary assessment	Host characteristics	Microbiota characteristics	Diet characteristics
Kong et al. 2014 (186)	•Healthy overweight and obese French adults n=45 (males n=6 and females n=39) •Healthy lean subjects (n=14) were used as a comparison for some aspects of the study	•16S rRNA sequencing real time quantitative PCR •Faecal samples were collected after a 12 hour fast before data collection	•Obese/overweight subjects completed a 7-day estimated food record and lean subjects completed a 3-day estimated food record •Dietitian reviewed the food records with participants •Food records were collapsed into 26 food groups •Cluster analysis used to identify dietary patterns	•BF% assessed with DXA •Fasting bloods assessed for insulin, glucose, lipids and inflammatory markers •Cluster 1 n=14, 86 % females, age: 34.4 ± 2.7 years •Cluster 2 n=18, 83 % females •Cluster 3 n=13, 92 % females, age: 52.2 ± 2.3 years •No difference in adiposity, body weight, HOMA-IR between the clusters •Cluster 1 was significantly younger than 2 and 3	•No difference between the 3 clusters identified and the 7 bacterial groups identified •Overweight/obese individuals had lower levels of *Clostridia leptum, Clostridia coccoides,* and *Bacteroides/Prevotella* groups in comparison to the lean subjects. •Stratifying the population by high gene count versus low gene count, cluster 3 had higher gene richness and diversity in comparison to the other clusters •Adjusting for age there was a significant correlation between intake of fruit and soups and bacterial gene counts	•Cluster 1 was characterised by higher intake of potatoes, sweets, SSB, and lower consumption of yoghurt, fruit and water •Cluster 3 was characterised by higher intake of fruit, yoghurt, soup, vegetables and lower intake of sweets and SSB •Cluster 2 was characterised by higher intake of water and yoghurt and considered in between healthfulness for cluster 1 and 3 •Cluster 3 was determined to be the 'healthier' cluster •Lean subjects were like the patterns observed in Cluster 3 – the 'healthier' pattern characterised by older females •There was no difference in energy intake across the clusters •Cluster 1 had lower fibre intake and cluster 3 had higher fibre intake

De Filippis et al. (170), analysed the dietary intake of a healthy population of Italian adults (n=153) in three ways (170). A 7-day weighed food record was used to explore *a priori* dietary patterns (the HFD and the Italian Mediterranean Index (IMD) (187)), and participants were also stratified by self-reported habitual adherence to diet category (vegan, vegetarian or omnivore). In comparison to omnivores, vegetarians and vegans had higher dietary fibre intake and IMD scores; however, like other studies (170,175–177), stratifying the population by self-reported diet type did not translate to differential abundance in the microbiota taxa. Higher adherence to the IMD was associated with higher faecal SCFA, and lower adherence to the IMD was associated with higher urinary trimethylamine oxide (TMAO). Higher TMAO is considered a risk factor for CVD (188), and TMAO levels were observed to be higher in omnivores. The microbiota metabolise choline and L-caritine consumed in the diet (e.g., eggs, beef, pork and fish) to trimethylamine (TMA), which is further oxidised to TMAO by the liver (189). *Anaerococcus hydrogenalis* and *Clostridium asparagiforme* are examples of bacterial species that can metabolise choline to TMA and encode genes (including choline-TMA lyase (cutC)) required for this specialised activity (190). The study by De Filippis et al. highlights the association of dietary intake and the microbiota with the potential to influence the metabolic health of the host, and is a strong example of how multiple methods of dietary assessment strengthen the quality and validity of the dietary data and observed associations.

Due to the composition of the Mediterranean diet and health benefits associated with higher adherence (94), the Mediterranean diet has often been explored in relation to the composition of the microbiota (170,181,183). Higher adherence to the Mediterranean diet has been associated with higher relative abundance of the species *Faecalibacterium prausnitzii* (181). *Faecalibacterium prausnitzii* is a butyrate producing bacteria essential for the health of colonocytes (120) and often thought of as a marker of health status of the host, as lower abundance of *Faecalibacterium prausnitzii* has been observed in IBD patients (191). In a population of Greek adults (n=116), higher adherence to a Mediterranean style diet was inversely associated with discretionary food intake (e.g., snacks, junk food and sugar sweetened beverages), and discretionary food intake was inversely associated with the relative abundance of *Faecalibacterium prausnitzii* (183). Thus, despite the heterogeneity in methodology across studies, similar themes are emerging where higher intake of core and

lower of discretionary foods appear to be associated with health of the host and diversity of the microbiota.

2.6.5. *A posteriori* dietary patterns

In a population of older American men (n=517) Shikany et al. (184) reportedly identified four patterns, but discussed only two. The Western pattern was characterised by: processed meats, refined grains, potatoes, eggs, and salty snacks; and the prudent pattern was characterised by fruit, vegetables, nuts, fish, skinless chicken, and turkey (184). Alpha diversity was not associated with adherence to either pattern, nor was BMI. However, BMI was positively associated with the Western pattern and negatively with the prudent pattern. In addition, after adjustments for co-variates the Western dietary pattern was inversely associated with numerous genera such as *Faecalibacterium*; in contrast to the prudent pattern which was positively associated with the abundance of the genera *Faecalibacterium* (184) (Table 2.4). The participants were part of a cross centre osteoporotic fractures risk study and over half of the participants indicated multimorbidity, thus increasing the likelihood that these observations are contextual. Further, *a posteriori* patterns are only generalisable to the study population, and the limited detail provided regarding analyses of the dietary patterns hinders clear conclusions to be drawn. However, the trend for *Faecalibacterium* to be positively associated with consuming core foods and negatively with the intake of discretionary foods across different populations is intriguing.

In a population of older German adults, (a convenience sample from the Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort), Oluwagbemigun et al. (185) identified five *a posteriori* dietary patterns based on three 24-hour recalls. The methodology of how they derived these dietary patterns is explained, and the characteristics (factor loadings) of the patterns were reported. Sensitivity analysis was conducted to explore the validity of patterns identified and all multivariate analyses were adjusted for covariates, thereby increasing the transparency of their dietary analysis. None of the dietary patterns predicted any of the seven bacterial patterns they identified, although the dietary patterns were associated with the serum metabolite profiles. In a small sample (n=45) of healthy overweight and obese French adults, Kong et al. (60) identified three dietary patterns with

cluster analysis. The healthier pattern, characterised by fruit, yoghurt, soup and lower intake of sugar sweetened foods and beverages, was associated with higher microbiota gene richness and diversity (60). There was no difference in adiposity, body weight, insulin resistance or energy intake between the clusters, and despite the healthier pattern consuming more dietary fibre there was no difference in taxa abundance between the clusters. The sample size may have been too small to observe an association with the dietary patterns and microbiota taxa; however, the inclusion of males and females within the clusters could have further obscured results due to potential differences in dietary intake and physiology. Overall, there appears to be limited adequately powered studies exploring *a posteriori* dietary patterns in relation to the microbiota of healthy, free-living younger populations.

Broadly summarising the three dietary assessment themes of the literature review, and without focusing on the heterogeneity between studies (i.e., participants, dietary assessment, or sequencing techniques). The evidence regarding diet-microbiota and health currently suggests that a lower BMI will be associated with higher alpha diversity (121), and that lower diversity will be associated with obesity (148). Higher dietary diversity (164) (e.g., higher adherence to the HEI (182) and Mediterranean diet (170)) and higher fibre intake (121) will also be positively associated with diversity. Consuming fish and raw fruit (161), coffee, tea, and wine (160) will be positively be associated with diversity; in contrast to consuming fried products, full fat milk, and SSB (161) will all be negatively associated with diversity.

Higher relative abundance of *Ruminococcaceae* will be associated with lower BMI (121,159), and higher adherence to a Western dietary pattern (184). Higher relative abundance of *Faecalibacterium prausnitzii* will be inversely associated with BMI (121), and positively associated with higher intake of fibre (179), red wine (161), and consuming over 30 plant-based foods (178), as well as higher adherence to a prudent dietary pattern (184). In contrast, consuming junk food and snacks (183) and higher adherence to a Western dietary pattern will be associated with higher BMI and lower relative abundance of *Faecalibacterium prausnitzii* (184). However, these observations are across different populations employing different analyses techniques and ideally need to be replicated in one setting.

2.7. Summary of the literature review

The causes of obesity are multifactorial and complex and the disparities in the prevalence of obesity are strongly associated with levels of deprivation. The rates of obesity continue to rise globally, and in NZ, one in three adults are obese. Obesity is associated with an increased risk of premature death and is a major public health concern. The obesogenic environment is driving a widespread increase in the consumption of inexpensive energy dense and nutrient poor foods, which are positively associated with increases in BMI. Diet is considered a key modifiable risk factor for NCDs such as obesity, and evidence from animal models suggest the microbiota may play a role in diet-induced obesity. Habitual diet appears to influence the composition of the microbiota, and dietary fibre intake has been consistently associated with microbiota composition. Most studies assess the composition of the faecal microbiota by analysis of 16S rRNA gene sequencing. However, to explore bacterial species abundances accurately, greater analytical depth is required. The majority of studies utilise a FFQ to assess habitual diet, and only a few smaller studies have employed food records, likely due to higher participant and researcher burden.

Typically, across studies the dietary assessment methodology and processing of dietary data are not accurately described, making it challenging to clearly interpret the significance of reported results. The appropriate controls and adjustments must be employed throughout the data collection and analysis process to reduce misreporting and, above all, methodology and data analyses should be clearly reported to increase transparency. Despite the inconsistencies in the analyses of diet and differences in sequencing methods across studies. The evidence suggests that higher intake of core foods and dietary fibre, would be associated with lower BMI and risk of obesity, as well as higher alpha diversity and relative abundance of *Faecalibacterium prausnitzii* and *Ruminococcaceae*. In contrast, consuming more discretionary foods, and lower fibre intake, will be associated with higher BMI, increased risk of obesity, lower diversity, and relative abundance of *Faecalibacterium prausnitzii*.

Further research is required to explore the diet-microbiota relationship, notably by conducting and reporting robust dietary assessment and improving the analytical depth of the microbiota analyses. Capturing multiple days of dietary intake can reduce misreporting and improve the

validity of the dietary data. Analysing the composition of the microbiota with metagenomic shotgun sequencing will improve analytical depth. Numerous studies have explored nutrient and food intake in relation to the microbiota but there are limited number of studies exploring how these dietary components are consumed together. Dietary pattern analysis considers the whole diet and may, in part, overcome the limitations of focusing on individual nutrients and foods. Statistical modelling of habitual dietary data is a valuable method, to reduce intra-individual variation of intake and improve the precision in the estimation of habitual intake, which needs to be explored in the context of diet-microbiota research. Exploring the diet-microbiota relationship in a healthy population whom have different metabolic disease risk, and body fat profiles, would provide insight into potential associations with body weight and metabolic health. In addition, ensuring factors which, can either influence diet and/or microbiota composition are assessed, and adjusted for in analyses, will improve the clarity of observations and further advance understanding.

In conclusion an extensive amount of research has been conducted exploring the association with habitual diet and microbiota. Despite this research, questions remain. Improving understanding of the diet-microbiota association in the context of health, may help guide therapeutic targets in the future. To advance understanding of the association between habitual diet, and the composition of the microbiota in relation to health outcomes, more research within one setting is required.

2.8. References

1. World Health Organization. Obesity and overweight [Internet]. 2018 [cited 2019 Apr 12].
2. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].
3. Organization for Economic Co-operation and Development (OECD). OECD Obesity update 2017. Paris, France; 2017.
4. Zhang Q, Wang Y. Socioeconomic inequality of obesity in the United States: Do gender, age, and ethnicity matter? Soc Sci Med. 2004 Mar;58(6):1171–80. doi:10.1016/S0277-9536(03)00288-0
5. The National Advisory Committee on Health and Disability. The social, cultural and economic determinants of health in New Zealand: Action to improve health a report from the National Advisory Committee on Health and Disability. Wellington; 1998.
6. Ljungvall Å, Zimmerman FJ. Bigger bodies: Long-term trends and disparities in obesity and body-mass index among U.S. adults, 1960-2008. Soc Sci Med [Internet]. 2012;75(1):109–19. doi:10.1016/j.socscimed.2012.03.003
7. The GBD Obesity Collaborators. Health effects of overweight and obesity in 195 countries over 25 years [Internet]. Vol. 377, New England Journal of Medicine. New England Journal of Medicine (NEJM/MMS); 2017. p. 13–27. doi:10.1056/nejmoa1614362
8. Kent S, Fusco F, Gray A, Jebb SA, Cairns BJ, Mihaylova B. Body mass index and healthcare costs: a systematic literature review of individual participant data studies. Obes Rev [Internet]. 2017;18(8):869–79. doi:10.1111/obr.12560
9. Lal A, Moodie M, Ashton T, Siahpush M, Swinburn B. Health care and lost productivity costs of overweight and obesity in New Zealand. Aust N Z J Public Heal [Internet]. 2012;36(6):550–6. doi:10.1111/j.1753-6405.2012.00931.x
10. King M, Smith A, Gracey M. Indigenous health part 2: The underlying causes of the health gap. Lancet [Internet]. 2009;374:76–85. doi:10.1016/S0140-6736(09)60827-8
11. Anderson I, Crengle S, Leialoha Kamaka M, Chen TH, Palafox N, Jackson-Pulver L. Indigenous health in Australia, New Zealand, and the Pacific. Lancet. 2006;367(9524):1775–85. doi:10.1016/S0140-6736(06)68773-4
12. Sheridan NF, Kenealy TW, Connolly MJ, Mahony F, Barber PA, Boyd MA, et al. Health equity in the New Zealand health care system: A national survey. Int J Equity Health. 2011;10:1–14. doi:10.1186/1475-9276-10-45
13. Williams WM. Racial inequities in cardiovascular disease in New Zealand. Divers Equal Heal Care. 2017;14(1):23–33. doi:10.21767/2049-5471.100088
14. Blakely T, Tobias M, Atkinson J, Yeh J-C, Huang K. Tracking Disparity: Trends in ethnic and socioeconomic inequalities in mortality, 1981-2004. Wellington; 2007.
15. Ministry of Health. Living well with diabetes: A plan for people at high risk of or living with diabetes 2015-2020. Wellington; 2015.
16. Ministry of Health. Quality standards for diabetes care toolkit. Wellington; 2014.
17. Statistics New Zealand and Ministry of Pacific Island Affairs. Health and Pacific peoples in New Zealand [Internet]. Health and Pacific peoples in New Zealand. Wellington; 2011.

18. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, et al. Obesity pathogenesis: An endocrine society scientific statement. Endocr Rev. 2017;38(4):267–96. doi:10.1210/er.2017-00111
19. Mechanick JI, Hurley DL, Garvey WT. Adiposity-based chronic disease as a new diagnostic term: The American association of clinical endocrinologists and American college of endocrinology position statement. Endocr Pract. 2017;23(3):372–8. doi:10.4158/EP161688.PS
20. Oliveros E, Somers VK, Sochor O, Goel K, Lopez-Jimenez F. The concept of normal weight obesity. Prog Cardiovasc Dis. 2014;56(4):426–33. doi:10.1016/j.pcad.2013.10.003
21. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y, et al. Abdominal visceral and subcutaneous adipose tissue compartments. Circulation. 2007;116(1):39–48. doi:10.1161/circulationaha.106.675355
22. Dickey RA, Bartuska DG, Bray GW, Callaway CW, Davidson ET, Feld S, et al. AACE/ACE position statement on the prevention, diagnosis, and treatment of obesity (1998 Revision) AACE/ACE Obesity Task Force. Endocr Pract. 1998;4(5):297–350.
23. Després J-P. Body fat distribution and risk of cardiovascular disease. Circulation. 2012;126(10):1301–13. doi:10.1161/circulationaha.111.067264
24. Kramer C, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions? A systematic review and meta-analysis. Ann Intern Med. 2013;159(11):758–69. doi:10.7326/0003-4819-159-11-201312030-00008
25. Ashwell M, Gunn P, Gibson S. Waist-to-height ratio is a better screening tool than waist circumference and BMI for adult cardiometabolic risk factors: Systematic review and meta-analysis. Obes Rev. 2012;13(3):275–86. doi:10.1111/j.1467-789X.2011.00952.x
26. Jo A, Mainous AG. Informational value of percent body fat with body mass index for the risk of abnormal blood glucose: A nationally representative cross-sectional study. BMJ Open. 2018;8(4):1–6. doi:10.1136/bmjopen-2017-019200
27. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001
28. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev Med (Baltim) [Internet]. 1999 Dec;29(6):563–70. doi:10.1006/pmed.1999.0585
29. 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 [Internet]. 2011;378(9793):804–14. doi:10.1016/S0140-6736(11)60813-1
30. Hopkins M, Blundell JE. Energy balance, body composition, sedentariness and appetite regulation: Pathways to obesity. Clin Sci. 2016;130(18):1615–28. doi:10.1042/CS20160006
31. Finlayson G, King N, Blundell JE. Liking vs. wanting food: Importance for human appetite control and weight regulation. Neurosci Biobehav Rev. 2007;31(7):987–1002. doi:10.1016/j.neubiorev.2007.03.004
32. Nestle M, Wing R, Birch L, DiSogra L, Drewnowski A, Middleton S, et al. Behavioral and social influences on food choice. Nutr Rev [Internet]. 2009 Apr 27;56(5):50–64. doi:10.1111/j.1753-4887.1998.tb01732.x
33. Kieffer TJ, Habener JF. The adipoinsular axis: effects of leptin on pancreatic B-cells. Am

J Physiol Metab. 2000;278:1–14.

34. Klok MD, Jakobsdottir S, Drent ML. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: A review. Obes Rev. 2007;8(1):21–34. doi:10.1111/j.1467-789X.2006.00270.x

35. D'souza A, Neumann UH, Glavas MM, Kieffer TJ. The glucoregulatory actions of leptin. Mol Metab [Internet]. 2017;6(9):1052–65. doi:10.1016/j.molmet.2017.04.011

36. 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 (80-) [Internet]. 2007 May 11;316(5826):889–94. doi:10.1126/science.1141634

37. Price RA, Li W, Zhao H. FTO gene SNPs associated with extreme obesity in cases, controls and extremely discordant sister pairs. BMC Med Genet. 2008;5:5–9. doi:10.1186/1471-2350-9-4

38. Livingstone KM, Celis-morales C, Papandonatos GD, Erar B, Florez JC, Jablonski KA, et al. FTO genotype and weight loss: Systematic review and meta-analysis of 9563 individual participant data from eight randomised controlled trials. BMJ. 2016;354:i4707. doi:10.1136/bmj.i4707

39. Wardle J, Carnell S, Haworth CM, Plomin R. Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. Am J Clin Nutr [Internet]. 2008 Feb 1;87(2):398–404. doi:10.1093/ajcn/87.2.398

40. Locke AEA, Kahali B, Berndt SIS, Justice AEA, Pers THT, Day FRF, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature [Internet]. 2015 Feb 11;518:197. doi:10.1038/nature14177

41. Silventoinen K, Rokholm B, Kaprio, J, Sørensen TIA. The genetic and environmental influences on childhood obesity: A systematic review of twin and adoption studies. Int J Obes [Internet]. 2010 Jan 15;34(1):29–40. doi:10.1038/ijo.2009.177

42. Stunkard AJ, Sørensen TIA, Hanis C, Teasdale TW, Chakraborty R, Schull WJ, et al. An adoption study of human obesity. N Engl J Med [Internet]. 1986 Jan 23;314(4):193–8. doi:10.1056/NEJM198601233140401

43. Wang T, Heianza Y, Sun D, Zheng Y, Huang T, Ma W, et al. Improving fruit and vegetable intake attenuates the genetic association with long-term weight gain. Am J Clin Nutr. 2019;1–10. doi:10.1093/ajcn/nqz166

44. 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 [Internet]. 2018;360:j5644. doi:10.1136/bmj.j5644

45. Qi Q, Chu AY, Kang JH, Huang J, Rose LM, Jensen MK, et al. Fried food consumption, genetic risk, and body mass index: Gene-diet interaction analysis in three US cohort studies. Br Med J [Internet]. 2014 Mar 19;348:g1610. doi:10.1136/bmj.g1610

46. Li S, Zhao JH, Luan J, Ekelund U, Luben RN, Khaw K-T, et al. Physical activity attenuates the genetic predisposition to obesity in 20,000 men and women from EPIC-Norfolk prospective population study. PLOS Med [Internet]. 2010 Aug 31;7(8):e1000332.

47. Martínez JA, Milagro FI, Claycombe KJ, Schalinske KL. Epigenetics in adipose tissue, obesity, weight. Adv Nutr. 2014;5:71–81. doi:10.3945/an.113.004705.71

48. Swinburn BA, Caterson I, Seidell J, James W. Diet, nutrition and the prevention of excess

weight gain and obesity. Public Health Nutr [Internet]. 2004 Feb 2;7(1a):123–46. doi:10.1079/PHN2003585

49. Swinburn BA, Kraak VI, Allender S, Atkins VJ, Baker PI, Bogard JR, et al. The global syndemic of obesity, undernutrition, and climate change: The lancet commission report. Lancet [Internet]. 2019 Feb;393(10173):791–846. doi:10.1016/S0140-6736(18)32822-8
50. Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK. American College of Sports Medicine Position Stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. Med Sci Sports Exerc. 2009;41(2):459–71.
51. World Health Organization. Physical activity and adults [Internet]. Global strategy on diet, physical activity and health. 2019 [cited 2019 Aug 19].
52. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, et al. Heart disease and stroke statistics - 2018 update: A report from the American Heart Association. Vol. 137, Circulation. 2018. doi:10.1161/CIR.0000000000000558
53. Swinburn B, Sacks G, Vandevijvere S, Kumanyika S, Lobstein T, Neal B, et al. INFORMAS (International Network for Food and Obesity/non-communicable diseases Research, Monitoring and Action Support): Overview and key principles. Obes Rev. 2013;14(October):1–12. doi:10.1111/obr.12087
54. Moodie R, Stuckler D, Monteiro C, Sheron N, Neal B, Thamarangsi T, et al. Profits and pandemics: Prevention of harmful effects of tobacco, alcohol, and ultra-processed food and drink industries. Lancet [Internet]. 2013;381(9867):670–9. doi:10.1016/S0140-6736(12)62089-3
55. Monteiro CA, Moubarac J, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. Obes Rev. 2013;14:21–8. doi:10.1111/obr.12107
56. Headey DD, Alderman HH. The relative caloric prices of healthy and unhealthy foods differ systematically across income levels and continents. J Nutr [Internet]. 2019 Nov 1;149(11):2020–33. doi:10.1093/jn/nxz158
57. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8
58. Vandevijvere S, Jaacks LM, Monteiro CA, Moubarac J, Girling-Butcher M, Lee AC, et al. Global trends in ultraprocessed food and drink product sales and their association with adult body mass index trajectories. Obes Rev [Internet]. 2019 Nov 17;20(S2):10–9. doi:10.1111/obr.12860
59. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr [Internet]. 2008 May 1;87(5):1107–17. doi:10.1093/ajcn/87.5.1107
60. Food and Agriculture Organization of the United Nations, Food Climate Reseach Network. Plates, pyramids, planet: Developments in national healthy and sustainable dietary guidelines: A state of play assessment. 2016. doi:ISBN 978-92-5-109222-4
61. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr. 2016;7(3):445–54. doi:10.3945/an.115.011718

62. Food and Agriculture Organization of the United Nations, World Health Organization. Preparation and use of food-based dietary guidelines. Nicosia, Cyprus; 1998.
63. Ministry of Health. Eating and activity guidelines for New Zealand adults [Internet]. Wellington; 2015.
64. Darmon N, Briend A, Drewnowski A. Energy-dense diets are associated with lower diet costs: A community study of French adults. Public Health Nutr. 2004;7(1):21–7. doi:10.1079/PHN2003512
65. Imamura F, Micha R, Khatibzadeh S, Fahimi S, Shi P, Powles J. Dietary quality among men and women in 187 countries in 1990 and 2010: A systematic assessment. Lancet. 2015;3:132–42. doi:10.1016/S2214-109X(14)70381-X
66. National Health and Medical Research Council. Australian Dietary Guidelines. Canberra: National Health and Medical Council; 2013.
67. Gregory E, Jackson PR, Chambers EC, McLaughlin ECG. New Zealand dietary studies. 1. Dietary studies of one hundred and seventy families of varying occupations. New Zeal Med Jour. 1943;42(228):64–70.
68. Amoah I, Cairncross C, Rush E. The battle for better nutrition: the role of the escalating fruit and vegetable prices. N Z Med J. 2017;130(1465):107–8.
69. University of Otago and Ministry of Health. A focus on nutrition: Key findings of the 2008/09 New Zealand adult nutrition survey [Internet]. Wellington: Ministry of Health; 2011. 1–332 p.
70. Lakerveld J, Ben Rebah M, Mackenbach JD, Charreire H, Compernolle S, Glonti K, et al. Obesity-related behaviours and BMI in five urban regions across Europe: Sampling design and results from the SPOTLIGHT cross-sectional survey. BMJ Open. 2015;5(10). doi:10.1136/bmjopen-2015-008505
71. Larson N, Story M. A review of environmental influences on food choices. Ann Behav Med. 2009;38(SUPPL.):56–73. doi:10.1007/s12160-009-9120-9
72. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45. doi:10.3945/jn.115.219634
73. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.
74. Schwingshackl L, Hoffmann G. Diet quality as assessed by the healthy eating index, the alternate healthy eating index, the dietary approaches to stop hypertension score, and health outcomes: A systematic review and meta-analysis of cohort studies. J Acad Nutr Diet [Internet]. 2015;115(5):780-800.e5. doi:10.1016/j.jand.2014.12.009
75. Schlesinger S, Neuenschwander M, Schwedhelm C, Hoffmann G, Bechthold A, Boeing H, et al. Food groups and risk of overweight, obesity, and weight gain: A systematic review and dose-response meta-analysis of prospective studies. Adv Nutr. 2019;10(2):205–18. doi:10.1093/advances/nmy092
76. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.
77. Lachat C, Nago E, Verstraeten R, Roberfroid D, Van Camp J, Kolsteren P. Eating out of home and its association with dietary intake: A systematic review of the evidence. Obes Rev [Internet]. 2012 Apr;13(4):329–46. doi:10.1111/j.1467-789X.2011.00953.x

78. Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake: Basal metabolic rate. A practical guide to its calculation, use and limitations. Int J Obes. 2000;24(9):1119–30. doi:10.1038/sj.ijo.0801376

79. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, et al. DIET@NET: Best practice guidelines for dietary assessment in health research. BMC Med [Internet]. 2017 Dec 15;15(1):202. doi:10.1186/s12916-017-0962-x

80. Goldberg GR, Black AE, Jebb SA, Cole TJ, Murgatroyd PR, Coward WA, et al. Critical evaluation of energy intake data using fundamental principles of energy physiology: 1. Derivation of cut-off limits to identify under-recording. Eur J Clin Nutr. 1991;45(12):569–81.

81. Kolonel LN, Henderson BE, Hankin JH, Nomura AMY, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: Baseline characteristics. Am J Epidemiol. 2000;151(4):346–57. doi:10.1093/oxfordjournals.aje.a010213

82. Chacko George G, Milani TJ, Hanss-Nuss H, Kim M, Freeland-Graves JH. Development and validation of a semi-quantitative food frequency questionnaire for young adult women in the southwestern United States. Nutr Res [Internet]. 2004 Jan;24(1):29–43. doi:10.1016/j.nutres.2003.09.006

83. Tooze JA, Freedman LS, Carroll RJ, Midthune D, Kipnis V. The impact of stratification by implausible energy reporting status on estimates of diet-health relationships. Biometrical J. 2016;58(6):1538–51. doi:10.1002/bimj.201500201

84. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr [Internet]. 1997 Apr 1;65(4):1220S-1228S. doi:10.1093/ajcn/65.4.1220S

85. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

86. Dodd KW, Guenther PM, Freedman LS, Subar AMYF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. J Am Diet Assoc. 2006;106(10):1640–50. doi:10.1016/j.jada.2006.07.011

87. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

88. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009;65(4):1003–10. doi:10.1111/j.1541-0420.2009.01223.x

89. Hu FB. Dietary pattern analysis: A new direction in nutritional epidemiology. Curr Opin Lipidol [Internet]. 2002 Feb;13(1):3–9. doi:10.1097/00041433-200202000-00002

90. Food and Nutrition Service United States Department of Agriculture. Healthy Eating Index (HEI) [Internet]. 2019 [cited 2019 Aug 12].

91. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. Nutr Rev. 2004;62(5):177–203. doi:10.1301/nr.2004.may.177

92. Moeller SM, Reedy J, Millen AE, Dixon LB, Newby PK, Tucker KL, et al. Dietary patterns:

Challenges and opportunities in dietary patterns research. J Am Diet Assoc. 2007;107(7):1233–9. doi:10.1016/j.jada.2007.03.014

93. Tucker KL. Dietary patterns, approaches, and multicultural perspective. Appl Physiol Nutr Metab. 2010;35:211–8. doi:10.1139/H10-010

94. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts. N Engl J Med [Internet]. 2018;378(25):e34. doi:10.1056/NEJMoa1800389

95. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhouser ML, et al. The dietary patterns methods project: synthesis of findings across cohorts and relevance to dietary guidance. J Nutr [Internet]. 2015;145(3):393–402. doi:10.3945/jn.114.205336

96. Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, et al. Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. Lancet [Internet]. 2019; doi:10.1016/s0140-6736(18)31788-4

97. Wang DD, Li Y, Afshin A, Springmann M, Mozaffarian D, Stampfer MJ, et al. Global improvement in dietary quality could lead to substantial reduction in premature death. J Nutr [Internet]. 2019;149(6):1065–74. doi:10.1093/jn/nxz010

98. Food and Nutrition Board: Institute of Medicine. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids (macronutrients). Washington, D.C; 2005.

99. National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. Nutrient reference values for Australia and New Zealand. Nutrition. Canberra; 2006.

100. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

101. Schwab U, Lauritzen L, Tholstrup T, Haldorssoni T, Riserus U, Uusitupa M, et al. Effect of the amount and type of dietary fat on cardiometabolic risk factors and risk of developing type 2 diabetes, cardiovascular diseases, and cancer: A systematic review. Food Nutr Res. 2014;58. doi:10.3402/fnr.v58.25145

102. Astrup A, Bertram HCS, Bonjour JP, De Groot LCP, De Oliveira Otto MC, Feeney EL, et al. WHO draft guidelines on dietary saturated and trans fatty acids: Time for a new approach? BMJ [Internet]. 2019;366(July):1–6. doi:10.1136/bmj.l4137

103. Satija A, Bhupathiraju SN, Spiegelman D, Chiuve SE, Manson JAE, Willett W, et al. Healthful and unhealthful plant-based diets and the risk of coronary heart disease in U.S. adults. J Am Coll Cardiol. 2017;70(4):411–22. doi:10.1016/j.jacc.2017.05.047

104. World Cancer Research Fund/American Institute for Cancer Research. Diet, nutrition, physical activity and cancer: A global perspective [Internet]. Continuous update project expert report 2018. 2018.

105. Chen J-P, Chen G-C, Wang X-P, Qin L, Bai Y. Dietary fiber and metabolic syndrome: A meta-analysis and review of related mechanisms. Nutrients. 2017;10(1):24. doi:10.3390/nu10010024

106. Wei B, Liu Y, Lin X, Fang Y, Cui J, Wan J. Dietary fiber intake and risk of metabolic syndrome: A meta-analysis of observational studies. Clin Nutr [Internet].

2018;37(6):1935–42. doi:10.1016/j.clnu.2017.10.019

107. Ye EQ, Chacko SA, Chou EL, Kugizaki M, Liu S. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. J Nutr. 2012;142(7):1304–13. doi:10.3945/jn.111.155325
108. Farvid MS, Chen WY, Rosner BA, Tamimi RM, Willett WC, Eliassen AH. Fruit and vegetable consumption and breast cancer incidence: Repeated measures over 30 years of follow-up. Int J Cancer. 2019;144(7):1496–510. doi:10.1002/ijc.31653
109. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9
110. McRorie JW, McKeown NM. Understanding the physics of functional fibers in the gastrointestinal tract: An evidence-based approach to resolving enduring misconceptions about insoluble and soluble fiber. J Acad Nutr Diet. 2017;117(2):251–64. doi:10.1016/j.jand.2016.09.021
111. Mann J, Cummings JH, Englyst HN, Key T, Liu S, Riccardi G, et al. FAO/WHO Scientific update on carbohydrates in human nutrition: Conclusions. Eur J Clin Nutr. 2007;61(S1):S132–7. doi:10.1038/sj.ejcn.1602943
112. FAO/WHO. CODEX Alimentarius (CODEX) Guidelines on nutrition labelling CAG/GL 2-1985 [Internet]. 2017.
113. Cummings JH, Stephen AM. The role of the dietary fibre in the human colon. Can Med Assoc J. 1980;123(11):1109–14.
114. He L-X, Zhao J, Huang Y-S, Li Y. The difference between oats and beta-glucan extract intake in the management of HbA1c, fasting glucose and insulin sensitivity: A meta-analysis of randomized controlled trials. Food Funct. 2016;7(3):1413–28. doi:10.1039/c5fo01364j
115. Tucker LA, Thomas KS. Increasing total fiber intake reduces risk of weight and fat gains in women. J Nutr. 2009;139(3):576–81. doi:10.3945/jn.108.096685
116. Stephen AM, Cummings JH. Mechanism of action of dietary fibre in the human colon. Nature [Internet]. 1980 Mar;284(5753):283–4. doi:10.1038/284283a0
117. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. Cell. 2016;165(6):1332–45. doi:10.1016/j.cell.2016.05.041
118. Cummings JH, Macfarlane GT. The control and consequences of bacterial fermentation in the human colon. J Appl Bacteriol [Internet]. 1991 Jun;70(6):443–59. doi:10.1111/j.1365-2672.1991.tb02739.x
119. Makki K, Deehan EC, Walter J, Bäckhed 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
120. Litvak Y, Byndloss MX, Bäumler AJ. Colonocyte metabolism shapes the gut microbiota. Science (80-) [Internet]. 2018 Nov 30;362(6418):eaat9076. doi:10.1126/science.aat9076
121. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int J Obes. 2017;41(7):1099–105. doi:10.1038/ijo.2017.66
122. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal

cancer. Nat Rev Microbiol. 2014;12(10):661–72. doi:10.1038/nrmicro3344

123. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut [Internet]. 2016;65(1):57–62. doi:10.1136/gutjnl-2015-309618

124. Sekirov I, Russell SL, Antunes LCM, Finlay BB. Gut microbiota in health and disease. Physiol Rev. 2010;90:859–904. doi:10.1152/physrev.00045.2009.

125. Tannock GW. Understanding the gut microbiota. Hoboken (NJ): John Wiley & Sons; 2017. 1–174 p.

126. Walter J, Ley R. The human gut microbiome: Ecology and recent evolutionary changes. Annu Rev Microbiol. 2011;65:411–29. doi:10.1146/annurev-micro-090110-102830

127. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220–30. doi:10.1038/nature11550

128. Thorburn AN, Macia L, Mackay CR. Diet, Metabolites, and "Western-Lifestyle" Inflammatory Diseases. Immunity [Internet]. 2014 Jun;40(6):833–42. doi:10.1016/j.immuni.2014.05.014

129. Kovatcheva-Datchary P, Nilsson A, Akrami R, Lee YS, De Vadder F, Arora T, et al. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. Cell Metab. 2015;22(6):971–82. doi:10.1016/j.cmet.2015.10.001

130. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. Cell. 2015;163:1079–94. doi:10.1016/j.cell.2015.11.001

131. Korem T, Zeevi D, Zmora N, Weissbrod O, Bar N, Lotan-Pompan M, et al. Bread affects clinical parameters and induces gut microbiome-associated personal glycemic responses. Cell Metab [Internet]. 2017;25:1243–53. doi:10.1016/j.cmet.2017.05.002

132. Martínez I, Kim J, Duffy PR, Schlegel VL, Walter J. Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. PLoS One [Internet]. 2010;5(11):e15046. doi:10.1371/journal.pone.0015046

133. Venkataraman A, Sieber JR, Schmidt AW, Waldron C, Theis KR, Schmidt TM. Variable responses of human microbiomes to dietary supplementation with resistant starch. Microbiome [Internet]. 2016 Dec 29;4(1):33. doi:10.1186/s40168-016-0178-x

134. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science (80-) [Internet]. 2016 Apr 29;352(6285):560–4. doi:10.1126/science.aad3503

135. Kootte RS, Levin E, Salojärvi J, Smits LP, Hartstra A V., Udayappan SD, et al. Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. Cell Metab. 2017;26(4):611-619.e6. doi:10.1016/j.cmet.2017.09.008

136. O'Toole PW, Claesson MJ. Gut microbiota: Changes throughout the lifespan from infancy to elderly. Int Dairy J [Internet]. 2010;20(4):281–91. doi:10.1016/j.idairyj.2009.11.010

137. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science (80-) [Internet]. 2011 Oct

7;334(6052):105–8. doi:10.1126/science.1208344

138. Johnson AJ, Vangay P, Al-Ghalith GA, Hillmann BM, Ward TL, Shields-Cutler RR, et al. Daily sampling reveals personalized diet-microbiome associations in humans. Cell Host Microbe. 2019;25(6):789-802.e5. doi:10.1016/j.chom.2019.05.005

139. Scott FI, Horton DB, Mamtani R, Haynes K, Goldberg DS, Lee DY, et al. Administration of antibiotics to children before age 2 years increases risk for childhood obesity. Gastroenterology [Internet]. 2016;151(1):120–9. doi:10.1053/j.gastro.2016.03.006

140. Gerber JS, Bryan M, Ross RK, Daymont C, Parks EP, Localio AR, et al. Antibiotic exposure during the first 6 months of life and weight gain during childhood. J Am Med Assoc. 2016;315(12):1258–65. doi:10.1001/jama.2016.2395

141. 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

142. Bäckhed F, Ding H, Wang T, Hooper L V., Gou YK, Nagy A, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A. 2004;101(44):15718–23. doi:10.1073/pnas.0407076101

143. Bäckhed F, Manchester JK, Semenkovich CF, Gordan JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. Proc Natl Acad Sci. 2007;104(3):979–84. doi:10.1007/BF00540675

144. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science (80-) [Internet]. 2013 Sep 6;341(6150):1241214–1241214. doi:10.1126/science.1241214

145. Cani PD, Van Hul M, Lefort C, Depommier C, Rastelli M, Everard A. Microbial regulation of organismal energy homeostasis. Nat Metab [Internet]. 2019;1(1):34–46. doi:10.1038/s42255-018-0017-4

146. Ley RE, Turnbaugh PJ, Klein S, Gordan JI, Gordon JI. Human gut microbes associated with obesity. Nature [Internet]. 2006;444(7122):1022–3. doi:10.1038/nature4441021a

147. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci. 2005;102(31):11070–5. doi:10.1073/pnas.0504978102

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

149. Cotillard A, Kennedy SP, Kong LC, Prifti E, Pons N, Le Chatelier E, et al. Dietary intervention impact on gut microbial gene richness. Nature [Internet]. 2013;500(7464):585–8. doi:10.1038/nature12480

150. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: Relationship with gut microbiome richness and ecology. Gut [Internet]. 2016;65(3):426–36. doi:10.1136/gutjnl-2014-308778

151. Schwiertz A, Taras D, Schäfer K, Beijer S, Bos NA, Donus C, et al. Microbiota and SCFA in lean and overweight healthy subjects. Obesity [Internet]. 2010;18(1):190–5. doi:10.1038/oby.2009.167

152. Duncan SH, Lobley GE, Holtrop G, Ince J, Johnstone AM, Louis P, et al. Human colonic microbiota associated with diet, obesity and weight loss. Int J Obes. 2008;32(11):1720–

4. doi:10.1038/ijo.2008.155

153. Sze MA, Schloss PD. Looking for a signal in the noise: Revisiting obesity and the microbiome. MBio [Internet]. 2016 Sep 7;7(4):1–10. doi:10.1128/mBio.01018-16

154. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. Nature [Internet]. 2013;500(7464):541–6. doi:10.1038/nature12506

155. Beaumont M, Goodrich JK, Jackson MA, Yet I, Davenport ER, Vieira-Silva S, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol [Internet]. 2016;17(1). doi:10.1186/s13059-016-1052-7

156. Manichanh C. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut [Internet]. 2006;55(2):205–11. doi:10.1136/gut.2005.073817

157. Lambeth SM, Carson T, Lowe J, Ramaraj T, Leff JW, Luo L, et al. Composition diversity and abundance of gut microbiome in prediabetes and type 2 diabetes. J Diabetes Obes. 2015 Dec 26;2(2):108–14. doi:10.15436/2376-0949.15.031

158. Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, et al. Structure, function and diversity of the healthy human microbiome. Nature [Internet]. 2012;486(7402):207–14. doi:10.1038/nature11234

159. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, et al. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. Front Genet [Internet]. 2019;10:454. doi:10.3389/fgene.2019.00454

160. Partula V, Mondot S, Torres MJ, Kesse-Guyot E, Deschasaux M, Assmann K, et al. Associations between usual diet and gut microbiota composition: Results from the Milieu Intérieur cross-sectional study. Am J Clin Nutr. 2019;109(5):1472–83. doi:10.1093/ajcn/nqz029

161. 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 (80-) [Internet]. 2016 Apr 29;352(6285):565–9. doi:10.1126/science.aad3369

162. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature [Internet]. 2011;473(7346):174–80. doi:10.1038/nature09944

163. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol [Internet]. 2018 Jan 18;3(1):8–16. doi:10.1038/s41564-017-0072-8

164. Claesson M, Jeffery IB, Conde S, Power S, O'Connor E, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;(488):178–84. doi:10.1038/nature11319

165. Dhakan DB, Maji A, Sharma AK, Saxena R, Pulikkan J, Grace T, et al. The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. Gigascience. 2019;8(3):1–20. doi:10.1093/gigascience/giz004

166. Hjorth MF, Roager HM, Larsen TM, Poulsen SK, Licht TR, Bahl MI, et al. Pre-treatment microbial Prevotella-to-Bacteroides ratio, determines body fat loss success during a 6-

month randomized controlled diet intervention. Int J Obes [Internet]. 2018;42:580–3. doi:10.1038/ijo.2017.220

167. Hjorth MF, Blædel T, Bendtsen LQ, Lorenzen JK, Holm JB, Kiilerich P, et al. Prevotella-to-Bacteroides ratio predicts body weight and fat loss success on 24-week diets varying in macronutrient composition and dietary fi ber: Results from a post-hoc analysis. Int J Obes. 2019;43:149–57. doi:10.1038/s41366-018-0093-2

168. 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 [Internet]. 2018;24(10):1526–31. doi:10.1038/s41591-018-0160-1

169. He Y, Wu W, Zheng HM, Li P, McDonald D, Sheng HF, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat Med [Internet]. 2018;24(10):1532–5. doi:10.1038/s41591-018-0164-x

170. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut. 2016;65(11):1812–21. doi:10.1136/gutjnl-2015-309957

171. Pendyala S, Walker JM, Holt PR. A high-fat diet is associated with endotoxemia that originates from the gut. Gastroenterology [Internet]. 2012;142:1100–1. doi:10.1053/j.gastro.2012.01.034

172. Albenberg LG, Wu GD. Diet and the intestinal microbiome: Associations, functions, and implications for health and disease. Gastroenterology [Internet]. 2014;146(6):1564–72. doi:10.1053/j.gastro.2014.01.058

173. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature [Internet]. 2014;505(7484):559–63. doi:10.1038/nature12820

174. So D, Whelan K, Rossi M, Morrison M, Holtmann G, Kelly JT, et al. Dietary fiber intervention on gut microbiota composition in healthy adults: a systematic review and meta-analysis. Am J Clin Nutr. 2018;965–83. doi:10.1093/ajcn/nqy041

175. de Moraes ACF, Fernandes GR, da Silva IT, Almeida-Pititto B, Gomes EP, Pereira A da C, et al. Enterotype may drive the dietary-associated cardiometabolic risk factors. Front Cell Infect Microbiol. 2017;7(FEB):1–9. doi:10.3389/fcimb.2017.00047

176. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. BMJ. 2016;65(1):63–72. doi:10.1136/gutjnl-2014-308209

177. Losasso C, Eckert EM, Mastrorilli E, Villiger J, Barrucci F, Pernthaler J, et al. Assessing the influence of vegan, vegetarian and omnivore oriented westernized dietary styles on human gut microbiota: A cross sectional study. Front Microbiol. 2018;9(March):1–12. doi:10.3389/fmicb.2018.00317

178. Mcdonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G, et al. American gut: an open platform for citizen science microbiome research. Am Soc Microbiol. 2018;3(3):1–28. doi:10.1128/mSystems.00031-18

179. Lin D, Peters BA, Friedlander C, Freiman HJ, Goedert JJ, Sinha R, et al. Association of dietary fibre intake and gut microbiota in adults. Br J Nutr [Internet]. 2018;120(9):1014–22. doi:10.1017/s0007114518002465

180. Koo SH, Chu CW, Khoo JJC, Cheong M, Soon GH, Ho EXP, et al. A pilot study to examine the association between human gut microbiota and the host's central obesity. JGH Open [Internet]. 2019 Dec 16;3(6):480–7. doi:10.1002/jgh3.12184
181. Gutiérrez-Díaz I, Fernández-Navarro T, Salazar N, Bartolomé B, Moreno-Arribas MV, de Andres-Galiana EJ, et al. Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish cohort of middle-age and older people. J Agric Food Chem [Internet]. 2017 Jan 25;65(3):586–95. doi:10.1021/acs.jafc.6b04408
182. Bowyer RCE, Jackson MA, Pallister T, Skinner J, Spector TD, Welch AA, et al. Use of dietary indices to control for diet in human gut microbiota studies. Microbiome [Internet]. 2018 Dec 25;6(1):77. doi:10.1186/s40168-018-0455-y
183. Mitsou EK, Kakali A, Antonopoulou S, Mountzouris KC, Yannakoulia M, Panagiotakos DB, et al. Adherence to the Mediterranean diet is associated with the gut microbiota pattern and gastrointestinal characteristics in an adult population. Br J Nutr [Internet]. 2017;117(12):1645–55. doi:10.1017/s0007114517001593
184. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. Am J Clin Nutr [Internet]. 2019 Oct 1;110(4):1003–14. doi:10.1093/ajcn/nqz174
185. Oluwagbemigun K, Foerster J, Watkins C, Fouhy F, Stanton C, Bergmann MM, et al. Dietary patterns are associated with serum metabolite patterns and their association is influenced by gut bacteria among older german adults. J Nutr [Internet]. 2019 Sep 2;1–10. doi:10.1093/jn/nxz194
186. Kong LC, Holmes BA, Cotillard A, Habi-Rachedi F, Brazeilles R, Gougis S, et al. Dietary patterns differently associate with inflammation and gut microbiota in overweight and obese subjects. PLoS One. 2014;9(10). doi:10.1371/journal.pone.0109434
187. Agnoli C, Krogh V, Grioni S, Sieri S, Palli D, Masala G, et al. A priori–defined dietary patterns are associated with reduced risk of stroke in a large Italian cohort. J Nutr. 2011;141:1552–8. doi:10.3945/jn.111.140061
188. Tang W, Wang Z, Levison B, Koeth R, Britt E, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med. 2013;368(17):1575–84. doi:10.1056/NEJMoa1109400
189. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med [Internet]. 2013;472:57–63. doi:10.1038/nm.3145
190. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite Trimethylamine-N-Oxide. Blaser MJ, editor. MBio [Internet]. 2015 Mar 17;6(2):1–8. doi:10.1128/mBio.02481-14
191. Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier M, Beaugerie L, et al. Low counts of Faecalibacterium prausnitzii in colitis microbiota. Inflamm Bowel Dis. 2009;15(8):1183–9. doi:10.1002/ibd.20903

3. Chapter 3. Methods

This thesis describes PhD research that utilised a cross-sectional study; the PRedictors linking Obesity and gut MIcrobiomE (PROMISE) study, which was funded by the Health Research Council of New Zealand. The purpose of this methods section is to highlight the PROMISE methods which are relevant to this PhD research, and to clearly distinguish data processing and analyses completed for this PhD research; all of which will be outlined below. Please see Appendix 1 for dissemination of this PhD research and Appendix 2 for PROMISE team members contributions to this PhD research.

3.1. Study design

The PROMISE study was conducted between July 2016 and September 2017, at the Human Nutrition Unit (HNU) at Massey University in Albany, New Zealand. The objective of the PROMISE study was to characterise the gut microbiome in two populations (Pacific and NZE women) with different metabolic disease risk and body weight profiles (normal and obese); and to further explore the association with dietary intake, eating behaviour, taste perception, sleep, and physical activity. Details of the PROMISE study procedures and recruitment have been reported elsewhere (1). This PhD research utilised the cross-sectional PROMISE study, to explore habitual dietary intake in relation to body fat content, metabolic health biomarkers, and the gut microbiota.

3.1.1. Participants

Eligible participants were post-menarche and premenopausal women, aged between 18 and 45 years, who self-identified as either NZE ethnicity (and lived in NZ for over 5 years) or Pacific ethnicity (with one parent of Pacific ethnicity) and were generally healthy. Participants were selectively recruited to the PROMISE study based on BMI, therefore exclusion criteria included BMI outside of the predefined normal and obese BMI ranges, presence of any chronic illness or disease (e.g., CVD), previous bariatric surgery, pregnant or lactating, smokers, antibiotic use in the past month or medications that could interfere with the immune system or appetite, and severe dietary restrictions or avoidances (e.g., vegan). Participants were recruited from the wider Auckland region, either online, in-person or via the phone and details of the PROMISE study recruitment have been reported elsewhere (1). Research staff utilised self-reported weight and height to calculate BMI to screen eligible participants to invite to participate in the PROMISE study.

3.1.2. Ethics

The PROMISE study was approved by the Southern Health Disability Ethics Committee (16/STH/32) and conducted according to the guidelines of the declaration of Helsinki. The trial was registered at anzctr.org.au (ACTRN12618000432213). All participants were provided with detailed information about the study procedures, study measurements, and gave written informed consent to participate. Access to data was restricted to the immediate research team, and only deidentified coded data was used for data analysis.

3.2. Study procedure

Briefly, participants attended two study visits and participants completed at home data collection between the study visits. On arrival to the HNU at Massey University at visit one, participants were welcomed, asked to carefully read, and sign a consent form and had the opportunity to ask questions prior to commencing the study. Each participant was then allocated a unique study ID, which was used for all later analysis to ensure all participant data was deidentified. Following the consent process, visit one included a demographic interview with a researcher, anthropometric measurements, blood sampling, and instructions for at home data collection. At home data collection included a 5-day non-consecutive estimated food record and faecal sample collection. Visit two was scheduled 11-14 days after the first visit; where participants returned all the data they collected at home, completed a semi quantitative food frequency questionnaire, a one-on-one food record interview and body composition measurements by DXA. Once participants had completed all aspects of the PROMISE study they were thanked for their time and provided with a *koha* (a gift or gratuity) to acknowledge their contribution to the study.

3.2.1. Demographics

Participants completed a one-on-one interview with a researcher to capture a range of demographic and health information including their address, occupation, personal/household income, number of children, dietary supplement use, birth weight and delivery method (if known) (see Appendix 3). Deprivation index was assessed in this study as a measurement of socioeconomic status. The New Zealand Deprivation 2013 index (NZDep2013), is an area-based measure of socioeconomic deprivation, and combines information from the national

census data including, housing, qualifications, employment, income, family structure, and access to transport (2). The NZDep2013 provides a deprivation score for each meshblock in NZ. Meshblocks are small geographical areas defined by Statistics NZ, that consist of a population of around 60 to 110 individuals. The NZDep2013 groups deprivation into deciles; ranging from decile one “least deprived” and decile ten “most deprived” (2).

3.2.2. Anthropometric measurements

Research staff who were level 1 International Society for the Advancement of Kinanthropometry (ISAK) trained, conducted anthropometric measurements (e.g., fasted weight, stretched height) according to ISAK protocols (3). Fasting weight (kg) and height (cm) were used to calculate BMI kg/m^2 (weight (kg)/height (m^2)) using the Quetelet index. At study visit two, a whole-body scan was conducted with Dual-energy X-ray Absorptiometry (DXA) (Hologic QDR Discovery A, Hologic Inc, Bedford, MA with APEX V. 3.2 software) to assess total and regional body fat percentage (BF%) (e.g., visceral, android, and gynoid fat percentage).

3.2.3. Blood samples

At study visit one, following an overnight fast of at least 10 hours, fasting blood samples were collected by a trained phlebotomist between 7:30 and 9am. A tourniquet was applied moderately to the arm before venipuncture. Blood was drawn from the same arm into four vacutainers (maximum total blood volume was 30 mL), to obtain plasma and serum for analysis of metabolic markers and endocrine regulators. Blood samples were collected into Ethylenediaminetetraacetic acid (EDTA) 10 mL vacutainers (Becton Dickinson) and silicone coated (with clot activator) vacutainer tubes, respectively. Immediately following blood collection, an aliquot of EDTA whole blood was frozen at -80°C for HbA1c analysis, and serum tubes were kept at room temperature for 30 minutes for clot formation. The remaining samples were kept at 4°C, and within an hour of collection, all four vacutainer tubes were centrifuged at 3500 rpm for 15 minutes at 4°C, and plasma and serum samples were frozen at -80°C until further analysis.

Analysis of blood samples was completed at the Liggins Institute, Auckland, New Zealand. Serum levels of glucose (enzymatic UV method), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C) (enzymatic colorimetric method) were measured using a Hitachi c311 autoanalyser (Hitachi High Technologies Corporation, Tokyo, Japan) with Roche Diagnostic reagents (Mannheim, Germany) for all assays (4). Using EDTA whole blood, HbA1c levels were measured by turbidimetric inhibition immunoassay method (Roche Diagnostic, Mannheim, Germany) on a Hitachi c311 autoanalyser (Hitachi High Technologies Corporation, Tokyo, Japan). The average inter-assay coefficient of variation for all the above analyses performed using the Hitachi c311 autoanalyser was 1.5 %. Serum levels of insulin were measured using the electrochemiluminescence immunoassay method (Roche Diagnostics, Mannheim, Germany) on the Cobas e411 analyser (Hitachi High Technologies Corporation, Tokyo, Japan). The inter-assay coefficient of variation was 0.5 %. Homeostasis model assessment (HOMA-IR) index for insulin resistance was calculated (fasting blood glucose [mmol/L] X fasting plasma insulin [μU/mL]/22.5) (5).

3.2.4. Blood pressure

At visit two, at rest blood pressure was measured after a 10-minute (sitting) period with an Omron HEM 907 digital blood pressure monitor. The measurement was repeated three times and the average of the second and third measurement was recorded.

3.3. Dietary assessment

At the end of visit one, participants watched a ten-minute video on how to complete their five-day, non-consecutive, estimated food record (5DFR), at home (see Appendix 4 for 5DFR). The in-house training video was developed by researchers at Massey University HNU. Trained researchers further clarified the food recording procedure and answered any participant questions. A standardised portion guide booklet (6), and standard household measures (e.g., metric cup), were provided to aid participants in completing the 5DFR. During participants second study visit, all 5DFRs were reviewed by a New Zealand Registered dietitian (NZRD, either Jo Slater or Nikki Renall), followed by an individual one-on-one interview with each participant and a NZRD. The purpose of these interviews was to clarify the food recorded, food

preparation and cooking methods used, portion sizes, and brands of foods consumed, or any potential ambiguities identified to inform accurate dietary data entry. Participants also completed a validated semi-quantitative, NZ Women's Food Frequency Questionnaire (NZWFFQ) (7) regarding the past thirty days food intake (see Appendix 5). The 220-item NZWFFQ provided standard portion sizes (e.g., 1 teaspoon of sugar, 1 slice of bread) as a reference for participants, and nine standard frequency categories (ranging from 1 = "never" to 9 = "4 or more times a day") to enable intake estimation. The NZWFFQ was hosted by SurveyMonkey© software (SurveyMonkey Inc, San Mateo, California, USA) online, which enabled the research staff the ability to monitor live progress.

3.4. Dietary data processing

Energy, macro- and micronutrient analyses of the 5DFR and NZWFFQ were completed using FoodWorks9 (Xyris Software (Australia) Pty Ltd, Queensland, Australia). FoodWorks9 hosts multiple food composition databases including NZ's food composition database (FOODFiles 2016), developed by the NZ Institute for Plant & Food Research and the NZ Ministry of Health. The NZRDs (Jo Slater and Nikki Renall) entered the 5DFR and NZWFFQ independently into the Foodworks9 software, using a SOP that was developed by the NZRDs to ensure consistent and reliable dietary data entry. The NZRDs utilised a hierarchy of food choices from the different food composition databases. If a direct match to the reported food item could not be found in NZ FOODFiles 2016, then the Xyris brandname database AusFoods 2017 and AusBrands 2017 (based on the Australian food composition databases AUSNUT 2011-13, developed by Food Standards Australia New Zealand) were used. During dietary data entry, a "PROMISE dietary data assumption dictionary" was generated, detailing the reported food in a participants 5DFR, and the selected option within the software. This dictionary ensured standardised dietary data entry for all participants in the PROMISE study.

Food items that did not have a direct compositional match within the FoodWorks9 databases were discussed at weekly meetings with the nutrition research team to resolve data entry. For example, if a participant was unable to report an accurate recipe (all/quantities of ingredients, method of cooking, etc.) or reported purchased food items (e.g., savoury scone), then a "standard recipe" was created from existing analysed foods within the databases. For a participant recipe (e.g., chocolate cake with a direct compositional match in the database),

the existing analysed recipe was selected and entered as “chocolate cake” from NZ FOODfiles2016. Where appropriate, all food items were entered in the software as the cooked portion consumed. Cooking factors were used from McCance and Widdowson’s (8), to convert the raw weights of the ingredients reported in participants meals, to allow for water and nutrient losses in the cooking process.

To ensure consistent data entry for all NZWFFQs, a standardised template of the NZWFFQs food items and corresponding portion sizes was created in the FoodWorks9 software by the NZRDs. Participants had the option to report any additional foods they consumed habitually (which were not included in the NZWFFQ) in the free text space at the end of the NZWFFQ. Additional foods were entered manually into the template for each individual following the same SOP developed for the 5DFR. The nine standard frequency categories were converted to units of daily consumption (e.g., “never” = 0, “less than once a month” = 0.01, “1-3 times a month” = 0.07, “once a week” = 0.14, “2-3 times a week” = 0.36, “4-6 times a week” = 0.71, “once a day” = 1, “2-3 times a day” = 2.5 to “4 or more times a day” = 4). All dietary data entry (5DFR and NZWFFQ) were independently checked by a second NZRD to confirm reliability of data entry. All exported data sets were extensively reviewed by Nikki Renall for misreporting using the raw dietary data. Following extensive review of the raw and processed dietary data for plausibility of intake, the cut-offs of >2100 kJ/day and <27000 kJ/day were considered valid completion of the 5DFR and NZWFFQ for the PROMISE study. These cut offs are higher than Willet’s epidemiological recommendations (9); however, were considered more reflective of the foods consumed by the study population (10–12). Reported intakes below or above these cut-offs were excluded from further dietary analysis.

Total energy intake is reported as kilojoules (kJ) and includes the energy contribution from total dietary fibre, all macronutrients and alcohol. Total dietary fibre (reported as grams (g)) includes all non-starch polysaccharides (cellulose, hemicellulose, and resistant starches) and NZFOODFILES uses AOAC Prosky method to measure dietary fibre). Total fat (reported in grams (g)) includes all unsaturated and saturated fats. Total carbohydrate (reported in grams (g)) includes starch, free sugars and glycogen and dextrins. To assess the composition of the total diet, the percentage of energy intake from all macronutrients was calculated (13) and compared to the AMDRs for NZ and Australia (14).

3.4.1. Food groups

For this PhD research in order to assess the complex interrelationships between nutrients and foods, all reported foods items from the 5DFR (n>2850) and the NZWFFQ were summarised into 55 food groups (in g/day) based on similar nutritional composition and characteristics (see Supplementary Table 5.1). Mixed dishes reported in the participants' 5DFRs were assigned into the appropriate food group as follows: when participants provided us with the specific details (e.g., weights of raw ingredients, cooking methods) of the meals they consumed, the individual components of these "Participant meals" were assigned to the corresponding food groups. For example, for the meal 'spaghetti bolognaise', the cooked portion consumed of pasta was assigned to "refined grains", mince to "red meat", tomatoes to "tomatoes", and onion to "other-non starchy vegetables" etc. The same approach was taken for beverages; for example, milk and sugar added to coffee and/or tea was assigned to the appropriate food group (e.g., "full fat milk" and "added sugar to food and beverages" respectively). In contrast, participant and standard recipes were assigned to the corresponding food groups in their entirety (e.g., banana cake was assigned to the food group "cakes and biscuits", and bliss balls were assigned to "sweet snacks"). Primarily because the existing analysed foods and recipes, which were used to create "Participant recipes" and "Standard recipes", were unable to be broken down to individual components. In addition, for descriptive purposes, the 55 food groups were broadly classified as either 'core' or 'discretionary' foods based on NZ (Ministry of Health) and Australian (Department of Health) dietary guidelines (15,16). 'Core' foods are considered nutrient rich foods, from the four core food groups (e.g., vegetables, wholegrains, dairy) (17); whereas, 'discretionary' foods do not fit into these core or 'healthy' food groups; and are characteristically energy dense, processed nutrient poor foods (e.g., confectionary, potato chips/crisps) (15–17).

3.5. Dietary data analysis

3.5.1. Calculating habitual dietary intake

For this PhD research, each participant with valid dietary data (within the specified kJ cut off range), individual habitual dietary intake of nutrients was estimated using the National Cancer Institute (NCI) method. The three NCI SAS macros (version 2.1 - available from the NCI website

(18)) were adapted to the PROMISE dietary data. Utilising the 5DFR as the primary dietary data, the NCI method uses a two-part modelling approach to estimate the probability of consumption of a nutrient, food or food group, respectively, and the respective consumed amount, considering covariates that are affecting either the probability of consumption or the amount of consumed. For episodically consumed dietary components (e.g., components that are not consumed every day i.e., alcohol) the first part of the model estimates the probability of consumption using logistic regression model allowing for a residual between-person variation (covariates) to be included. The second part of the model uses a linear regression model on a transformed scale (as the NCI method requires near normal distribution of the residuals) to estimate the amount consumed, and, like the first model, considers covariates. The individual habitual intake was then defined as the product of probability of consumption x the consumed amount (*Usual intake = Probability x Amount*). Estimation of a dietary component consumed daily (e.g., energy (kJ)) requires a one-part model, where the process is the same as mentioned above, however there is no need to model probability of consumption, so this step was omitted.

In this PhD research, the three NCI SAS macros (Version 2.1 - available from the NCI website) (18) were adapted for the PROMISE data in SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). Utilising the 5DFR as the primary dietary data source, the covariates age, ethnicity, BMI, season (summer, winter, spring, autumn), NZWFFQ information (in standard units/day) and weekend information (Weekday = Monday-Thursday, Weekend = Friday-Sunday) were considered. If a participant reported never consuming a nutrient or food group in both the 5DFR and NZWFFQ they were assigned 0g/day for the respective nutrient/food group. The average daily intake of 36 nutrients (in units/day) and 55 food groups (g/day) consumed within the last month was estimated for each participant.

3.5.2. Dietary pattern analysis

Principal component analysis (PCA) was conducted to identify dietary patterns for this PhD research, where the principal component scores extracted from PCA represent dietary pattern scores for use in subsequent analyses. The estimated daily intake values of the 55 food groups (described above) were individually log transformed to approach normality prior to dietary

pattern analysis (using SPSS 25.0 for Windows (SPSS inc., Chicago, IL, USA)). All participants with available habitual dietary data were included in the exploratory PCA. The break-even point of the scree plot, the Kaiser criterion (eigenvalues greater than 1), varimax rotation, and the magnitude of the component loadings themselves, were all used to identify distinct and interpretable patterns, with component loadings of >0.3 considered to significantly contribute to the pattern. Food groups are weighted onto a component, either positively or negatively, by their contribution to explain the variance in the dietary pattern, with highly positive component loadings contributing significantly to the pattern. Cronbach's alpha (≥0.8 = good, 0.7 = moderate) (19) was used to assess the inter item reliability of each food group and the pattern they are associated with (e.g., whether the inclusion of a food group could skew the reliability of a pattern). Each participant was assigned an individual score of adherence to each extracted dietary pattern, which ranged from positive to negative values. Thus, an individual may score positively in one pattern and negatively in another. To account for higher energy intakes correlating with higher nutrient intakes (9,20), dietary pattern scores were adjusted for total energy (kJ) intake, and energy-adjusted dietary patterns were used for all subsequent analyses. These individual energy adjusted scores were then used to rank participants into tertiles of adherence to the different patterns for further analyses.

3.6. Faecal samples

3.6.1. Faecal sample collection

At the end of study visit one, participants received detailed verbal and written instructions (see Appendix 7) about how to collect their samples at home, during the period between study visit one and two and were provided with a collection kit. The kit contained two pre-labelled screw-top containers with a scoop in the lid (LBS3805 25 mL, ThermoFisher NZ), two larger pre-labelled screw-top plastic containers (LBS30130 130mL PP, ThermoFisher NZ), ice-sheets, kidney dishes, gloves and zip-lock plastic bags, all within an insulated carry bag. Participants were instructed to collect their faecal sample after they had completed their food record, and to record the date and time of sample collection. Participants were asked to take two samples from the same faecal sample, and to place the smaller sealed sample container into the larger container with 2-3 cm of cold water, which acted as a water jacket. Participants stored their samples within their household freezer (-20°C) until they returned for visit two (transporting

the samples in the insulated bag with ice sheets), where the samples were immediately stored in the -80**°**C freezer until laboratory analysis. Samples will be kept for 10 years and will not be analysed for anything that participants have not provided consent for.

3.6.2. Faecal water content

Faecal water content was used within this PhD research as a proxy for colonic transit time. To determine faecal water content, approximately 200mg of each faecal sample was placed in a pre-weighed microfuge tube, the weight recorded, and the tube with cap open placed in a 37**°**C incubator. The tubes were dried until a constant dry weight was obtained, and percentage water content was then calculated.

3.7. Microbiota analysis

3.7.1. DNA extraction and sequencing

Emeritus Professor Gerald Tannock and Dr Blair Lawley conducted the DNA extraction, analyses of the microbiota with MetaPhlan 2.0 and QIIME2, and the prediction of enterotypes (Microbiome Otago, University of Otago) as follows. DNA was extracted from 250 mg faeces according to the kit protocol provided by the manufacturer (PowerSoil DNA isolation kit, Mo Bio, Carlsbad, CA, USA), with the following modification. Faecal samples were suspended in 1mL of TN150 buffer (containing 10 mM TRIS-CL pH 8.0, 150 mM NaCl). The suspension was centrifuged at 14,600×*g* (3 min, 5°C) and then suspended in 700 µl solution from the PowerBead Tubes, from the PowerSoil DNA isolation kit. The suspension was added back to the PowerBead Tubes and the standard protocol followed. The modification of an additional centrifuging step was to improve the quality of the sample. DNA was eluted in 100 µl of elution buffer (warmed to 70^{o}C) and stored at -80°C.

Quality and quantity of genomic DNA was checked on a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and on a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) prior to sending the cleaned DNA to New Zealand Genomics Ltd. (NZGL) for shotgun metagenome sequencing. Both of these methods were employed as a high-level quality control. NZGL prepared 384 Thruplex DNA libraries and carried out 2×125

bp paired-end sequencing across 24 lanes on an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA, USA). Libraries were sequenced across a minimum of six HiSeq lanes, and multiple libraries were prepared for several samples to test for library preparation and sequence run bias. An average of 13,150,561 (range 7,6940,894-17,081,755) reads were recovered for each sample.

For quality control BBDuk (https://sourceforge.net/projects/bbmap/) was used to trim adapters, remove low quality reads and remove reads <100 bp after trimming. KneadData (http://huttenhower.sph.harvard.edu/kneaddata) was used as quality control to remove human genome reads (i.e., DNA) from bacterial reads, implementing the hg19 database. Sequence data used in this study will be deposited with the short-read archive (SRA).

3.7.2. Bioinformatic analysis

Microbiota taxonomic profiles were created from DNA sequences using MetaPhlan 2.0 (version 2.6.0) according to default parameters (21).

Microbiota composition and diversity was further analysed with QIIME2 (version 2018.8, Bolyen E et al., (22) (https://qiime2.org/) using converted output tables from MetaPhlan 2.0. Beta diversity group significance for each metric (Bray-Curtis Dissimilarity index, and Jacard similarity matrix) was measured with PERMANOVA (23) and group dispersion was measured with PERMDISP (24).

Enterotypes were predicted in R using the approach described in Arumugum et al. (25) and following the tutorial provided by EMBL (http://enterotyping.embl.de). Differential abundance testing to determine which species were driving enterotypes was carried out with Statistical Analysis of Taxonomic and Functional Profiles (STAMP) (26). Each enterotype was compared to all other samples using Welch's *t*-test using Benjamini-Hochberg for multiple testing correction.

3.8. Statistical analysis

3.8.1. Data handling

The data processing and handling discussed below were completed as part of this PhD research. Participants were selectively recruited by BMI; however, to control for the fact individuals with the same BMI can have different metabolic disease risks (27,28), participants were stratified into low and high BF% groups using the median as the cut point; that is, low-BF% (<35 %) versus high-BF% (≥35 %) (27,29,30). The NZDEP2013 index was collapsed into quintiles (e.g., quintile one = decile one and two, quintile two = decile three and four) and these quintiles were used for subsequent analyses. All data were tested for normality. Measurements that did not follow a normal distribution were reported as medians [25^{th}, 75^{th} percentiles] and non-parametric tests were used to assess the differences between groups. Where relevant, comparisons of participant characteristics focused on the differences between BF% groups within an ethnic group (e.g., NZE low-BF% versus NZE high-BF%), and between ethnicities within a BF% group (e.g., NZE low-BF% versus Pacific low-BF%). Statistical analyses were conducted using SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). For all analyses *p*-values <0.05 were considered statistically significant.

3.8.1.1. Energy adjustment of dietary data

Higher energy intake correlates with higher nutrient intake, therefore adjusting nutrient and food group intake for total energy intake allows for the evaluation of diet composition and health outcomes independent of energy intake. All dietary data analyses (nutrients and food groups) were adjusted for energy intake. Dietary pattern scores were adjusted for energy intake and energy adjusted dietary patterns were used for all subsequent analyses.

3.8.1.2. Covariates

All regression models were adjusted for age and NZDep2013 quintiles. Due to the design of the study (i.e., selectively recruited based on ethnicity) where appropriate combined analyses were adjusted for ethnicity. All independent variables were assessed for collinearity by assessing tolerance and the variance inflation factor.

3.8.2. Chapter 4

To assess the association between diet (macronutrient intake) and BF%, BF% was used as the dependent variable and macronutrients were used as the independent variables. Firstly, univariate linear regression analyses were conducted followed by multivariate linear analyses controlled for potential confounders including total dietary fibre, protein, total carbohydrate, total fat, energy intake, age, and NZDep2013 quintiles. Analyses were conducted separately for NZE and Pacific participants, as well as both groups combined. To assess the association of habitual dietary fibre intake and participants' characteristics, all participants were stratified by tertiles of energy adjusted dietary fibre intake. Comparisons focused on the lowest tertile of dietary fibre intake (lowest 33 % of the population) versus the highest tertile of dietary fibre intake (tertile three).

3.8.3. Chapter 5

To assess the association between habitual diet (adherence to dietary patterns) and BF%, BF% was used as the dependent variable and dietary pattern scores were used as the independent variables. Univariate linear regression analyses were conducted followed by multivariate linear regression analyses controlling for potential confounders including: NZDep2013 quintiles, age and ethnicity. Participants were stratified into tertiles of adherence to the energy-adjusted dietary patterns and comparisons focused on low adherence (lowest 33 % of the population) *versus* high adherence (tertile three). Spearman's rank correlation analysis was used to explore the associations between the adjusted dietary patterns scores, and food groups and nutrients. To assess the macronutrient intake in relation to the total energy intake of the diet, the percentage of total energy intake from macronutrients were calculated (14).

3.8.4. Chapter 6

To assess the association between habitual diet and microbiota composition (expressed as enterotypes and relative abundance) the associations with energy-adjusted dietary patterns (determined with PCA as described previously), the 55 food groups, and 36 nutrients were explored. Logistic regression analyses were conducted for dichotomous outcomes (e.g., relative abundance of bacteria species) and multinomial regression for (outcome) variables

that were categorical (e.g., enterotype) but not dichotomous. For species that were either present or absent in the study population, “presence” versus “absence” was used as the outcome variable. Species that were relatively abundant in all participants were stratified as either “lower” or “higher” relative abundance (using the median as the cut point). Univariate analysis was conducted followed by multivariate analyses controlling for potential confounders including age, NZDep2013 quintiles, faecal water content, and energy intake. Analyses were conducted separately for NZE and Pacific participants, as well as for both groups combined. For each participant the *Firmicutes* to *Bacteriodetes* ratio was calculated by dividing the relative abundance of the phyla *Firmicutes* by the relative abundance of *Bacteriodetes.*

3.9. References

1. Kindleysides S, Kruger R, Douwes J, Tannock G, Renall N, Slater J, et al. PRedictors linking Obesity and the gut MIcrobiomE (The PROMISE study): Protocol and recruitment strategy of a cross-sectional study on pathways that affect the gut microbiome and its impact on obesity. JMIR Res Protoc. 2019;8:1–16. doi:10.2196/14529

2. Atkinson J, Salmond C, Crampton P. NZDep2013 index of deprivation. NZDep2013 Index of Deprivation. Wellington; 2014.

3. Stewart A, Marfell-Jones M, Olds T, Hans De Ridder J. International standards for anthropometric assessment. Lower Hutt, New Zealand: International Society for the Advancement of Kinanthropometry; 2011.

4. Rodríguez-Cabaleiro D, Van Uytfanghe K, Stove V, Fiers T, Thienpont LM. Pilot study for the standardization of insulin immunoassays with isotope dilution–liquid chromatography/tandem mass spectrometry. Clin Chem [Internet]. 2007 Aug 1;53(8):1462–9. doi:10.1373/clinchem.2007.088393

5. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: Insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28(7):412–9. doi:10.1007/BF00280883

6. Nelson M, Atkinson M, Meyer J. A photographic atlas of food portion sizes. London: Ministry of Agriculture, Fisheries and Food; 1997. 1–109 p.

7. Beck KL, Houston ZL, McNaughton SA, Kruger R. Development and evaluation of a food frequency questionnaire to assess nutrient intakes of adult women in New Zealand. Nutr Diet. 2018;1–7. doi:10.1111/1747-0080.12472

8. Finglas P., Roe M., Pinchen H., Berry R, Church S., Dodhia S., et al. McCance and Widdowson's the composition of foods, seventh summary edition. 7th ed. Cambridge: The Royal Society of Chemistry; 2015. 1–630 p. doi:10.1039/9781849737562

9. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001

10. Gemming L, Jiang Y, Swinburn B, Utter J, Mhurchu CN. Under-reporting remains a key limitation of self-reported dietary intake: An analysis of the 2008/09 New Zealand Adult Nutrition Survey. Eur J Clin Nutr [Internet]. 2014;68(2):259–64. doi:10.1038/ejcn.2013.242

11. Smith C, Gray AR, Mainvil LA, Fleming EA, Parnell WR. Secular changes in intakes of foods among New Zealand adults from 1997 to 2008/09. Public Health Nutr [Internet]. 2015 Dec 10;18(18):3249–59. doi:10.1017/S1368980015000890

12. Metcalf PA, Scragg RRK, Schaaf D, Dyall L, Black PN, Jackson R. Dietary intakes of European, Māori, Pacific and Asian adults living in Auckland: The Diabetes, Heart and Health Study. Aust N Z J Public Health [Internet]. 2008;32(5):454–60. doi:10.1111/j.1753-6405.2008.00279.x

13. Food Standards Australia New Zealand (FSANZ). Australia New Zealand Food Standards

Code-Section 11-Calculation of values for nutrition information panel [Internet]. F2018C00953. 2018 [cited 2019 Aug 20].

14. National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. Nutrient reference values for Australia and New Zealand. Nutrition. Canberra; 2006.

15. National Health and Medical Research Council. Australian Dietary Guidelines. Canberra: National Health and Medical Council; 2013.

16. Ministry of Health. Eating and activity guidelines for New Zealand adults [Internet]. Wellington; 2015.

17. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr. 2016;7(3):445–54. doi:10.3945/an.115.011718

18. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

19. Field A. Discovering statistics using SPSS. 3rd ed. London: SAGE Publications Ltd; 2009.

20. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr [Internet]. 1997 Apr 1;65(4):1220S-1228S. doi:10.1093/ajcn/65.4.1220S

21. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods [Internet]. 2012;9(8):811–4. doi:10.1038/nmeth.2066

22. Bolyen E, Dillon M, Bokulich N, Abnet C, Al-Ghalith G, Alexander H, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Prepr. 2018; doi:10.7287/peerj.preprints.27295

23. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol [Internet]. 2001 Feb;26(1):32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

24. Anderson MJ. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics [Internet]. 2006;62(1):245–53. doi:10.1111/j.1541-0420.2005.00440.x

25. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature [Internet]. 2011;473(7346):174–80. doi:10.1038/nature09944

26. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics [Internet]. 2014;30(21):3123–4. doi:10.1093/bioinformatics/btu494

27. Oliveros E, Somers VK, Sochor O, Goel K, Lopez-Jimenez F. The concept of normal weight obesity. Prog Cardiovasc Dis. 2014;56(4):426–33. doi:10.1016/j.pcad.2013.10.003

28. Kramer C, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions? A systematic review and meta-analysis. Ann Intern Med. 2013;159(11):758–69. doi:10.7326/0003-4819-159-11-201312030-00008

29. Dickey RA, Bartuska DG, Bray GW, Callaway CW, Davidson ET, Feld S, et al. AACE/ACE position statement on the prevention, diagnosis, and treatment of obesity (1998 Revision) AACE/ACE Obesity Task Force. Endocr Pract. 1998;4(5):297–350.

30. Jo A, Mainous AG. Informational value of percent body fat with body mass index for the risk of abnormal blood glucose: A nationally representative cross-sectional study. BMJ Open. 2018;8(4):1–6. doi:10.1136/bmjopen-2017-019200

4. Chapter 4. Higher habitual dietary fibre intake is associated with lower body fat percentage and metabolic disease risk in New Zealand European and Pacific women.

4.1. Abstract

Introduction: Diet is a modifiable risk factor for noncommunicable diseases like obesity. This study assessed relationships between habitual dietary intake, body fat content and metabolic disease risk of New Zealand European (NZE) and Pacific women.

Methods Healthy premenopausal Pacific (*n*=126) and NZE (*n*=161) women (*n*=287) aged 18-45 years were recruited based on body mass index (lean and obese) and stratified as either low (<35 % body fat) or high (≥35 % body fat) body fat percentage (BF%). Habitual dietary intake was calculated with the National Cancer Institute method, using a 5-day estimated food record and a semi-quantitative food frequency questionnaire. BF% was assessed by dual-energy x-ray absorptiometry. Fasting blood samples were analysed for metabolic health markers (e.g., fasting glucose, insulin and lipid profile). Adjusted multivariate linear regression models were conducted to explore the association of BF% and habitual macronutrient intake.

Results: Out of all the macronutrients explored habitual dietary fibre intake was the most strongly and inversely associated with BF% for both Pacific and NZE women. Higher energy adjusted habitual dietary fibre intake was associated with lower BF% (β= -0.35, $p \leq 0.001$), and this relationship became stronger after further adjustments for protein (g/day), total carbohydrate (g/day) and fat (g/day) intake (β= -0.47, $p \leq 0.001$). There was no difference in energy intake (kJ/day) between women in the highest (27.6 g/day [24.9, 30.6]) and lowest tertile (16 g/day [13.3, 17.6]) of dietary fibre intake; yet women in the highest tertile consumed a higher percentage of their total energy intake from protein, total fat, and total carbohydrate. In addition, women in the highest tertile of dietary fibre intake were older, with lower deprivation levels and had lower visceral fat (27.8 % [19.2, 35.4] versus 35.3 % [24.7, 41.3]) and android fat (29.7 % [22.4, 38.1] versus 37.8 % [27.4, 42.5]) percentage in comparison to women in the lowest tertile.

Conclusions: In a healthy population of Pacific and NZE women with different metabolic disease risks, higher dietary fibre intake was associated with lower BF% (including visceral and android fat), and higher nutrient intake. This suggests that higher habitual dietary fibre intake is a marker of food choice, which is associated with a reduced risk of metabolic disease. These

results support current dietary guidelines advocating for higher habitual dietary fibre intake from nutrient dense foods. Further research is required to explore what foods are contributing to the dietary fibre intake, and the causal mechanism of this protective effect.

4.2. Introduction

Obesity and associated non-communicable diseases (NCDs) are a global epidemic (1,2). New Zealand (NZ) has one of the highest rates of obesity in the world, where one in three people over the age of 15 are obese (3), ranking NZ third highest among the Organization for Economic Co-operation and Development (OECD) countries (1). Within NZ's culturally diverse population, obesity rates vary by ethnic group; for example, 29 % of NZ Europeans (NZE), 48 % of indigenous Māori, and 67 % of Pacific people were classified as obese (body mass index (BMI) ≥30 kg/m^2) in the 2018/2019 annual NZ health survey (3). Similar to global trends (4,5), in NZ, higher levels of deprivation are associated with an increased vulnerability to developing obesity and women are more likely to be obese compared to men (3).

Obesity manifests as a complex interaction between physiological, genetic, environmental, and behavioural factors across the lifespan, leading to the excess accumulation of fat mass (6). Excess adipose tissue can impair metabolic function (2), with increased metabolic risks associated with higher abdominal fat which surrounds vital organs (i.e., android and visceral fat) (7,8). Obesity is a consequence of the disruption to a complex energy regulation system (6), and overnutrition and sedentary behaviours are considered principle factors driving this disruption (9). The subsequent energy imbalance arises when energy consumed exceeds the energy expended for a prolonged period of time, facilitating the accumulation of excess adipose tissue. Thus, diet is considered a modifiable risk factor for NCDs such as type 2 diabetes mellitus (T2DM) and obesity (10–12). Further, in contrast to other risk factors, diet affects all people, regardless of age, gender, ethnicity, or socioeconomic status (SES) (12).

Across 195 countries worldwide, the 2017 Global Burden of Disease Study explored the effect of food consumption on NCDs and mortality. The authors emphasised how suboptimal dietary intakes, defined as inadequate intake of core or 'healthy' foods and nutrients (e.g., wholegrains, fruit, vegetables, dietary fibre), and excessive intake of non-essential 'discretionary' components (e.g., red and processed meat, sugar sweetened beverages, sodium), cause more deaths than any other risk factor for NCDs, including tobacco smoking (12). Diets high in sodium, sugar, and saturated fat (SFA), typically consumed in the form of highly processed, refined energy-dense and nutrient poor 'discretionary foods', are key

features of a suboptimal 'Western' diet (11). Overconsumption of these non-essential components and foods likely displaces 'core' foods, thereby driving the increased burden of NCDs, in particular obesity and T2DM (10,12,13). Furthermore, the report proposed that improvements in diet quality, by consuming 'optimal' levels of healthy foods and nutrients, could potentially prevent one in every five deaths globally (12).

Increasing evidence suggests there are key drivers and many ways to consume a healthy diet; for example, diets characterised by high intake of fruits, vegetables, wholegrains, legumes, and moderate intake of dairy and alcohol are associated with a more favourable weight and reduced risk of obesity (11,14). Across populations, these key drivers are shaped by, and relevant to, sociocultural and nutrition practices, including SES. However, globally, the key drivers have consistently been associated with higher quality diets and a reduced risk of NCDs, obesity, and mortality (11,12,14–16). Habitually consuming more core foods such as wholegrains, unrefined carbohydrates, and plant based foods, which are characterised by higher essential nutrients (e.g., dietary fibre), is protective against and inversely associated with components of the metabolic syndrome (e.g., abdominal obesity, high blood pressure, and insulin resistance) (14,17–20). A recent meta-analysis by Reynolds et al. (21) further emphasised the importance of the quality of the food source consumed, as higher dietary fibre intake was associated with a reduced risk of NCDs and all-cause mortality. The authors recommended adults aspire to consume at least 25g of dietary fibre a day from quality carbohydrate foods, whilst consuming more than 30 g/day would confer additional health benefits such as reducing the risk of cardiometabolic diseases (21).

The quality and quantity of foods habitually consumed together directly influences health outcomes. A recent meta-analysis by Seidelmann et al. (22) highlighted that habitually consuming 50-55 % of energy from carbohydrates was associated with decreased risk of all-cause mortality, with high (≥70 % energy) and low (≤40 % energy) carbohydrate intake increasing the risk. Seidelmann et al. (22) observed that replacing the energy intake from carbohydrates with animal based proteins and fats further increased the risk, in contrast to switching to plants based proteins and fats which decreased the risk of mortality. Consuming a balanced energy intake from quality food sources (e.g., the key drivers) can reduce the risk of mortality and poor health outcomes (21,22). However, reference guidelines such as the

acceptable macronutrient distribution ranges (AMDRs), which are recommended ranges of energy intakes to consume from macronutrients (expressed as a percentage of energy intake), have long been established and promoted by health agencies and professionals (23). The AMDRs complement food based dietary guidelines and the cut-offs values can differ between countries (23,24); however, they are evidenced based guidelines to reduce the risk of chronic illness and disease, while ensuring adequate micronutrient intake for healthy living (23).

The causes of obesity are complex and multifactorial. Considering habitual dietary intake directly influences health outcomes and risk of developing NCDs globally, it does pose the question of whether habitual diet is a driver of NZ's high obesity rates. Further, as the prevalence of obesity varies across the different ethnic groups in NZ, are there different dietary drivers of obesity for these ethnicities? The aim of this present study was to explore habitual dietary intake in a population of healthy NZ women (Pacific and NZE) aged 18-45 years with different metabolic disease risks and body weight profiles (lean and obese) and to investigate the relationship between habitual macronutrient intake and body fat content.

4.3. Materials and methods

4.3.1. Study design

The participants were part of the PRedictors linking Obesity and gut MIcrobiomE (PROMISE) cross-sectional study, which was conducted between July 2016 and September 2017, at the Human Nutrition Unit at Massey University in Albany, New Zealand. The PROMISE study aimed to characterise the gut microbiome and related parameters in two populations (Pacific and NZE women) with different metabolic disease risk and body weight profiles (lean and obese). This sub analysis aimed to further explore the dietary intake of these women. Eligible participants were healthy and free from any chronic illness or disease, and were post-menarche and premenopausal women, aged between 18 and 45 years, from the Auckland region. The study procedures and recruitment are reported in detail elsewhere (25). Briefly, participants attended two study visits and completed at home data collection between the study visits.

The PROMISE study was approved by the Southern Health Disability Ethics Committee (16/STH/32) and conducted according to the guidelines of the declaration of Helsinki. The trial was registered at anzctr.org.au (ACTRN12618000432213). All participants were provided with detailed information about the study procedures, study measurements, and gave written informed consent to participate.

4.3.2. Demographic and anthropometric information, blood sampling and laboratory analyses

At study visit one, trained research staff followed standard operating procedures (SOP) to conduct one-on-one interviews with participants to capture a range of demographic information (e.g., address, occupation, personal/household income). The New Zealand Deprivation 2013 index (NZDep2013), is an area based measure of socioeconomic deprivation, and combines information from the national census data including, housing, qualifications, employment, income, family structure, and access to transport (26). The NZDep2013 index was used to assign a socioeconomic deprivation score ranging between decile one "least deprived" and decile ten "most deprived" (26), and details of score calculation are discussed in the PROMISE protocol paper (25). Accredited research staff conducted anthropometric measurements (e.g., fasted weight, stretched height) according to the International Society for the Advancement of Kinanthropometry (ISAK) protocols (27). Fasting weight (kg) and height (cm) were used to calculate BMI kg/m^2 (weight (kg)/height (m^2)) using the Quetelet index. At study visit two, a whole-body scan was conducted with Dual-energy X-ray Absorptiometry (DXA) (Hologic QDR Discovery A, Hologic Inc, Bedford, MA with APEX V. 3.2 software) to assess total and regional body fat percentage (BF%) (e.g., visceral, android, and gynoid fat percentage).

At study visit one, following an overnight fast of at least 10 hours, fasting blood samples were collected by a trained phlebotomist. Each participant provided three fasting blood samples between 7:30 and 9am. To obtain plasma and serum, blood samples were collected into EDTA and silicone coated (with clot activator) vacutainer tubes, respectively. Immediately following blood collection, an aliquot of EDTA whole blood was frozen at -80°C for HbA1c analysis, and serum tubes were kept at room temperature for 30 minutes for clot formation. The remaining

samples were kept at 4°C, and within an hour of collection, all three vacutainer tubes were centrifuged at 3500 rpm for 15 minutes at 4°C, and plasma and serum samples were frozen at -80°C until further analysis.

Serum levels of glucose (enzymatic UV method), total cholesterol (TC), high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C) (enzymatic colorimetric method) were measured using a Hitachi c311 autoanalyser (Hitachi High Technologies Corporation, Tokyo, Japan) with Roche Diagnostic reagents (Mannheim, Germany) for all assays. Using EDTA whole blood, HbA1c levels were measured by turbidimetric inhibition immunoassay method (Roche Diagnostic, Mannheim, Germany) on a Hitachi c311 autoanalyser (Hitachi High Technologies Corporation, Tokyo, Japan). The average inter-assay coefficient of variation for all the above analyses performed using the Hitachi c311 autoanalyser was 1.5 %. Serum levels of insulin were measured using the electrochemiluminescence immunoassay method (Roche Diagnostics, Mannheim, Germany) on the Cobas e411 analyser (Hitachi High Technologies Corporation, Tokyo, Japan). The inter-assay coefficient of variation was 0.5 %. Homeostasis model assessment (HOMA-IR) index for insulin resistance was calculated (fasting blood glucose [mmol/L] X fasting plasma insulin [µU/mL]/22.5) (28). At rest, blood pressure was measured after a 10-minute (sitting) period with an Omron HEM 907 digital blood pressure monitor. The measurement was repeated three times and the average of the second and third measurement was recorded.

4.3.3. Dietary assessment

Participants watched a ten-minute video on how to complete their five-day, non-consecutive, estimated food record (5DFR), at home. Trained researchers further clarified the food recording procedure and answered any participant questions. A standardised portion guide booklet, and standard household measures (e.g., metric cup), were provided to aid participants in completing the 5DFR. During their second study visit, all 5DFRs were reviewed by a dietitian, followed by an individual interview to clarify the food recorded, food preparation and cooking methods used, portion sizes, and brands of foods consumed, or any potential ambiguities identified to inform accurate dietary data entry. Participants also completed a validated semi-quantitative, NZ Women's Food Frequency Questionnaire

(NZWFFQ) (29) regarding the past thirty days food intake. The 220-item NZWFFQ provided standard portion sizes (e.g., 1 teaspoon of sugar, 1 slice of bread) as a reference for participants, and nine standard frequency categories (ranging from 1 = "never" to 9 = "4 or more times a day") to enable intake estimation. The NZWFFQ was hosted by SurveyMonkey© software (SurveyMonkey Inc, San Mateo, California, USA) online, which enabled the research staff the ability to monitor live progress.

4.3.4. Dietary data processing

Energy, macro- and micronutrient analyses of the 5DFR and NZWFFQ were completed using FoodWorks9 (Xyris Software (Australia) Pty Ltd, Queensland, Australia). FoodWorks9 hosts multiple food composition databases including NZ's food composition database (FOODFiles 2016), developed by the NZ Institute for Plant & Food Research and the NZ Ministry of Health. Trained research staff entered the 5DFR and NZWFFQ independently into the Foodworks9 software, using a SOP that was developed to ensure consistent and reliable dietary data entry. Researchers utilised a hierarchy of food choices from the different food composition databases. If a direct match to the reported food item could not be found in NZ FOODFiles 2016, then the Xyris brandname database AusFoods 2017 and AusBrands 2017 (based on the Australian food composition databases AUSNUT 2011-13, developed by Food Standards Australia New Zealand) were used. During dietary data entry, a "PROMISE dietary data assumption dictionary" was generated, detailing the reported food in a participants 5DFR, and the selected option within the software. This dictionary ensured standardised dietary data entry for all participant's in the PROMISE study.

Food items that did not have a direct compositional match within the FoodWorks9 databases were discussed at weekly meetings with the nutrition research team to resolve data entry. For example, if a participant was unable to report an accurate recipe (all/quantities of ingredients, method of cooking) or reported purchased food items (e.g., savoury scone), then a "standard recipe" was created from existing analysed foods within the databases. For a participant recipe (e.g., chocolate cake with a direct compositional match in the database), the existing analysed recipe was selected and entered as "chocolate cake" from NZ FOODfiles2016. Where appropriate, all food items were entered in the software as the cooked portion consumed.

Cooking factors were used from McCance and Widdowson's (30), to convert the raw weights of the ingredients reported in participants meals, to allow for water and nutrient losses in the cooking process.

To ensure consistent data entry for all NZWFFQs, a standardised template of the NZWFFQs food items and corresponding portion sizes was created in the FoodWorks9 software. Participants had the option to report any additional foods they consumed habitually (which were not included in the NZWFFQ) in the free text space at the end of the NZWFFQ. Additional foods were entered manually into the template for each individual following the same SOP developed for the 5DFR. The nine standard frequency categories were converted to units of daily consumption (e.g., "never" = 0, "less than once a month" = 0.01, "1-3 times a month" = 0.07, "once a week" = 0.14, "2-3 times a week" = 0.36, "4-6 times a week" = 0.71, "once a day" = 1, "2-3 times a day" = 2.5 to "4 or more times a day" = 4). All dietary data entry (5DFR and NZWFFQ) were independently checked by a second researcher to confirm reliability of data entry. All exported data sets were extensively reviewed for misreporting using the raw dietary data. Following extensive review of the raw and processed dietary data for plausibility of intake, the cut-offs of >2100 kJ/day and <27000 kJ/day were considered valid completion of the 5DFR and NZWFFQ. These cut offs are higher than Willet's epidemiological recommendations (31); however, were considered more reflective of the foods consumed by the study population (32–34). Reported intakes below or above these cut-offs were excluded from further dietary analysis.

Total energy intake is reported as kilojoules (kJ) and includes the energy contribution from total dietary fibre, all macronutrients and alcohol. Total dietary fibre (reported as grams (g)) includes all non-starch polysaccharides (cellulose, hemicellulose, and resistant starches) and NZFOODFILES uses AOAC Prosky method to measure dietary fibre). Total fat (reported in grams (g)) includes all unsaturated and saturated fats. Total carbohydrate (reported in grams (g)) includes starch, free sugars and glycogen and dextrins. To assess the composition of the total diet, the percentage of energy intake from all macronutrients was calculated (35) and compared to the AMDRs for NZ and Australia (23).

4.3.5. Calculating habitual dietary intake

For each participant with valid dietary data (within the specified kJ cut off range), individual habitual dietary intake of nutrients was estimated using the National Cancer Institute (NCI) method. The three NCI SAS macros (version 2.1 - available from the NCI website (36)) were adapted to the PROMISE dietary data. Utilising the 5DFR as the primary dietary data, the NCI method uses a two-part modelling approach to estimate the probability of consumption of a nutrient, food or food group, respectively, and the respective consumed amount, considering covariates that are affecting either the probability of consumption or the amount of consumed. For episodically consumed dietary components (e.g., components that are not consumed every day i.e., alcohol) the first part of the model estimates the probability of consumption using logistic regression model allowing for a residual between-person variation (covariates) to be included. The second part of the model uses a linear regression model on a transformed scale (as the NCI method requires near normal distribution of the residuals) to estimate the amount consumed, and, like the first model, considers covariates. The individual habitual intake was then defined as the product of probability of consumption x the consumed amount (*Usual intake = Probability x Amount*). Estimation of a dietary component consumed daily (e.g., energy (kJ)) requires a one-part model, where the process is the same as mentioned above, however there is no need to model probability of consumption, so this step was omitted.

In the current analyses, covariates age, ethnicity, BMI, season, NZWFFQ information (in standard units/day), and weekend information (Weekday = Monday-Thursday, weekend = Friday-Sunday) were considered. If a participant reported never consuming a nutrient in both the NZWFFQ and 5DFR they were considered a “true non-consumer” and were assigned 0 g/day for the nutrient. The average daily intake of 36 nutrients (in units/day) consumed for one month (the timeframe of the NZWFFQ) was calculated for each participant.

4.3.6. Statistical analysis

Participants were selectively recruited by BMI; however, to control for the fact individuals with the same BMI can have different metabolic disease risks (37,38), participants were stratified into low and high BF% groups using the median as the cut point; that is, low-BF% (<35 %)

versus high-BF% (≥35 %) (37,39,40). Measurements that did not follow a normal distribution were reported as medians [25th, 75th percentiles] and non-parametric tests were used to assess the differences between groups. Analyses were conducted using SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). The NZDep2013 index was collapsed into quintiles (e.g., quintile one = decile one and two, quintile two = decile three and four) and these quintiles were used for subsequent analyses. Analyses were conducted separately for NZE and Pacific participants, as well as both groups combined. Comparisons of participant characteristics focused on the differences between BF% groups within an ethnic group (e.g., NZE low-BF% versus NZE high-BF%), and between ethnicities within a BF% group (e.g., NZE low-BF% versus Pacific low-BF%).

To assess the association between diet (macronutrient intake) and BF%, BF% was used as the dependent variable and macronutrients were used as the independent variables. Firstly, univariate linear regression analyses were conducted followed by multivariate linear analyses controlled for potential confounders including total dietary fibre, protein, total carbohydrate, total fat, energy intake, age, and NZDep2013. Due to the design of the study (i.e., selectively recruited based on ethnicity) we adjusted combined analyses for ethnicity. All independent variables were assessed for collinearity by assessing tolerance and the variance inflation factor (VIF). No collinearity was detected. To assess the association of habitual dietary fibre intake and participants' characteristics, all participants were stratified by tertiles of energy adjusted dietary fibre intake. Higher energy intake correlates with higher nutrient intake, therefore adjusting nutrient intake for total energy intake allows for the evaluation of diet composition/nutrient intake and health outcomes independent of energy intake. Comparisons focused on the lowest tertile of dietary fibre intake (lowest 33 % of the population) versus the highest tertile of dietary fibre intake (tertile three). *P*-values <0.05 were considered statistically significant.

4.4. Results

4.4.1. Characteristics of the study participants

A total of 351 women were eligible to participate in the PROMISE study, and a total of 304 women completed all aspects of the study. Seventeen women were excluded prior to calculation of the NCI dietary data based on misreporting (>27000 kJ/day) their energy intake on the NZWFFQ. We therefore had complete data for 287 participants: 126 (44 %) Pacific and 161 (56 %) NZE women, with a median age of 23 (18-44) and 32 (18-45) years respectively (see Table 4.1).

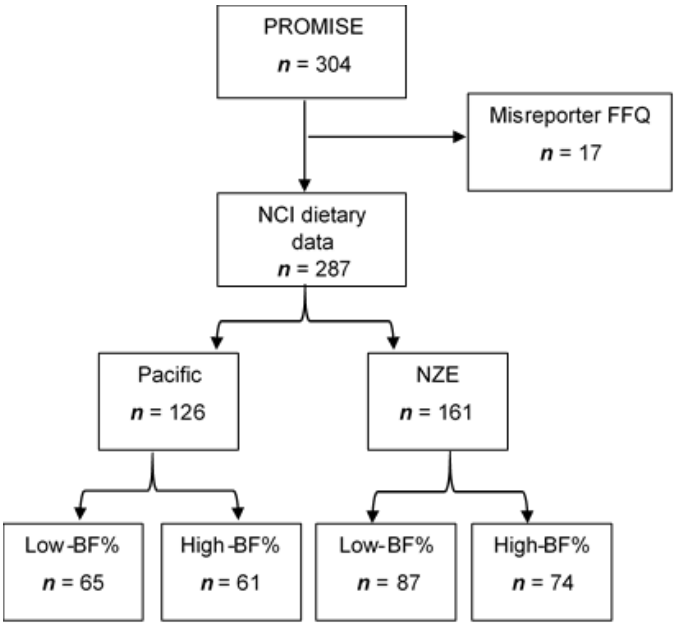


Figure 4.1. Overview of participant's included in the study

NZE women were older and had lower NZDep2013 scores in comparison to Pacific women. Between BF% groups there was no significant difference in TC or LDL-C for Pacific women; however, NZE women in the high-BF% group had higher TC and LDL-C in comparison to NZE women in the low-BF% group. Regarding the metabolic health markers assessed, Pacific women in the low- and high-BF% groups were within the recommended ranges (41); however, NZE women in the high-BF% were slightly above the reference ranges for TC and fasting glucose in comparison to NZE women in the low-BF% group. In the low-BF% group, NZE women had lower BMI, BF%, visceral (VAT) and android fat (AF%) percentage, including lower glycosylated haemoglobin (HbA1c), fasting glucose, insulin and HOMA-IR index, compared to Pacific women. In contrast, in the high-BF% group, NZE women had higher BF% and gynoid fat percentage (GF%) in comparison to Pacific women; however, NZE women had lower HbA1c, fasting insulin and HOMA-IR index.

Table 4.1. Characteristics of Pacific and NZE women stratified by body fat percentage groups

	All participants	Pacific		NZE	
		Low-BF%	High-BF%	Low-BF%	High-BF%
	n=287	n=65	n=61	n=87	n=74
Age (y)	28 [22, 35]	23 [20, 29]	23 [21, 29]	29 [24, 36]π	35 [28, 40]*π
Body composition					
Weight (kg)	77.6 [65.6, 96.0]	72.4 [67.3, 79.1]	97.0 [87.4, 109.9]*	62.4 [58.1, 66.6]π	94.1 [86.8, 101.7]*
Height (cm)	167.7 [163.7, 172.2]	168.7 [164.7, 174.2]	168.4 [163.3, 173.4]	167.7 [163.9, 170.7]	166.7 [162.5, 171.9]
BMI (kg/m^2)	28.0 [23.0, 33.4]	25.0 [23.6, 27.6]	33.8 [31.1, 39.9]*	22.5 [20.9, 23.5]π	33.5 [31.7, 36.3]*
Body fat (%)	34.6 [28.8, 39.6]	29.6 [27.9, 32.3]	39.5 [36.6, 42.4]*	28.0 [24.2, 31.9]π	40.3 [38.7, 44.2]*π
Visceral fat (%)	32.3 [23.7, 38.9]	26.8 [23.1, 31.4]	40.3 [35.6, 43.3]*	21.5 [16.8, 27.3]π	39.7 [35.7, 44.0]*
Android fat (%)	34.3 [26.1, 40.7]	29.1 [25.6, 33.5]	41.2 [38.0, 44.8]*	24.2 [20.2, 29.1]π	41.5 [38.0, 45.7]*
Gynoid fat (%)	37.3 [34.1, 41.8]	35.1 [32.7, 37.1]	41.0 [38.5, 43.7]*	34.3 [30.5, 36.7]	42.5 [39.8, 45.1]*π
Blood pressure €					
SBP (<130 mm Hg)+	115 [106, 123]	113 [106, 119]	117 [111, 128]*	113 [105, 119]	120 [111, 129]*
DBP (<80 mm Hg)+	74 [68, 80]	71 [65, 74]	77 [71, 84]*	69 [66, 76]	80 [74, 85]*
Metabolic markers ∞					
TC (<5 mmol/L)+	4.8 [4.3, 5.4]	4.5 [4.1, 5.1]	4.6 [4.2, 5.1]	4.9 [4.3, 5.4]π	5.2 [4.7, 6.1]*π
HDL-C (>1 mmol/L)+	1.6 [1.3, 1.9]	1.5 [1.3, 1.8]	1.3 [1.2, 1.6]*	1.8 [1.6, 2.0]π	1.4 [1.3, 1.7]*π
LDL-C (0-3.4 mmol/L)+	3.0 [2.4, 3.5]	2.8 [2.4, 3.2]	3.0 [2.5, 3.3]	2.8 [2.4, 3.4]	3.4 [2.7, 4.1]*π
TG (<2 mmol/L)+	0.9 [0.7, 1.2]	0.8 [0.7, 1.1]	1.0 [0.9, 1.5]*	0.7 [0.6, 0.9]	1.1 [0.8, 1.5]*
TC:HDL-C (<4 mmol/L)	3.1 [2.6, 3.7]	2.9 [2.5, 3.3]	3.4 [2.9, 4.0]*	2.7 [2.3, 3.1]π	3.5 [3.0, 4.3]*
HbA1c (<40 mmol/mol)+	31.7 [30.0, 33.6]	32.1 [30.5, 33.8]	34.8 [32.3, 36.7]*	30.6 [29.0, 31.9]π	31.0 [29.8, 33.3]*π
Glucose (3.5-5.4 mmol/L)+	5.3 [5.0, 5.6]	5.3 [5.0, 5.5]	5.4 [5.1, 5.9]*	5.1 [4.9, 5.3]π	5.5 [5.1, 5.7]*
Insulin (3-25 uU/mL)+	10.9 [7.3, 17.3]	11.2 [7.9, 16.0]	21.4 [13.1, 31.9]*	7.1 [5.2, 8.7]π	12.6 [10.0, 17.9]*π
HOMA-IR	2.6 [1.7, 4.0]	2.6 [1.9, 3.6]	4.8 [3.1, 8.5]*	1.6 [1.1, 2.1]π	3.1 [2.3, 4.1]*π
Deprivation Index#	6 [3, 8]	7 [5, 9]	8 [7, 9]	3 [2, 6]π	5 [3, 6]*π

All values are reported as medians [25th, 75th percentiles]. Mann Whitney statistical test used to identify a significant difference (p<0.05). *Statistically significant difference between body fat groups within ethnic group. πStatistically significant difference between ethnic groups within body fat group. €NZE women (n=2), ∞Pacific woman (n=1) and #Pacific women (n=2) have not been included in analysis due to missing data. BF% = Total Body Fat Percentage (% measured with DXA), low-BF%: BF% <35 %, high-BF%: BF% ≥35 %. NZE: New Zealand European, BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, TC: total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TG: Triglycerides, TC:HDL ratio: total cholesterol/high density lipoprotein cholesterol, HbA1c: glycosylated haemoglobin. Deprivation Index NZ 2013. +Normal healthy ranges for metabolic biomarkers (41)

4.4.2. Habitual dietary intake

The NCI calculated individual habitual intake of 36 nutrients for all 287 participants represents a period of one month, respective numbers are presented as a median daily intake (in units/day) Table 4.2.

4.4.3. Nutrient intake

There was no difference in nutrient intake between Pacific women in the low- and high-BF% groups (see Table 4.2). However, NZE women in the high-BF% group consumed more energy (kJ/day), yet less PUFA (g/day), dietary fibre (g/day), and alcohol (g/day) in comparison to NZE women in the low-BF% group. In the low-BF% group NZE women consumed less energy, total carbohydrate (g/day), sugar (g/day), and starch (g/day), yet more PUFA, total cholesterol (g/day), dietary fibre and alcohol in comparison to Pacific women in the low-BF% group. In the high-BF% group, NZE women consumed more PUFA, dietary fibre and alcohol, and less total carbohydrate and starch in comparison to Pacific women (see Table 4.2).

4.4.4. Percentage of energy intake from nutrients

There was no difference in percentage of total energy intake consumed from the nutrients between Pacific women in the low- and high-BF% groups (see Table 4.2). However, NZE women in the low-BF% group consumed a higher percentage of total energy from PUFA and alcohol in comparison to NZE women in the high-BF% group. In the low-BF% group, NZE women consumed a higher percentage of total energy intake from protein, PUFA and alcohol; yet a lower percentage of energy from total carbohydrate intake in comparison to Pacific women (see Table 4.2). In the high-BF% group, NZE women consumed a higher percentage of total energy intake from PUFA and alcohol, and a lower percentage of total energy from total carbohydrate in comparison to Pacific women. In addition, when comparing to the AMDR, the median percentage of protein intake was within the AMDRs (15-25 % of total energy intake) for Pacific and NZE women. Whereas the median percentage of total carbohydrate intake was below the AMDRs (45-65 % of total energy intake), in contrast to the percentage of total and SFA intake, which were both above the AMDRs (20-35 % and <10 % of total energy intake, respectively) for Pacific and NZE women.

Table 4.2. Habitual nutrient intake of Pacific and NZE women stratified by body fat percentage groups

	AMDR % of energy intake	Pacific		NZE	
		Low-BF% n=65	High-BF% n=61	Low-BF% n=87	High-BF% n=74
Energy (kJ/day)		8547.5 [8132.8, 8854.7]	8748.6 [8404.5, 8985.7]	8555.4 [8045.2, 8894.4]	8307.3 [8032.8, 8659.9]π
Protein (g/day)		82.4 [73.2, 91.6]	77.9 [69.5, 91.6]	80.2 [71.3, 92.7]	83.0 [76.1, 91.0]
Protein (E%/day)	15-25 %	16.3 [14.3, 18.1]	15.1 [13.5, 17.2]	15.9 [14.0, 18.4]	16.7 [15.4, 18.1]π
Total fat (g/day)		89.3 [75.7, 102.3]	90.6 [76.7, 102.0]	89.6 [70.8, 98.7]	89.4 [75.7, 104.2]
Total fat (E%/day)	20 -35 %	39.6 [34.4, 45.6]	39.2 [33.7, 45.1]	39.1 [31.5, 44.0]	41.1 [35.7, 47.1]
SFA (g/day)		33.7 [28.7, 39.9]	33.7 [29.7, 40.7]	35.0 [27.2, 39.2]	32.5 [26.6, 39.9]
SFA (E%/day)	<10 %	15.1 [12.8, 17.4]	14.9 [12.8, 17.4]	15.5 [12.8, 16.9]	14.8 [12.4, 17.4]
PUFA (g/day)		12.3 [9.9, 14.3]	12.3 [9.3, 13.9]	10.7 [8.6, 12.9]	13.3 [10.9, 15.4]π
PUFA (E%/day)		5.4 [4.3, 6.4]	5.3 [4.1, 6.0]	4.7 [3.9, 5.8]	6.0 [5.0, 7.0]π
MUFA (g/day)		33.7 [28.3, 38.6]	33.9 [28.7, 38.0]	33.2 [27.7, 37.1]	34.7 [28.7, 39.6]
MUFA (E%/day)		14.8 [12.7, 17.2]	14.9 [12.7, 16.4]	14.3 [12.3, 17.0]	15.6 [13.2, 17.7]
Cholesterol (g/day)		294.6 [241.9, 363.2]	269.0 [230.5, 356.5]	274.3 [223.3, 332.9]	296.6 [250.4, 363.2]π
CHO (g/day)		194.3 [163.0, 227.0]	212.4 [180.0, 246.6]	213.3 [177.7, 238.9]	181.7 [152.2, 202.6]π
CHO (E%/day)	45-65 %	37.9 [32.4, 44.4]	40.6 [34.1, 47.5]	42.7 [33.8, 45.6]	35.5 [30.4, 40.3]π
Sugar (g/day)		80.7 [67.4, 95.3]	82.5 [74.0, 102.7]	78.9 [66.1, 96.2]	80.2 [66.0, 89.5]π
Starch (g/day)		110.2 [91.5, 132.2]	125.7 [105.1, 145.2]	128.1 [109.1, 146.3]	102.5 [82.0, 119.0]π
Dietary Fibre (g/day)		20.6 [17.5, 24.9]	18.8 [15.6, 22.1]	17.8 [15.0, 20.8]	23.7 [20.1, 29.9]π
Alcohol (g/day)		0.3 [0.1, 0.9]	0.2 [0.1, 0.5]	0.1 [0.1, 0.3]	0.8 [0.2, 3.9]π
Alcohol (E%/day)		0.1 [0.0, 0.3]	0.1 [0.0, 0.2]	0.0 [0.0, 0.1]	0.3 [0.1, 1.4]π

All values are reported as medians [25th, 75th percentiles]. Habitual nutrient intake calculated for a period of one month with the National Cancer Institute method (36). Mann Whitney statistical test used to identify a significant difference ($p<0.05$). *Statistically significant difference between body fat groups within ethnic group. πStatistically significant difference between ethnic groups within body fat group. BF% = Total Body Fat Percentage (% measured with DXA), low-BF%: BF% <35 %, high-BF%: BF% ≥35 %, NZE: New Zealand European, SFA = Saturated fat, PUFA = Polyunsaturated fat, MUFA = Monounsaturated fat, CHO = total carbohydrate. AMDR: acceptable macronutrient distribution ranges (23).

4.4.5. Association of nutrient intake and body fat percentage

Higher habitual dietary fibre intake was inversely associated with BF% for Pacific and NZE women (Table 4.3). This association became stronger following adjustments for the intake of the other nutrients (total carbohydrate, total fat, protein, and energy intake), for both the stratified (within ethnicity) and combined (Pacific and NZE adjusted for ethnicity) analyses. In both the stratified and combined analyses, only after adjustment for the other nutrients was higher total carbohydrate intake positively associated with BF% for Pacific and NZE women. In addition, only following adjustment for intake of the other nutrients was higher protein intake positively, and higher total fat intake inversely associated with BF% for Pacific women. Adjustments for alcohol intake did not alter the results (for either stratified or combined analyses). Considering habitual dietary fibre was strongly associated with BF% for both Pacific and NZE women, all participants were stratified into tertiles of energy adjusted dietary fibre intake to further explore associations with metabolic health markers.

Table 4.3. The association of habitual nutrient intake and total body fat percentage for Pacific and NZE women

Nutrient	Pacific n=126		NZE n=161		All participants≠ n=287	
	Model 1 β [95 % CI]	Model 2 β [95 % CI]	Model 1 β [95 % CI]	Model 2 β [95 % CI]	Model 1≠ β [95 % CI]	Model 2≠ β [95 % CI]
Total carbohydrate	0.01 [-0.02, 0.03]	0.04 [0.01, 0.07]*	0.02 [-0.01, 0.06]	0.05 [0.01, 0.08]*	0.01 [-0.01, 0.03]	0.04 [0.02, 0.06]*
Total fat	-0.03 [-0.09, 0.03]	-0.10 [-0.19, -0.01]*	-0.04 [-0.10, 0.03]	0.001 [-0.07, 0.07]	-0.03 [-0.08, 0.01]	-0.03 [-0.09, 0.02]
Protein	0.03 [-0.04, 0.10]	0.14 [0.04, 0.23]*	-0.003 [-0.11, 0.10]	0.04 [-0.07, 0.16]	0.01 [-0.05, 0.07]	0.07 [-0.002, 0.14]
Dietary fibre	-0.24 [-0.47, -0.01]*	-0.48 [-0.78, -0.18]*	-0.40 [-0.58, -0.22]**	-0.49 [-0.69, -0.29]**	-0.35 [-0.49, -0.21]**	-0.47 [-0.62, -0.31]**

Total n=284: Pacific n= 124 (n=2 with missing deprivation data), NZE n= 160 (n=1 with missing deprivation data) *P value <0.05, *P value ≤0.01, **P value ≤0.001. NZE: ≠Adjusted for ethnicity. New Zealand European. Univariate: Model1: body fat(%)=β1*nutrient (unit) (adjusted for: + age+ NZDep2013 + energy intake (kJ/day)). Multivariate: Model2: body fat(%)=β1*nutrient1 + β2*nutrient2+ β3*nutrient3+ β4*nutrient4 (adjusted for: + age+ NZDep2013 + energy intake (kJ/day))

4.4.6. Characteristics of participants stratified by tertiles of dietary fibre intake

Women in the highest tertile of habitual dietary fibre intake were older with lower deprivation levels, and had a lower body weight, BMI, and BF% (including lower VAT, AF%, and GF%) in comparison to women in the lowest tertile. These women also had higher fasting HDL-C, including lower HbA1c and TC/HDL-C ratio, in comparison to women in the lowest tertile. Women in the lowest tertile of dietary fibre intake had higher fasting insulin and HOMA-IR in comparison to women in the highest tertile of intake (Table 4.4).

4.4.7. Nutrient intake of participants stratified by tertiles of dietary fibre intake.

The estimated median intake for women in the highest tertile of dietary fibre intake was 27.6 g/day [24.9, 30.6], which was higher than women's median and lowest tertile of intake (20.6 [19.6, 22.0] and 16 g/day [13.3, 17.6], respectively) (see Table 4.5). Women in the median tertile of dietary fibre intake consumed more energy than women in the lowest tertile of dietary fibre intake; however, there was no difference in comparison to women in the highest tertile. Unexpectedly, there was no difference in energy intake between the highest and lowest tertile of dietary fibre intake. Despite this finding, there was a positive association between dietary fibre intake and the intake of protein and total fat (including PUFA and MUFA) intake, as well as the total percentage of energy intake consumed from these nutrients. In addition, women in the lowest tertile of dietary fibre intake consumed less SFA, total carbohydrate and percentage of total energy consumed from these nutrients, including less starch, sugar, and alcohol intake, where there was no difference between women in the median and highest tertiles of intake.

Table 4.4. Characteristics of Pacific and NZE women stratified by tertiles of energy (kJ) adjusted dietary fibre intake

	Tertile 1 n=95	Tertile 2 n=96	Tertile 3 n=96
Ethnicity n (%)			
Pacific	67 (71 %)	39 (41 %)	20 (21 %)
NZE	28 (29 %)	57 (59 %)	76 (79 %)
Age (y)	23 [21, 29]*	29 [23, 37]	30 [25, 35]+
Body composition			
Weight (kg)	85.3 [68.2, 100.5]*	76.7 [66.2, 92.1]	70.1 [63.2, 92.5]+
Height (cm)	168.4 [163.5, 172.5]	166.5 [162.7, 172.1]	168.1 [164.8, 172.0]
BMI (kg/m^2)	31.1 [24.0, 35.8]	27.4 [23.4, 33.3]	25.0 [22.2, 31.9]+
Body fat (%)	35.4 [29.4, 40.1]	35.2 [29.6, 39.7]^	32.1 [26.4, 38.9]+
Low-BF% <35 %	44 (46 %)	47 (49 %)	61 (64 %)
High-BF% ≥35 %	51 (54 %)	49 (51 %)	35 (36 %)
Visceral fat (%)	35.3 [24.7, 41.3]	33.3 [26.0, 39.6]^	27.8 [19.2, 35.4]+
Android fat (%)	37.8 [27.4, 42.5]	35.2 [28.1, 40.3]^	29.7 [22.4, 38.1]+
Gynoid fat (%)	37.8 [34.4, 41.9]	37.9 [35.4, 42.1]^	36.2 [31.5, 40.9]+
Blood pressure €			
SBP (<130 mm Hg)+	116 [107, 124]	115 [109, 123]	114 [106, 123]
DBP (<80 mm Hg)+	74 [67, 83]	74 [69, 80]	73 [68, 80]
Metabolic markers ∞			
TC (<5 mmol/L)+	4.6 [4.2, 5.4]	4.9 [4.3, 5.6]	4.8 [4.5, 5.5]
HDL-C (>1 mmol/L)+	1.4 [1.3, 1.7]*	1.6 [1.3, 1.9]	1.7 [1.4, 2.0]+
LDL-C (0-3.4 mmol/L)+	2.9 [2.4, 3.4]	3.0 [2.5, 3.5]	3.0 [2.4, 3.5]
TG (<2 mmol/L)+	0.9 [0.7, 1.2]	0.9 [0.7, 1.1]	0.8 [0.6, 1.2]
TC:HDL-C (<4 mmol/L)	3.2 [2.6, 3.9]	3.1 [2.7, 3.6]	2.9 [2.4, 3.5]+
HbA1c (<40 mmol/mol)+	32.2 [30.5, 35.0]	31.8 [29.5, 33.7]	31.2 [29.6, 32.6]+
Glucose (3.5-5.4 mmol/L)+	5.4 [5.1, 5.6]	5.3 [5.0, 5.7]	5.2 [4.9, 5.5]+
Insulin (3-25 uU/mL)+	13.6 [9.1, 21.4]*	10.7 [7.8, 17.4]^	8.7 [6.0, 13.2]+
HOMA-IR	3.3 [2.1, 5.3]*	2.5 [1.8, 4.1]^	2.1 [1.3, 3.1]+
Deprivation Index#	8 [4, 9]*	5 [3, 8]^	5 [2, 6]+

All values are reported as medians [25th, 75th percentiles]. Mann Whitney statistical test used to identify a significant difference between tertiles ($p<0.05$). *Statistically significant difference between tertile 1 and 2 +Statistically significant difference between tertile 1 and 3. ^Statistically significant difference between tertile 2 and 3. €NZE women (n=2), ∞Pacific woman (n=1) and #Pacific women (n=2) have not been included in analysis due to missing data. NZE: New Zealand European, BF% = Total Body Fat Percentage (% measured with DXA), BMI: body mass index, SBP: systolic blood pressure, DBP: diastolic blood pressure, TC: total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TG: Triglycerides, HbA1c: glycosylated haemoglobin. Deprivation Index NZ 2013. Daily total dietary fibre intake calculated for a period of one month with the National Cancer Institute method (36).

Table 4.5. Habitual nutrient intake of Pacific and NZE women stratified by tertiles of energy (kJ) adjusted dietary fibre intake

	AMDR % of energy intake	Tertile 1 n=95	Tertile 2 n=96	Tertile 3 n=96
Energy (kJ/day)		8436.5 [8041.3, 8801.5]*	8652.5 [8182.3, 8898.1]	8492.9 [8192.3, 8837.0]
Protein (g/day)		75.4 [67.8, 84.5]*	83.8 [76.4, 91.6]^	90.2 [78.8, 97.1]+
Protein (E%/day)	15-25 %	14.7 [13.3, 17.0]*	16.5 [15.0, 18.1]^	17.4 [15.8, 19.1]+
Total fat (g/day)		79.7 [67.3, 90.9]*	91.6 [78.7, 102.1]^	99.1 [83.1, 112.7]+
Total fat (E%/day)	20 -35 %	35.9 [30.3, 40.8]*	40.3 [35.8, 45.2]^	44.2 [37.6, 49.1]+
SFA (g/day)		31.2 [25.2, 35.6]*	35.7 [30.0, 41.0]	35.2 [29.9, 42.2]+
SFA (E%/day)	<10 %	14.0 [11.3, 16.3]*	15.6 [13.4, 17.9]	15.4 [13.7, 18.6]+
PUFA (g/day)		9.8 [8.0, 11.4]*	12.3 [10.9, 13.8]^	14.4 [12.6, 17.0]+
PUFA (E%/day)		4.3 [3.6, 5.1]*	5.4 [4.7, 6.1]^	6.4 [5.6, 7.6]+
MUFA (g/day)		28.8 [26.2, 35.1]*	34.7 [29.4, 37.9]^	37.5 [30.3, 41.9]+
MUFA (E%/day)		13.5 [11.6, 15.7]*	15.1 [13.1, 17.0]^	16.4 [13.9, 18.4]+
CHO (g/day)		180.3 [145.5, 213.3]*	208.1 [172.6, 232.5]	194.2 [170.1, 226.4]+
CHO (E%/day)	45-65 %	34.6 [29.6, 42.8]*	39.1 [33.6, 45.1]	38.2 [33.5, 44.6]+
Sugar (g/day)		72.7 [59.4, 91.2]*	80.2 [70.9, 97.5]	85.6 [76.2, 95.8]+
Starch (g/day)		105.1 [84.4, 118.9]*	119.8 [96.7, 140.7]	113.4 [92.3, 132.6]+
Dietary Fibre (g/day)		16.0 [13.3, 17.5]*	20.6 [19.6, 22.0]^	27.6 [24.9, 30.6]+
Alcohol (g/day)		0.2 [0.1, 0.5]*	0.4 [0.2, 1.7]	0.4 [0.2, 1.4]+
Alcohol (E%/day)		0.1 [0.0, 0.2]*	0.1 [0.1, 0.6]	0.1 [0.1, 0.5]+

All values are reported as medians [25th, 75th percentiles]. Tertile 1: n=95 (Pacific n=67, NZE n=28), Tertile 2: n=96 (Pacific n=39, NZE n=57), Tertile 3: n=96 (Pacific n=20, NZE n=76), Mann Whitney statistical test used to identify a significant difference ($p<0.05$). *Statistically significant difference between tertile 1 and 2 +Statistically significant difference between tertile 1 and 3. ^Statistically significant difference between tertile 2 and 3. NZE: New Zealand European, SFA = Saturated fat, PUFA = Polyunsaturated fat, MUFA = Monounsaturated fat, CHO = total carbohydrate. Daily nutrient intake calculated for a eriod of one month with the National Cancer Institute method (36). AMDR: acceptable macronutrient distribution ranges (23).

4.5. Discussion

In the present study, habitual macronutrient intake was explored in relation to total BF% in a population of healthy Pacific and NZE women, whom have different metabolic disease risks and body fat profiles. There was no difference in macronutrient intake between Pacific women in the low- and high-BF% groups; however, for NZE women there were differences in energy, PUFA, dietary fibre, and alcohol intake between low- and high-BF% groups. In comparison to Pacific women, NZE women consumed less total carbohydrate (including starch and sugar), yet more dietary fibre, PUFA, and alcohol in both the low- and high-BF% groups. However, higher habitual dietary fibre intake was associated with lower total BF% for both Pacific and NZE women, and no other nutrient had such a strong association with BF% across all adjusted analyses. Further, unexpectedly, there was no difference in energy intake between women in the highest and lowest tertile of dietary fibre intake, yet women in the highest tertile consumed a higher intake (including percentage of total energy intake) of protein, total fat, and carbohydrate.

The observation of higher dietary fibre intake being associated with lower BF% agrees with the results of other prospective studies. For example, Ludwig et al. (42) observed in a large population of American women and men from the Coronary Artery Risk Development in Young Adults (CARDIA) study, that dietary fibre intake was a stronger predictor weight gain over a 10 year follow up in comparison to fat intake (42). Liu et al. (43) also observed that weight gain was inversely associated with higher dietary fibre intake over a 12 year follow up for women in the Nurses' Health Study Tucker et al. (44) similarly observed that in a population of Caucasian women, higher dietary fibre intake was associated with lower BF% and weight gain over the 20 month follow up period. However, Tucker et al. reported that higher fibre intake was associated with lower energy intake, and percentage of total energy consumed from total fat; whereas we observed no difference in energy intake between the low and high tertiles of dietary fibre intake, and the percentage of total energy intake consumed from total fat was positively associated with dietary fibre intake. Further, unlike Tucker et al. (44), the inclusion of energy intake in the analyses did not attenuate the observed association. Following the adjustments for energy and macronutrient intake, habitual dietary fibre explained more of the variation in BF%. Fibre can contribute to the weight and volume of food

without necessarily adding to the energy intake (e.g., lower energy density), thus may explain the observation of no difference in energy intake between the high and low tertiles of fibre intake (and why adjustment for energy intake had no effect). However, higher fibre intake was associated with higher nutrient intake, suggesting that although dietary fibre is strongly associated with BF%, it may indeed be a marker of food choice and thus diet quality (e.g., nutrient density) in this population.

When stratifying the study population by tertiles of energy adjusted dietary fibre intake, lower habitual dietary fibre intake was associated with higher adiposity: BF%, weight, BMI, VAT and AF%, including higher fasting insulin. This is an important observation as excess adipose tissue is a risk factor for hyperinsulinemia and insulin resistance, in addition to higher central adiposity being associated with increased metabolic risks (as the excess fat surrounds vital organs and can influence their function) (7,8). Previous studies have found that Pacific people have higher abdominal fat mass in comparison to NZE (45), which has been proposed to contribute to the high prevalence of diabetes within Pacific populations in comparison to Europeans (46,47). Thus, the observed inverse association of higher dietary fibre intake and lower VAT and AF% may, in part, reflect ethnic differences across the tertiles of dietary fibre intake. However, by design, the participants were selected to be healthy (free from any disease) and subsequently had HbA1c and fasting plasma glucose concentrations within the normal ranges (indicating euglycemia) (41). Hyperinsulinemia, in the absence of impaired glucose tolerance (e.g., fasting glucose above the normal range), is a risk factor for metabolic disease, as it precludes hyperglycaemia. Higher dietary fibre intake has been associated with lower insulin resistance in women, in part due to differences in BF% (48). Therefore, the observation that women in the lowest tertile of dietary fibre intake had the highest fasting insulin levels, suggests that lower habitual fibre intake is associated with, if not a risk factor for, hyperinsulinemia and higher BF%, which independently and combined are associated with increased metabolic health risks.

The health benefits associated with higher fibre intake have been attributed to the physiochemical and functional properties of dietary fibre (49,50). Higher fibre intake may reduce the risk of weight gain by reducing energy intake (by lowering the energy density of the diet). Additionally, increasing faecal bulk, can slow gastric emptying and rate of nutrient

absorption (51,52) which, in turn, can influence gastrointestinal hormones that influence appetite regulation and insulin secretion (49,50). Taken together, it aligns with higher habitual fibre intake being associated with lower weight gain and BF% over time (11,19,44,53). Dietary fibre can be classified as either indigestible polysaccharides (e.g., cellulose, hemicellulose, and resistant starch) or as insoluble and soluble fibre (49,50,54,55). However, dietary fibre is largely indigestible to human digestive processes in the small intestine. The host indigestible substrates reach the large colon where they are degraded and fermented by gut microbiota (the community of bacteria residing in the large bowel) (56,57). The gut microbiota are able to metabolise such complex substrates, including breaking down plant cell walls releasing phytochemicals, as they are equipped with carbohydrate degrading enzymes (56). Thus, although yet to be elucidated, emerging evidence suggests the beneficial effects of consuming plant-based diets (rich in key drivers and host indigestible substrates) may be mediated by the functions of the microbiota.

Higher fibre intake provides more substrates for the gut microbiota to ferment and products of microbial fermentation are metabolites called short chain fatty acids (SCFAs) (56,57). These SCFAs (acetate, propionate, and butyrate) influence host physiology via a wide variety of effects such as promoting satiety (via interactions with glucagon-like peptide-1 (GLP-1) and peptide YY), including improving glucose control (56,58). In small cohorts of overweight adults, supplementation of resistant starch has been associated with increasing first phase insulin secretion (59), and targeted delivery of an inulin-propionate-ester increased postprandial GLP-1 and peptide YY including decreasing rate of weight gain at the 24 week follow up (60). In addition, higher fibre intake has also been associated with lower weight gain at the 9 year follow up in a population of adult female twins (61). Within these studies the beneficial associations were attributed to the functions of the microbiota (59–61). Thus, the observation of higher fibre intake being associated with lower adiposity and fasting insulin may in part be mediated by the functions of the microbiota.

Only after adjustments for the intake of macronutrients was total carbohydrate intake positively associated with higher BF% for both Pacific and NZE women. However, these observations were relatively weak in comparison to the observed association with dietary fibre intake. An unexpected observation was the trend for the study population to consume

below the recommended AMDRs for total carbohydrate intake and over the AMDRs for total fat and SFA intake. It has been proposed that there is no ideal total carbohydrate to total fat ratio for the general population (62), due to the individualised responsiveness to diets with varying macronutrient ratios (63), which is thought to be driven in part by the gut microbiota (64). Nonetheless, the quality of the food sources (i.e., what the energy is replaced with) is perhaps more important than the ratio consumed (22). For example, the Mediterranean dietary pattern is associated with reduced risk of cardiometabolic disease and includes up to nearly 40 % of total energy intake from total fat (11). The important differentiation is that the source of the fats are unsaturated, as high saturated fat intake (>10 % of total energy intake) is associated with an increased risk of cardiovascular disease (CVD) (65,66). Thus, although total fat intake was high in this study population there was a positive trend of higher PUFA and MUFA intake associated with higher dietary fibre intake; albeit, the concerning factor is still the observed trend of high SFA intake given the increased risk of CVD. Potentially the high unsaturated fat intake, coupled with high dietary fibre intake, may reduce the risk; however, this is speculation and due to the design of the study we are unable to infer causality. What appears to be consistently more important is the quality of the food sources habitually consumed (more so than ratios) (10,11,21,22), and the strong public health message is to reduce and substitute saturated fat intake with unsaturated fat to decrease cardiometabolic health risks (11,65,66).

Individuals choose to eat foods not nutrients, and food choice is determined by accessibility and cost. Women in the lowest tertile of dietary fibre intake also had a higher level of deprivation. In NZ, living in a deprived area increases the risk of obesity by over 50 % (3), and decreases the likelihood of consuming the recommended daily servings of fruit and vegetables by 30 % (67), in comparison to individuals living in the least deprived areas after accounting for age, gender, and ethnicity. Higher quality diets are consistently associated with higher SES (68), and the high cost of fruit and vegetables (and nutrient dense foods) has long been cited as a barrier to consumption by NZ citizens (5,69,70). Individuals with limited financial resources will select the most financially viable option for themselves, which are often the inexpensive energy dense and refined food products (71). Although these results support the wealth of knowledge and public health messages (e.g., dietary guidelines (23)) which recommend consuming more nutrient rich foods rich in dietary fibre to reduce the risk of

cardiometabolic diseases (21), an individual's SES and the cost of food are always going to be key determinants of food choice, and thus habitual intake.

Despite having different metabolic disease risks, higher habitual fibre intake was associated with lower BF% for Pacific and NZE women after accounting for age, ethnicity, level of deprivation, energy, and macronutrient intake. With respect to the associations identified within this cross-sectional study, we are unable to predict an exact decrease in BF% attributable to a specific dietary fibre intake. Nor can we confirm the mechanism (s) of exactly how dietary fibre could be contributing to lower BF% in this population. However, we can postulate it could be related to consuming higher fibre foods which are of lower energy and higher nutrient density, where potentially the functions of the microbiota may be playing a role in stimulating appetite regulation.

There are several strengths to the current study. Primarily the detailed dietary data. Research personnel went to great lengths to address common critiques of self-reported dietary data throughout the collection and entering processes (31,72). Overall, this approach improved the validity of the data captured. In addition, the NCI method used to estimate habitual dietary data specifically addresses the intra individual day-to-day variation inherent in self-reported data (31). The habitual dietary data are based on actual intake (5DFR), considered person specific effects that can influence intake which, combined with the inclusion of NZWFFQ information, improves the estimation of intake (73). The results of these estimations for habitual intake are a smoother distribution of intake i.e., the NCI method reduces the long tails of a distribution curve which are often skewed by high and low intakes (non consumers). Most other studies that explore the relationship between dietary fibre intake and body weight profiles, rely on FFQs (42,43), and such tools are prone to recall bias and overestimation of intake (74,75). However, it must be considered that a FFQ typically reduces participant and researcher burden, and thus is favoured in large prospective studies which have the power to identify associations. In addition, all dietary components were considered within the models, as higher intake of one nutrient/dietary component often displaces others and vice versa (31). This study also had detailed information regarding the deprivation levels of participants, and adjusted analyses accordingly. Further, participants were selectively recruited based on BMI (lean and obese) and stratified by BF% to account for how individuals with similar BMIs can

have different metabolic disease risks (37,38). Multiple body composition measurements (e.g., weight, BMI, and BF% (and distribution)) and metabolic biomarkers (e.g., glucose and lipid profile) were assessed, which enabled the opportunity to explore their association with habitual diet to identify potential metabolic disease risks.

There are several limitations in the present study. The primary source of dietary information was self-reported dietary data, so results need to be interpreted with this in mind. However, self-reported dietary data offer insight into the complexities of what and how free-living individuals are eating that no set of biomarkers are currently able to do (76). The misreporting cut-offs of the wider PROMISE study were adapted and higher than the parameters recommended by Willet 2012 (31). However, higher cut-offs have been employed by other studies within ethnic minority groups (77,78), and the plausibility of energy intakes were extensively reviewed and considered reflective of the participants within the present study (32). Due to the cross-sectional design of the study results are not generalisable to other populations. Further, higher fibre intake may also be a marker of a healthier lifestyle in general (44), as factors other than diet can influence metabolic health profiles (e.g., physical activity levels) (48).

Taken together, the results of the present study support the current recommendations to consume 25-30 g/day of dietary fibre to reduce metabolic disease risk (21,23). However, the present study highlights the disconnect between public health messages and uptake by free-living individuals, evidenced by the observed deviations from the AMDRs. Suggesting as health researchers we need to more effectively translate our findings to accessible, realistic, and affordable behaviours. The variability and social contexts of how and when people eat, influences food choice and nutrition decisions must make sense in the context of people's lives. If we want to successfully tackle the obesity epidemic, we need to start treating the causes, not the symptoms. If adequate habitual total dietary fibre intake has the potential to reduce BF% and reduce the risk of NCDs, such as obesity (21), we need to explore how to improve dietary fibre intake, and thus diet quality at a population level. When considering the costs attributable to treatment and management of NCDs, prevention is a cheaper alternative (79).

4.6. Conclusion

In a population of Pacific and NZE women who have different metabolic disease risk and body fat profiles, higher dietary fibre intake was associated with lower BF%, including higher nutrient intake and lower levels of deprivation. Adjustments for macronutrient and energy intake strengthened the association with BF%, suggesting despite differences in metabolic disease risk a habitual diet characterised by higher fibre intake could reduce metabolic risks for these women. Further higher fibre intake was also associated with lower plasma insulin concentrations, VAT and AF%, suggesting the potential wider influence of higher fibre intake on metabolic health risk. Future longitudinal studies would be required to validate these observations in larger populations, ideally with multiple assessment time points to assess the inter- and intra-individual variation in diet and health outcomes. In addition, future work that explores dietary intake in totality should consider the combinations of nutrients consumed together, while accounting for the social determinants which influence food choice. Information generated from such trials could help equip health professionals with the knowledge of the foods (and combinations) to prescribe to promote better health outcomes. Until such time, with an appreciation of the determinants of food choice, it appears more important to focus on emphasising the quality of fibre rich foods consumed together to reduce the risk of NCDs such as obesity.

4.7. References

1. Organization for Economic Co-operation and Development (OECD). OECD Obesity update 2017. Paris, France; 2017.
2. World Health Organization. Obesity and overweight [Internet]. 2018 [cited 2019 Apr 12].
3. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].
4. Zhang Q, Wang Y. Socioeconomic inequality of obesity in the United States: Do gender, age, and ethnicity matter? Soc Sci Med. 2004 Mar;58(6):1171–80. doi:10.1016/S0277-9536(03)00288-0
5. The National Advisory Committee on Health and Disability. The social, cultural and economic determinants of health in New Zealand: Action to improve health a report from the National Advisory Committee on Health and Disability. Wellington; 1998.
6. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, et al. Obesity pathogenesis: An endocrine society scientific statement. Endocr Rev. 2017;38(4):267–96. doi:10.1210/er.2017-00111
7. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y, et al. Abdominal visceral and subcutaneous adipose tissue compartments. Circulation. 2007;116(1):39–48. doi:10.1161/circulationaha.106.675355
8. Després J-P. Body fat distribution and risk of cardiovascular disease. Circulation. 2012;126(10):1301–13. doi:10.1161/circulationaha.111.067264
9. Swinburn B, Sacks G, Vandevijvere S, Kumanyika S, Lobstein T, Neal B, et al. INFORMAS (International Network for Food and Obesity/non-communicable diseases Research, Monitoring and Action Support): Overview and key principles. Obes Rev. 2013;14(October):1–12. doi:10.1111/obr.12087
10. Ludwig DS, Hu FB, Tappy L, Brand-Miller J. Dietary carbohydrates: Role of quality and quantity in chronic disease. BMJ. 2018;361:k2340:1–6. doi:10.1136/bmj.k2340
11. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.
12. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8
13. Dahl WJ, Stewart ML. Position of the Academy of Nutrition and Dietetics: Health implications of dietary fiber. J Acad Nutr Diet. 2015;115(11):1861–70. doi:10.1016/j.jand.2015.09.003
14. Schlesinger S, Neuenschwander M, Schwedhelm C, Hoffmann G, Bechthold A, Boeing H, et al. Food groups and risk of overweight, obesity, and weight gain: A systematic review and dose-response meta-analysis of prospective studies. Adv Nutr.

2019;10(2):205–18. doi:10.1093/advances/nmy092

15. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhouser ML, et al. The dietary patterns methods project: synthesis of findings across cohorts and relevance to dietary guidance. J Nutr [Internet]. 2015;145(3):393–402. doi:10.3945/jn.114.205336

16. Wang DD, Li Y, Afshin A, Springmann M, Mozaffarian D, Stampfer MJ, et al. Global improvement in dietary quality could lead to substantial reduction in premature death. J Nutr [Internet]. 2019;149(6):1065–74. doi:10.1093/jn/nxz010

17. Chen J-P, Chen G-C, Wang X-P, Qin L, Bai Y. Dietary fiber and metabolic syndrome: A meta-analysis and review of related mechanisms. Nutrients. 2017;10(1):24. doi:10.3390/nu10010024

18. Wei B, Liu Y, Lin X, Fang Y, Cui J, Wan J. Dietary fiber intake and risk of metabolic syndrome: A meta-analysis of observational studies. Clin Nutr [Internet]. 2018;37(6):1935–42. doi:10.1016/j.clnu.2017.10.019

19. Ye EQ, Chacko SA, Chou EL, Kugizaki M, Liu S. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain. J Nutr. 2012;142(7):1304–13. doi:10.3945/jn.111.155325

20. Schwingshackl L, Hoffmann G. Long-term effects of low glycemic index/load vs. high glycemic index/load diets on parameters of obesity and obesity-associated risks: A systematic review and meta-analysis. Nutr Metab Cardiovasc Dis. 2013;23(8):699–706. doi:10.1016/j.numecd.2013.04.008

21. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9

22. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

23. National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. Nutrient reference values for Australia and New Zealand. Nutrition. Canberra; 2006.

24. Food and Nutrition Board: Institute of Medicine. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids (macronutrients). Washington, D.C; 2005.

25. Kindleysides S, Kruger R, Douwes J, Tannock G, Renall N, Slater J, et al. PRedictors linking Obesity and the gut MIcrobiomE (The PROMISE study): Protocol and recruitment strategy of a cross-sectional study on pathways that affect the gut microbiome and its impact on obesity. JMIR Res Protoc. 2019;8:1–16. doi:10.2196/14529

26. Atkinson J, Salmond C, Crampton P. NZDep2013 index of deprivation. NZDep2013 Index of Deprivation. Wellington; 2014.

27. Stewart A, Marfell-Jones M, Olds T, Hans De Ridder J. International standards for anthropometric assessment. Lower Hutt, New Zealand: International Society for the

Advancement of Kinanthropometry; 2011.

28. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: Insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28(7):412–9. doi:10.1007/BF00280883

29. Beck KL, Houston ZL, McNaughton SA, Kruger R. Development and evaluation of a food frequency questionnaire to assess nutrient intakes of adult women in New Zealand. Nutr Diet. 2018;1–7. doi:10.1111/1747-0080.12472

30. Finglas P., Roe M., Pinchen H., Berry R, Church S., Dodhia S., et al. McCance and Widdowson's the composition of foods, seventh summary edition. 7th ed. Cambridge: The Royal Society of Chemistry; 2015. 1–630 p. doi:10.1039/9781849737562

31. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001

32. Gemming L, Jiang Y, Swinburn B, Utter J, Mhurchu CN. Under-reporting remains a key limitation of self-reported dietary intake: An analysis of the 2008/09 New Zealand Adult Nutrition Survey. Eur J Clin Nutr [Internet]. 2014;68(2):259–64. doi:10.1038/ejcn.2013.242

33. Smith C, Gray AR, Mainvil LA, Fleming EA, Parnell WR. Secular changes in intakes of foods among New Zealand adults from 1997 to 2008/09. Public Health Nutr [Internet]. 2015 Dec 10;18(18):3249–59. doi:10.1017/S1368980015000890

34. Metcalf PA, Scragg RRK, Schaaf D, Dyall L, Black PN, Jackson R. Dietary intakes of European, Māori, Pacific and Asian adults living in Auckland: The Diabetes, Heart and Health Study. Aust N Z J Public Health [Internet]. 2008;32(5):454–60. doi:10.1111/j.1753-6405.2008.00279.x

35. Food Standards Australia New Zealand (FSANZ). Australia New Zealand Food Standards Code-Section 11-Calculation of values for nutrition information panel [Internet]. F2018C00953. 2018 [cited 2019 Aug 20].

36. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

37. Oliveros E, Somers VK, Sochor O, Goel K, Lopez-Jimenez F. The concept of normal weight obesity. Prog Cardiovasc Dis. 2014;56(4):426–33. doi:10.1016/j.pcad.2013.10.003

38. Kramer C, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions? A systematic review and meta-analysis. Ann Intern Med. 2013;159(11):758–69. doi:10.7326/0003-4819-159-11-201312030-00008

39. Dickey RA, Bartuska DG, Bray GW, Callaway CW, Davidson ET, Feld S, et al. AACE/ACE position statement on the prevention, diagnosis, and treatment of obesity (1998 Revision) AACE/ACE Obesity Task Force. Endocr Pract. 1998;4(5):297–350.

40. Jo A, Mainous AG. Informational value of percent body fat with body mass index for the risk of abnormal blood glucose: A nationally representative cross-sectional study. BMJ Open. 2018;8(4):1–6. doi:10.1136/bmjopen-2017-019200

41. New Zealand Guidelines Group. New Zealand primary care handbook 2012 [Internet]. 3rd ed. New Zealand Primary Care Handbook. Wellington: Ministry of Health; 2012. 45–64 p.

42. Ludwig DS, Pereira MA, Kroenke CH, Hilner JE, Van Horn L, Slattery ML, et al. Dietary fiber, weight gain, and cardiovascular disease risk factors in young adults. J Am Med Assoc. 1999;282(16):1539–46. doi:10.1001/jama.282.16.1539

43. Liu S, Willett WC, Manson JAE, Hu FB, Rosner B, Colditz G. Relation between changes in intakes of dietary fiber and grain products and changes in weight and development of obesity among middle-aged women. Am J Clin Nutr. 2003;78(5):920–7. doi:10.1093/ajcn/78.5.920

44. Tucker LA, Thomas KS. Increasing total fiber intake reduces risk of weight and fat gains in women. J Nutr. 2009;139(3):576–81. doi:10.3945/jn.108.096685

45. Rush EC, Freitas I, Plank LD. Body size, body composition and fat distribution: Comparative analysis of European, Maori, Pacific Island and Asian Indian adults. Br J Nutr [Internet]. 2009 Aug 10;102(04):632. doi:10.1017/S0007114508207221

46. Ministry of Health. Quality standards for diabetes care toolkit. Wellington; 2014.

47. Coppell KJ, Mann JI, Williams SM, Jo E, Drury PL, Miller J, et al. Prevalence of diagnosed and undiagnosed diabetes and prediabetes in New Zealand: Findings from the 2008/09 adult nutrition survey. N Z Med J. 2013;126(1370):23–42.

48. Breneman CB, Tucker L. Dietary fibre consumption and insulin resistance - The role of body fat and physical activity. Br J Nutr. 2013;110(2):375–83. doi:10.1017/S0007114512004953

49. McRorie JW, McKeown NM. Understanding the physics of functional fibers in the gastrointestinal tract: An evidence-based approach to resolving enduring misconceptions about insoluble and soluble fiber. J Acad Nutr Diet. 2017;117(2):251–64. doi:10.1016/j.jand.2016.09.021

50. Mann J, Cummings JH, Englyst HN, Key T, Liu S, Riccardi G, et al. FAO/WHO Scientific update on carbohydrates in human nutrition: Conclusions. Eur J Clin Nutr. 2007;61(S1):S132–7. doi:10.1038/sj.ejcn.1602943

51. Cummings JH, Stephen AM. The role of the dietary fibre in the human colon. Can Med Assoc J. 1980;123(11):1109–14.

52. Stephen AM, Cummings JH. Mechanism of action of dietary fibre in the human colon. Nature [Internet]. 1980 Mar;284(5753):283–4. doi:10.1038/284283a0

53. Wang T, Heianza Y, Sun D, Zheng Y, Huang T, Ma W, et al. Improving fruit and vegetable intake attenuates the genetic association with long-term weight gain. Am J Clin Nutr. 2019;1–10. doi:10.1093/ajcn/nqz166

54. Armet AM, Deehan EC, Thöne J V, Hewko SJ, Walter J. The effect of isolated and synthetic dietary fibers on markers of metabolic diseases in human intervention studies: A systematic review. Adv Nutr. 2019;(14):1–19. doi:10.1093/advances/nmz074

55. FAO/WHO. CODEX Alimentarius (CODEX) Guidelines on nutrition labelling CAG/GL 2-1985 [Internet]. 2017.

56. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. Cell. 2016;165(6):1332–45. doi:10.1016/j.cell.2016.05.041

57. Cummings JH, Macfarlane GT. The control and consequences of bacterial fermentation in the human colon. J Appl Bacteriol [Internet]. 1991 Jun;70(6):443–59. doi:10.1111/j.1365-2672.1991.tb02739.x

58. Makki K, Deehan EC, Walter J, Bäckhed 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

59. Bodinham CL, Smith L, Wright J, Frost GS, Robertson MD. Dietary fibre improves first-phase insulin secretion in overweight individuals. PLoS One. 2012;7(7):1–5. doi:10.1371/journal.pone.0040834

60. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SEK, et al. Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults. Gut. 2015;64(11):1744–54. doi:10.1136/gutjnl-2014-307913

61. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int J Obes. 2017;41(7):1099–105. doi:10.1038/ijo.2017.66

62. Ludwig DS, Willett WC, Volek JS, Neuhouser ML. Dietary fat: From foe to friend? Science (80-) [Internet]. 2018 Nov 16;362(6416):764–70. doi:10.1126/science.aau2096

63. De Souza RJ, Bray GA, Carey VJ, Hall KD, LeBoff MS, Loria CM, et al. Effects of 4 weight-loss diets differing in fat, protein, and carbohydrate on fat mass, lean mass, visceral adipose tissue, and hepatic fat: Results from the POUNDS LOST trial. Am J Clin Nutr. 2012;95(3):614–25. doi:10.3945/ajcn.111.026328

64. Kovatcheva-Datchary P, Nilsson A, Akrami R, Lee YS, De Vadder F, Arora T, et al. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. Cell Metab. 2015;22(6):971–82. doi:10.1016/j.cmet.2015.10.001

65. Schwab U, Lauritzen L, Tholstrup T, Haldorssoni T, Riserus U, Uusitupa M, et al. Effect of the amount and type of dietary fat on cardiometabolic risk factors and risk of developing type 2 diabetes, cardiovascular diseases, and cancer: A systematic review. Food Nutr Res. 2014;58. doi:10.3402/fnr.v58.25145

66. Astrup A, Bertram HCS, Bonjour JP, De Groot LCP, De Oliveira Otto MC, Feeney EL, et al. WHO draft guidelines on dietary saturated and trans fatty acids: Time for a new approach? BMJ [Internet]. 2019;366(July):1–6. doi:10.1136/bmj.l4137

67. New Zealand Ministry of Health. New Zealand Health Survey 2017/2018: Nutrition: Vegetable and fruit intake [Internet]. [cited 2019 Apr 11].

68. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr [Internet]. 2008 May 1;87(5):1107–17. doi:10.1093/ajcn/87.5.1107

69. University of Otago and Ministry of Health. A focus on nutrition: Key findings of the

2008/09 New Zealand adult nutrition survey [Internet]. Wellington: Ministry of Health; 2011. 1–332 p.

70. Amoah I, Cairncross C, Rush E. The battle for better nutrition: the role of the escalating fruit and vegetable prices. N Z Med J. 2017;130(1465):107–8.

71. Darmon N, Briend A, Drewnowski A. Energy-dense diets are associated with lower diet costs: A community study of French adults. Public Health Nutr. 2004;7(1):21–7. doi:10.1079/PHN2003512

72. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, et al. DIET@NET: Best practice guidelines for dietary assessment in health research. BMC Med [Internet]. 2017 Dec 15;15(1):202. doi:10.1186/s12916-017-0962-x

73. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

74. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

75. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.

76. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45. doi:10.3945/jn.115.219634

77. Kolonel LN, Henderson BE, Hankin JH, Nomura AMY, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: Baseline characteristics. Am J Epidemiol. 2000;151(4):346–57. doi:10.1093/oxfordjournals.aje.a010213

78. Chacko George G, Milani TJ, Hanss-Nuss H, Kim M, Freeland-Graves JH. Development and validation of a semi-quantitative food frequency questionnaire for young adult women in the southwestern United States. Nutr Res [Internet]. 2004 Jan;24(1):29–43. doi:10.1016/j.nutres.2003.09.006

79. Lal A, Moodie M, Ashton T, Siahpush M, Swinburn B. Health care and lost productivity costs of overweight and obesity in New Zealand. Aust N Z J Public Heal [Internet]. 2012;36(6):550–6. doi:10.1111/j.1753-6405.2012.00931.x

5. Chapter 5. Higher adherence to dietary patterns characterised by ‘core’ foods are associated with lower body fat percentage and metabolic risk factors for Pacific and New Zealand European women.

5.1. Abstract

Introduction: Dietary pattern analysis enables the assessment of the interrelatedness of foods and nutrients consumed together and how they relate to health outcomes. This study characterised the habitual dietary patterns of Pacific and New Zealand European (NZE) women and explored the association of adherence to habitual dietary patterns with body fat content and metabolic health markers.

Methods: Healthy Pacific (n=126) and NZE (n=161) women aged 18-45years living in Auckland, New Zealand were recruited based on body mass index (BMI; lean and obese) and stratified based on either low (<35 %) or high (≥35 %) body fat percentage (BF%). Habitual dietary intake data were assessed using the National Cancer Institute method considering a 5-day estimated food record and a semi quantitative food frequency questionnaire. *A posteriori* dietary patterns were derived using principal component analysis (PCA). Fasting blood samples were analysed for metabolic health markers (e.g., blood glucose and lipid profiles) and BF% was assessed by dual-energy x-ray absorptiometry. Adjusted multivariate linear regressions were conducted to explore the association of adherence to dietary patterns and BF%.

Results: Four dietary patterns which explained 30.9 % of the observed variance in habitual diet were identified. Higher adherence to the dietary patterns characterised by 'core' foods ("colourful vegetables, plant protein, and dairy" and "fruit, starchy vegetables, and nuts" patterns) were associated with lower BF% after adjusting for ethnicity, age, social deprivation, and energy intake. In contrast, higher adherence to patterns characterised by 'discretionary foods' ("sweet and fat rich carbohydrate" pattern) and less diversity of 'core' foods ("animal meat and fat" pattern), were associated with higher BF% for both Pacific and NZE women. Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was associated with lower homeostasis model assessment of insulin resistance (HOMA-IR) in comparison to low adherence to this pattern ($P \leq 0.001$). In contrast, higher adherence to the "sweet and fat rich carbohydrate" pattern was associated with higher HOMA-IR in comparison to low adherence to this pattern ($P \leq 0.01$). Across all dietary patterns, in comparison to the Acceptable Macronutrient Distribution Ranges (AMDRs), the median percentage of total and saturated fat intake exceeded the AMDRs (20-35 % and <10 % of total energy intake

respectively) and was below the AMDRs for percentage of total carbohydrate intake (45-65 % of total energy intake).

Conclusion: Despite having different metabolic disease risk, higher adherence to dietary patterns characterised by 'core' foods was associated with lower BF% and metabolic disease risks for both Pacific and NZE women. In contrast, higher adherence to patterns characterised by 'discretionary' foods or less diversity of 'core' foods were associated with higher BF% and metabolic disease risks. These findings highlight the importance of utilising dietary pattern analysis to assess the combinations of foods and nutrients consumed together in relation to health outcomes.

5.2. Introduction

Obesity is a global epidemic (1). Obesity is classified as a body mass index (BMI) ≥30 kg/m^2 and is defined as an excess accumulation of adipose tissue that could impair health (1). Obesity is considered a significant risk factor for developing non-communicable diseases (NCDs) such as diabetes, cancer, and cardiovascular disease (CVD) (2–4). In New Zealand (NZ), one in three people over the age of 15years are obese (5), ranking NZ as the country with the third highest adult obesity rate in the Organization for Economic Cooperation and Development (OECD) (6). Similar to global trends, in NZ higher levels of deprivation are associated with an increased risk of obesity and women are more likely to be obese than men (5). Of concern, in the past three decades the prevalence of obesity has substantially escalated in NZ while disproportionately affecting different population groups; for example, 67 % of Pacific people, 48 % of indigenous Māori and 29 % of NZE were classified as obese in the 2018/2019 health survey (5). Diet is considered a key modifiable risk factor for obesity and associated NCDs (3) and could be a key determinant driving NZ's high obesity rates.

The association between dietary intake and health outcomes is inherently complex to assess, because people choose to eat foods not nutrients, which, in turn, can also be influenced by environmental factors. Dietary pattern analysis enables the assessment of the interrelationship between foods and nutrients consumed together (e.g., meals consist of combinations of foods which deliver energy and nutrients), and how dietary patterns relate to disease and health outcomes (7–9). There are two main approaches to assess dietary patterns using self-reported dietary data, *a priori* or *a posteriori* approaches. *A priori* approaches use predefined criteria (based on current nutritional knowledge), and scores how the quality of the reported diet aligns with this criterion (e.g., Healthy Eating Index 2010 (HEI-2010), Alternative Healthy Eating Index 2010 (AHEI), Dietary Approaches to Stop Hypertension (DASH) (7,10–12)). In contrast, *a posteriori* approaches are data driven using statistical methods to characterise total diet by either identifying patterns of foods which are correlated with each other (e.g., principle component analysis) or clustering individuals into groups by frequency of intake (e.g., cluster analysis) (7,11). *A posteriori* derived patterns thereby enable the assessment of unique eating behaviours and enhance the understanding of a population's specific dietary intake (10).

Despite *a posteriori* patterns being unique to the study population, across studies similar patterns of foods have been identified and associated with distinct health outcomes. For example, 'unhealthy' or 'Western' patterns characterised by higher intake of 'discretionary' foods (e.g., refined grains, processed meat and foods, and sugar sweetened beverages) are associated with an increased risk of weight gain, obesity, and NCDs (10,12–14). In contrast, 'healthy' or 'prudent' patterns are characterised by higher intake of core or 'healthy' foods (e.g., fruit, vegetables, wholegrains, and lean proteins) which are associated with decreased risks of weight gain, obesity, and NCDs (10,13,14). Thus, despite the heterogeneity across studies (e.g., food group classification, naming of patterns, study populations), underlying eating patterns are consistently associated with health outcomes and appear to be markers of 'diet quality'.

Liese et al. (15) conducted the Dietary Patterns Methods Project (DPMP) to strengthen diet quality and mortality related conclusions by implementing standardised dietary analysis across three prospective cohorts. The four dietary pattern indexes (HEI-2010, AHEI-2010, DASH, and alternate Mediterranean Diet) used across the cohorts consistently identified that individuals with higher quality diets were associated with a reduced risk of mortality (15). Additionally, despite the variation in the combinations of foods across the indexes used, the indexes shared common food groups such as vegetables, fruit, wholegrains, and plant-based proteins, suggesting that these components may drive distinct health benefits. Furthermore, Liese et al. (15) highlighted observations of only small differences in diet quality within cohorts, which reflects that even modest changes in dietary intake could have a substantial influence on chronic disease risk.

The causes of obesity are complex and multifactorial. Yet, despite high obesity rates in NZ, there is limited insight into current dietary patterns, notably within population groups with different metabolic disease risks. A recent study by Beck et al. (16) derived dietary patterns from the last NZ National Nutrition Survey (NNS) conducted in 2008/09 and identified two patterns: 'healthy' and 'traditional'. The 'healthy' pattern was positively associated with age, NZE women, and inversely associated with BMI and higher socio-economic deprivation; in contrast, the 'traditional' pattern was positively associated with age, male gender, higher deprivation, and tobacco smoking (16). At the time of the last NZ NNS one in four adults were

classified as obese; now one in three adults are obese (5). Other studies exploring dietary patterns in NZ populations have focused on specific hypotheses such as nutrient status (17), pregnancy (18,19), NZE women (20), cardiorespiratory and muscular fitness in children (21) and dieting in adolescents (22).

A posteriori dietary patterns tend to be derived from Food Frequency Questionnaires (FFQ) (17–22) which are prone to overestimation of intake (9,23,24). Multiple days of dietary assessment are considered more accurate (24); however, there are no studies in NZ that have derived dietary patterns from multiple days and dietary assessment methods in a healthy population. The aim of this cross-sectional study was to characterise *a posteriori* dietary patterns, derived from multiple days of dietary assessment, and to explore the association with body fat content and metabolic health markers in a population of NZ women with different metabolic disease risk (Pacific women have a high risk of obesity (70 % obesity) and NZE women have a moderate risk of obesity (30 % obesity) (5)), and body fat profiles (lean and obese).

5.3. Materials and methods

5.3.1. Study design

Participants were part of The PRedictors linking Obesity and gut MIcrobiomE: PROMISE cross-sectional study, which was conducted between July 2016 and September 2017, at the Human Nutrition Unit at Massey University in Albany, Auckland, New Zealand (25). This sub analysis aimed to explore the dietary patterns of these women. Eligible participants were free from any chronic illness or disease, post-menarche and premenopausal Pacific and NZE women aged 18-45 years from the Auckland region. Participants were recruited based on their BMI (normal: <25 kg/m^2 or obese: ≥30 kg/m^2) and details of the study procedures and recruitment strategies have been published elsewhere (25). Participants attended two study visits and completed at home data collection between study visits.

The PROMISE study was approved by the Southern Health Disability Ethics Committee (16/STH/32) and conducted according to the guidelines of the declaration of Helsinki. The trial was registered at anzctr.org.au (ACTRN12618000432213). All participants received detailed information about the procedures, and study measurements, and gave written informed consent prior to their participation in the study.

5.3.2. Demographic and anthropometric information and blood sampling

During the first study visit trained research staff followed standard operating procedures (SOP) to conduct one-on-one interviews with participants to capture a range of demographic information (e.g. occupation, personal/household income). Accredited research staff collected anthropometric measures (e.g. fasted weight (kg), stretched height (cm)) using the International Society for the Advancement of Kinanthropometry (ISAK) protocols (26). BMI was calculated using the Quetelet index (weight (kg)/height $(cm)^2$) (27). Fasting blood samples were collected by a trained phlebotomist after an overnight fast of at least 10 hours, and blood metabolic markers (e.g., plasma glucose, insulin, glycosylated haemoglobin (HbA1c), and lipid profile) were assessed. Collection of the blood samples has been discussed in detail in the PROMISE protocol paper (25), as has processing of the blood samples (see chapter four). Homeostasis model assessment (HOMA-IR) index for insulin resistance was calculated (fasting blood glucose [mmol/L] X fasting plasma insulin [µU/mL]/22.5) (28).

The New Zealand Deprivation Index (NZDep2013) was used in this study as a measure of socioeconomic status. The NZDep2013 is an area-based measure of socioeconomic deprivation based on NZ census information. The NZDep2013 provides a deprivation score ranging between decile one “least deprived” and decile ten “most deprived”, thus a participant is assigned a deprivation score based on where they live (29) which has been discussed previously (see chapter four and (25)). At study visit two a whole-body scan was performed to assess total body fat percentage (BF%) using Dual-energy X-ray Absorptiometry (DXA) (Hologic QDR Discovery A, Hologic Inc, Bedford, MA with APEX V. 3.2 software), and regional BF% (i.e., visceral, android, and gynoid fat percentage). Blood pressure was measured after a 10-minute resting period with an Omron HEM 907 digital blood pressure monitor. The

measurement was repeated three times, and the average of the second and third measurement recorded.

5.3.3. Dietary assessment

Participants completed their five-day, non-consecutive, estimated food record (5DFR), at home. Details of dietary assessment methodology have been discussed previously (see chapter three and four). During their second study visit, (11 to 14 days after visit one) every completed 5DFR was reviewed by a research dietitian. This was followed by an individual interview to clarify, if necessary, any aspect of the food record, such as food preparation and cooking methods used, portion sizes or any apparent ambiguities. Participants also completed a validated semi-quantitative New Zealand Women's Food Frequency Questionnaire (NZWFFQ) regarding the previous 30 days dietary intake (30). The 220-item NZWFFQ provided standard portion sizes (e.g., 1 teaspoon of sugar) as a reference for participants, and the FFQ was hosted by SurveyMonkey© software (SurveyMonkey Inc, San Mateo, California, USA), which allowed research staff the ability to monitor live progress as participants completed it online.

5.3.4. Dietary data processing

Energy, macro- and micronutrient analyses of the 5DFR and FFQ were independently completed using FoodWorks9 (Xyris Software (Australia) Pty Ltd, Queensland, Australia), which hosts multiple food composition databases including New Zealand's food composition database (FOODFiles 2016), developed by the NZ Institute for Plant & Food Research and the NZ Ministry of Health. Trained research staff entered the 5DFR and NZWFFQ into the software, using a standardised data entry procedure to ensure consistent high-quality dietary data entry, utilising a hierarchy of food choice from the different food composition databases. Dietary data processing of the 5DFR and NZWFFQ have been discussed in detail previously (see chapter four).

All dietary data (5DFR and NZWFFQ) were entered twice into the software and were independently checked by a second researcher to confirm reliability of data entry. Energy, nutrient, and food intake data were exported from FoodWorks9 in standard units. All exported

data sets were extensively reviewed using the raw dietary data to assess plausibility of processed dietary data and to identify misreporting. Cut-offs of >2100 kJ/day and <27000 kJ/day were considered to indicate valid completion of the 5DFR and NZWFFQ. Although these cuts-offs are adapted from Willett's epidemiological recommendations (9), they were considered to be more reflective of the foods consumed by the participants (31–33). Reported intakes below or above these cut-offs were considered invalid and were excluded from further dietary analysis.

Food groups

To assess the complex interrelationships between nutrients and foods, all reported foods items from the 5DFR (n>2850) and the NZWFFQ were summarised into 55 food groups (in g/day) based on similar nutritional composition and characteristics (Supplementary Table 5.1). Mixed dishes reported in the participants' 5DFRs were assigned into the appropriate food group as follows: when participants provided us with the specific details (e.g., weights of raw ingredients, cooking methods) of the meals they consumed, the individual components of these "Participant meals" were assigned to the corresponding food groups. For example, for the meal 'spaghetti bolognaise', the cooked portion consumed of pasta was assigned to "refined grains", mince to "red meat", tomatoes to "tomatoes", and onion to "other-non starchy vegetables" etc. The same approach was taken for beverages; for example, milk and sugar added to coffee and/or tea was assigned to the appropriate food group (e.g., "full fat milk" and "added sugar to food and beverages" respectively). In contrast, participant and standard recipes were assigned to the corresponding food groups in their entirety (e.g., banana cake was assigned to the food group "cakes and biscuits", and bliss balls were assigned to "sweet snacks"). Primarily because the existing analysed foods and recipes, which were used to create "Participant recipes" and "Standard recipes", were unable to be broken down to individual components. In addition, for descriptive purposes, the 55 food groups were broadly classified as either 'core' or 'discretionary' foods based on NZ (Ministry of Health) and Australian (Department of Health) dietary guidelines (34,35). 'Core' foods are considered nutrient rich foods, from the four core food groups (40); whereas, 'discretionary' foods do not fit into these core or 'healthy' food groups; and are characteristically energy dense, processed nutrient poor foods (41,42).

Total energy (reported as kilojoules (kJ)) includes the energy contribution from all the macronutrients as well as total dietary fibre and alcohol. Total dietary fibre includes all non-starch polysaccharides (cellulose, hemicellulose, and resistant starches). The NZ FOODFiles 2016 uses AOAC Prosky method to analyse dietary fibre. Total carbohydrate (reported in grams (g)) includes free sugars, dextrins, starch, and glycogen. Total fat intake (reported in grams (g)) includes all saturated and unsaturated fats.

5.3.5. Calculating habitual dietary intake

For each participant with valid dietary data (within the specified range of kJ cut-offs), individual habitual dietary intake of nutrients and food groups was estimated using the National Cancer Institute (NCI) method which has previously been discussed in detail (see chapter three and four). Briefly, for episodically consumed foods (i.e., foods not consumed each day) the NCI method uses a two-part modelling approach to estimate the probability of consumption and the respective amount consumed, considering the effect of covariates which can influence the probability of consumption or the amount consumed. The individual habitual intake is then defined as the product of probability of consumption multiplied by the consumed amount (*Usual intake = Probability x Amount*). For foods or nutrients consumed daily, there is no need to model probability; therefore, the amount of usual intake is estimated considering person specific covariates.

In the current analysis, the three NCI SAS macros (Version 2.1 - available from the NCI website) (36) were adapted for the PROMISE data in SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). Utilising the 5DFR as the primary dietary data source, the covariates age, ethnicity, BMI, season (summer, winter, spring, autumn), NZWFFQ information (in standard units/day) and weekend information (Weekday = Monday-Thursday, Weekend = Friday-Sunday) were considered. If a participant reported never consuming a nutrient or food group in both the 5DFR and NZWFFQ they were assigned 0 g/day for the respective nutrient/food group. The average daily intake of 36 nutrients (in units/day) and 55 food groups (g/day) consumed within the last month was estimated for each participant.

5.3.6. Dietary pattern analysis

Principal component analysis (PCA) was conducted to identify dietary patterns, where the principal component scores extracted from PCA represent dietary pattern scores for use in subsequent analyses (see below). The estimated daily intake values of the 55 food groups (described above) were individually log transformed to approach normality prior to dietary pattern analysis (using SPSS 25.0 for Windows (SPSS inc., Chicago, IL, USA)). All participants with available habitual dietary data were included in the exploratory PCA. The break-even point of the scree plot, the Kaiser criterion (eigenvalues greater than 1), varimax rotation, and the magnitude of the component loadings themselves, were all used to identify distinct and interpretable patterns, with component loadings of >0.3 considered to significantly contribute to the pattern. Food groups are weighted onto a component, either positively or negatively, by their contribution to explain the variance in the dietary pattern, with highly positive component loadings contributing significantly to the pattern. Cronbach's alpha (≥0.8 = good, 0.7 = moderate) (37) was used to assess the inter item reliability of each food group and the pattern they are associated with (e.g., whether the inclusion of a food group could skew the reliability of a pattern). Each participant was assigned an individual score of adherence to each extracted dietary pattern, which ranged from positive to negative values. Thus, an individual may score positively in one pattern and negatively in another. To account for higher energy intakes correlating with higher nutrient intakes (9,38), dietary pattern scores were adjusted for total energy (kJ) intake, and energy-adjusted dietary patterns were used for all subsequent analyses. These individual energy adjusted scores were then used to rank participants into tertiles of adherence to the different patterns for further analyses.

5.3.7. Statistical analysis

Although participants were recruited based on their BMI, BF% was used to stratify the population because individuals with the same BMI can have different body compositions and metabolic disease risks (39,40). Participants were classified as either low-BF% (<35 %) or high-BF% (≥35 %) with the median as the cut point. Statistical analyses were conducted using SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). Measurements that did not follow a normal distribution were reported as medians [25^{th}, 75^{th}] and non-parametric tests were applied to assess the differences between groups. The NZDep2013 index was collapsed into

quintiles (e.g., quintile one = decile one and two, and quintile five = decile nine and ten) and these quintiles were used for subsequent analyses (29). Comparisons of participant characteristics focused on the differences between BF% groups within each ethnic group (e.g., NZE low-BF% versus NZE high-BF%) and between ethnicities within BF% groups (e.g., NZE low-BF% versus Pacific low-BF%). Participants were also stratified into tertiles of adherence to the energy-adjusted dietary patterns and comparisons focused on low adherence (lowest 33 % of the population) *versus* high adherence (tertile three). Spearman's rank correlation analysis was used to explore the associations between the adjusted dietary patterns scores, and food groups and nutrients. To assess the macronutrient intake in relation to the total energy intake of the diet, the percentage of total energy intake from macronutrients were calculated (41). Statistical analyses were conducted separately for Pacific and NZE participants, as well as for the groups combined. *P*-values <0.05 were considered statistically significant.

To assess the association with habitual dietary intake (dietary patterns as defined above) and total BF%, univariate linear regression analyses were conducted followed by multivariate linear regression analyses controlling for potential confounders including: NZDep2013 quintiles and age. Due to the design of the study (e.g., participants were selectively recruited based on ethnicity) analyses were adjusted for ethnicity. All independent variables were assessed for collinearity (by assessing tolerance and variance inflation factor). No collinearity was detected.

5.4. Results

5.4.1. Participants' characteristics and their metabolic health markers

A total of 351 participants were eligible and enrolled to participate in the PROMISE study, and 304 participants completed all aspects of the study. In this study, 17 participants were excluded prior to calculation of the habitual dietary data based on misreporting (>27000 kJ/day) their energy intake in the NZWFFQ. We therefore had complete data for 287 participants: 126 (44 %) Pacific and 161 (56 %) NZE women, with a median age of 23 and 32 years respectively. Participant characteristics stratified by ethnicity and BF% categories are reported in Table 5.1.

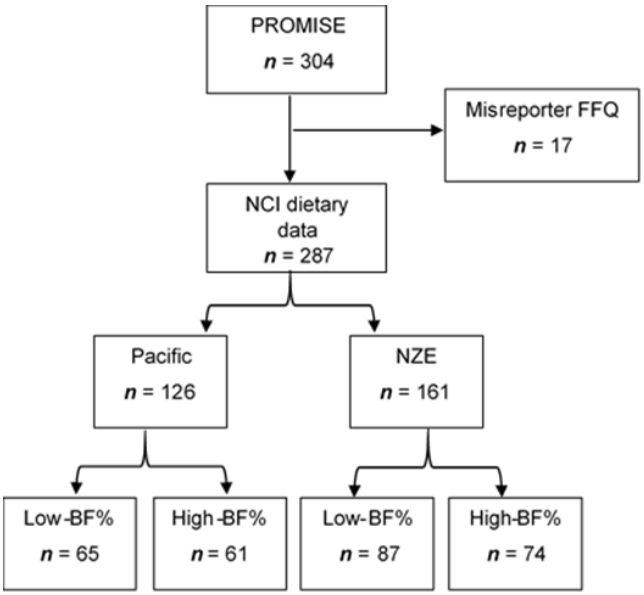


Figure 5.1. Overview of participant's included in the study

The median BF% for all participants, when stratified by ethnicity, was <35 % (Pacific 34.6 % [29.5, 39.4], NZE 34.4 % [27.8, 39.9]). In the low-BF% group, Pacific women had higher BMI, weight, visceral (VAT) and android fat percentage (AF%) compared to NZE women. In the high-BF% group, Pacific women had a lower BF% compared to NZE women, and in the low-BF% group NZE women had a lower BF% in comparison to Pacific women. In the high- and low-BF% groups NZE women had lower HbA1c, fasting glucose, insulin and HOMA-IR, but had higher total and HDL cholesterol in comparison to Pacific women (Table 5.1).

Table 5.1. Participant characteristics stratified by ethnicity and body fat percentage group

	All participants n = 287	Pacific Low-BF% n = 65	Pacific High-BF% n = 61	NZ European Low-BF% n = 87	NZ European High-BF% n = 74
Age (y)	28.0 [22.0, 35.0]	23 [20, 29]	23 [21, 29]	29 [24, 36]π	35 [28, 40]*π
Body composition					
Weight (kg)	77.6 [65.6, 96.0]	72.4 [67.3, 79.1]	97.0 [87.4, 109.9]*	62.4 [58.1, 66.6]π	94.1 [86.8, 101.7]*
Height (cm)	167.7 [163.7, 172.2]	168.7 [164.7, 174.2]	168.4 [163.3, 173.4]	167.7 [163.9, 170.7]	166.7 [162.5, 171.9]
BMI (kg/m^2)	28.0 [23.0, 33.4]	25.0 [23.6, 27.6]	33.8 [31.1, 39.9]*	22.5 [20.9, 23.5]π	33.5 [31.7, 36.3]*
Body fat (%)	34.6 [28.8, 39.6]	29.6 [27.9, 32.3]	39.5 [36.6, 42.4]*	28.0 [24.2, 31.9]π	40.3 [38.7, 44.2]*π
Visceral fat %	32.3 [23.7, 38.9]	26.8 [23.1, 31.4]	40.3 [35.6, 43.3]*	21.5 [16.8, 27.3]π	39.7 [35.7, 44.0]*
Android fat %	34.3 [26.1, 40.7]	29.1 [25.6, 33.5]	41.2 [38.0, 44.8]*	24.2 [20.2, 29.1]π	41.5 [38.0, 45.7]*
Gynoid fat %	37.3 [34.1, 41.8]	35.1 [32.7, 37.1]	41.0 [38.5, 43.7]*	34.3 [30.5, 36.7]	42.5 [39.8, 45.1]*π
Blood pressure€					
Systolic (<120 mmHg)	115.0 [107.5, 122.5]	113.0 [106.0, 118.5]	116.5 [110.5, 128.0]*	112.5 [104.5, 118.5]	120.0 [111.3, 127.8]*
Diastolic (<80 mmHg)	73.5 [68.0, 80.0]	70.5 [65.0, 74.0]	77.0 [70.5, 83.5]*	69.0 [66.0, 76.0]	79.5 [73.8, 85.0]*
Metabolic markers∞					
TC (<5 mmol/L)-	4.8 [4.3, 5.4]	4.5 [4.1, 5.1]	4.6 [4.2, 5.1]	4.9 [4.3, 5.4]π	5.2 [4.7, 6.1]*π
HDL-C (>1 mmol/L)-	1.6 [1.3, 1.9]	1.5 [1.3, 1.8]	1.3 [1.2, 1.6]*	1.8 [1.6, 2.0]π	1.4 [1.3, 1.7]*π
LDL-C (0-3.4 mmol/L)-	3.0 [2.4, 3.5]	2.8 [2.4, 3.2]	3.0 [2.5, 3.3]	2.8 [2.4, 3.4]	3.4 [2.7, 4.1]*π
TGS (<2 mmol/L)-	0.9 [0.7, 1.2]	0.8 [0.7, 1.1]	1.0 [0.9, 1.5]*	0.7 [0.6, 0.9]	1.1 [0.8, 1.5]*
TC:HDL (<4 mmol/L)-	3.1 [2.6, 3.7]	2.9 [2.5, 3.3]	3.4 [2.9, 4.0]*	2.7 [2.3, 3.1]π	3.5 [3.0, 4.3]*
HbA1c (<40 mmol/mol)-#	31.7 [30.0, 33.6]	32.1 [30.5, 33.8]	34.8 [32.3, 36.7]*	30.6 [29.0, 31.9]π	31.0 [29.8, 33.3]*π
Glucose (3.5-5.4 mmol/L)-	5.3 [5.0, 5.6]	5.3 [5.0, 5.5]	5.4 [5.1, 5.9]*	5.1 [4.9, 5.3]π	5.5 [5.1, 5.7]*
Insulin (3-25 uU/mL)-	10.9 [7.3, 17.3]	11.2 [7.9, 16.0]	21.4 [13.1, 31.9]*	7.1 [5.2, 8.7]π	12.6 [10.0, 17.9]*π
HOMA-IR	2.6 [1.7, 4.0]	2.6 [1.9, 3.6]	4.8 [3.1, 8.5]*	1.6 [1.1, 2.1]π	3.1 [2.3, 4.1]*π
NZDep2013	6.0 [3.0, 8.0]	7 [5, 9]	8 [7, 9]	3 [2, 6]π	5 [3, 6]*π

All values are reported as medians [25th, 75th percentiles]. Mann Whitney statistical test used to identify a significant difference ($p < 0.05$). *Statistically significant difference between body fat groups within ethnic group. πStatistically significant difference between ethnic groups within body fat group. € NZE women (n=2), ∞PI woman (n=1) and #PI women (n=2) have not been included in analysis due to missing data. Total body fat % (BF%) assessed with DXA, Low-BF%: BF% <35 %, High-BF% BF%: ≥35 %. BMI: body mass index, TC: total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TGS: Triglycerides, HbA1c: glycosylated haemoglobin. Deprivation Index NZ 2013. -Normal healthy ranges for metabolic markers (53)

5.4.2. Dietary patterns

Results are only presented for the groups combined as the results were similar for Pacific and NZE participants. A total of four habitual dietary patterns were identified. Combined, the four dietary patterns explained 30.9 % of the total variance of the habitual dietary intake data (see Table 5.2). These habitual dietary patterns were named according to the nutritional composition of the food groups that loaded highly (>0.4) onto the components (discussed below). No food groups were removed from the patterns as the overall reliability for all patterns ranged from good to moderate (0.8-0.7). Cross loadings were retained as they were considered to contribute to the patterns.

5.4.3. Characteristics of pattern 1

The “colourful vegetables, plant protein, and dairy” pattern explained the largest amount of variance (13.7 %), and was characterised by higher loadings of colourful vegetables, dairy products, plant proteins, alcohol and coffee, and ‘healthy’ foods. This pattern correlated positively with the following food groups: “other non-starchy vegetables”, “tomatoes”, “green vegetables”, “low fat cheese”, and “legumes and meat alternatives”, and inversely with; “sugar sweetened beverages”, “fast food burgers”, “unsweetened cereals”, “refined grains mixed dishes” and “crumbed and deep fried” (Supplementary Table 5.2). At the nutrient level this pattern correlated positively with potassium (mg/day), total vitamin A (μg/day), calcium (mg/day) and total dietary fibre (g/day); and negatively with total carbohydrate (including starch and total sugar (g/day)) intake (Supplementary Table 5.3).

Higher adherence to the “colourful vegetables, plant protein, and dairy” pattern was inversely associated with BF% after adjusting for age, ethnicity, NZDep2013 and energy intake (Table 5.3). When stratifying the pattern into tertiles of adherence, tertile 3 (higher adherence) was characterised by older NZE women (96 %), with lower NZDep2013 scores, including lower BMI, VAT, AF%, and gynoid fat % (GF%). These women also had lower circulating concentrations of fasting triglycerides, TC:HDL ratio, insulin, blood glucose, HbA1c and HOMA-IR index, including higher HDL (see Table 5.4).

Table 5.2. Component loadings of four habitual dietary patterns

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Explanation of variation of food intake (%)	**13.7**	**6.2**	**5.7**	**5.4**
Cronbachs's alpha	**0.8**	**0.7**	**0.7**	**0.7**
Other non-starchy vegetables	0.79			
Green vegetables	0.71		0.33	
Savoury sauces and condiments	0.66			
Low fat cheese	0.65			
Legumes and meat alternatives	0.64			-0.39
Yellow vegetables	0.59		0.44	
High fat cheese	0.58			
Tomatoes	0.57			
Fast food burgers	-0.57	0.38		
Nuts and seeds	0.56		0.41	
Plant based fats	0.56			
Coffee	0.55			
Refined grains mixed dishes	-0.52	0.30		
Alcoholic beverages	0.51			
Unsweetened cereals	-0.49			
Sugar sweetened beverages	-0.47	0.37		
Dairy Yoghurt	0.43			
Egg and egg products	0.41			
Crumbed and deep-fried food	-0.41			
Added sugar to food and beverages	-0.34			
Tea	0.33			
Sweet snack foods	0.31			
Water				
Fruit and vegetable juice				
Sweetened milk products		0.55		
Potatoes		0.54		
Discretionary breads	0.32	0.51		
Puddings and desserts		0.50		
Fast food salad and sushi		0.49		
Creamy based sauces and dressings		0.43		
Cakes and biscuits		0.41		
White bread		0.38		
Savoury snack foods		0.32		
Margarine		0.32		
Diet drinks				
Low fat milk				
Refined grains				
Sweet spreads				
Apple, banana, orange			0.57	
Starchy vegetables			0.45	
Other fruit			0.44	
Coconut products			0.44	
Peanut butter and peanuts	0.36		0.43	
Fish and seafood			0.37	0.34
Milk alternatives			0.35	
Wholegrain products		0.31	0.34	
Oats				
Soups and stocks				
Crackers				
Sweetened cereals				
Red meat				0.82
White meat				0.82
Processed meat				0.78
Animal fats				0.43
Full fat milk				

Patterns identified based on component loading >0.3. Kaiser-Meyer-Olkin measure of sampling adequacy=0.795, Bartlett's test of sphericity = 0.000. Total n=287 in each dietary pattern; Pacific: n=126, NZE: n=161. Food groups with loadings <0.3 are not shown.

Table 5.3. Association of adherence to energy adjusted dietary patterns and body fat percentage

Pattern 1 β+ (95 % CI)	Pattern 2 β+ (95 % CI)	Pattern 3 β+ (95 % CI)	Pattern 4 β+ (95 % CI)
-3.35 [-4.82, -1.88]***	1.66 [-0.82, 2.49]***	-1.26 [-2.06, -0.46]**	0.96 [0.12, 1.81]*

Total n=284 in each dietary pattern: Pacific n= 124 (n=2 with missing deprivation data), NZE n= 160 (n=1 with missing deprivation data) **P* value <0.05, ***P* value ≤0.01, ****P* value ≤0.001 **+**Adjusted for age, ethnicity and NZDep2013.

5.4.4. Characteristics of pattern 2

The "sweet and fat rich carbohydrate" pattern was characterised by higher loadings of 'discretionary' foods and correlated positively with "sweetened milk products", "creamy based sauces and dressings", "cake and biscuits" and "puddings and desserts", and inversely with "nuts and seeds", "plant-based fats", "water", "coconut products" and "green vegetables" (Supplementary Table 5.2). This pattern explained 6.2 % of the variance, and at the nutrient level correlated positively with total carbohydrate (including starch, and sugar (g/day)) and sodium (mg/day) (Supplementary Table 5.3). Higher adherence to the "sweet fat and rich carbohydrate" pattern was associated with higher BF% (seeTable 5.3). Women in the highest tertile of adherence to this pattern were characterised by higher BMI, BF%, VAT, GF% and AF%, including higher fasting TC, triglycerides, LDL, TC:HDL ratio, insulin, and HOMA-IR index (see Table 5.4).

Table 5.4. Participant characteristics stratified by tertiles of adherence to energy adjusted dietary pattern 1 and 2

	Pattern 1: Colourful vegetables, plant protein and dairy			Pattern 2: Sweet and fat rich carbohydrate		
	T1	**T2**	**T3**	**T1**	**T2**	**T3**
Age (years)	22 [20, 25]***	29 [24, 37]^	34 [28, 38]+++	29 [22, 38]	28 [22, 36]	27 [23, 33]
Ethnicity *n* (%)						
Pacific	92 (73 %)	30 (24 %)	4 (3 %)	41 (33 %)	54 (43 %)	31 (25 %)
NZE	3 (2 %)	66 (41 %)	92 (57 %)	54 (34 %)	42 (26 %)	65 (40 %)
Body composition						
Weight (kg)	84.2 [72.9, 100.8]**	74.3 [63.9, 93.5]	70.7 [62.5, 92.3]+++	72.9 [62.4, 90.2]**	82.1 [68.1, 99.9]	80.2 [67.3, 96.7]+
Height (cm)	168.4 [163.8,174.1]	166.6 [163.8, 171.6]	168.2 [163.6, 171.4]	167.7 [163.2, 172.5]	169.3 [164.4, 174.1]	166.5 [164.0, 170.7]
BMI (kg/m^2)	31.1 [25.3, 35.2]**	26.1 [22.7, 32.8]	24.9 [21.9, 32.4]+++	24.6 [21.9, 32.8]**	30.5 [23.6, 34.5]	29.3 [23.9, 33.6]++
Total body fat %	35.4 [31.9, 39.9]	32.7 [29.4, 39.7]	32.9 [25.0, 39.4]++	32.3 [26.6, 38.8]	35.6 [29.6, 39.4]	35.3 [29.5, 40.6]++
Visceral fat %	35.2 [27.7, 40.5]*	30.4 [23.4, 40.1]^	28.6 [18.4, 37.0]+++	28.9 [19.9, 37.1]*	34.2 [25.4, 39.6]	33.7 [25.8, 40.0]++
Android fat %	37.5 [29.7, 42.0]*	32.8 [26.2, 42.1]^	30.1 [21.3, 39.2]+++	30.3 [22.5, 39.4]*	35.9 [27.7, 41.1]	35.6 [28.1, 41.8]+
Gynoid fat %	38.5 [35.4, 41.6]	37.0 [34.4, 42.1]	36.3 [31.5, 41.6]+	35.9 [31.8, 40.9]	37.7 [34.7, 41.0]	38.1 [35.4, 43.1]++
Blood pressure€						
Systolic (<120 mmHg)	113.5 [107.5,122.0]	116.0 [107.0, 123.5]	114.0 [106.5, 123.0]	115.0 [110.0, 123.0]	114.0 [106.0, 122.5]	115.5 [107.0, 123.8]
Diastolic (<80 mmHg)	72.0 [67.0, 80.0]	75.0 [70.0, 81.0]	73.0 [68.0, 81.0]	73.8 [67.5, 81.0]	71.5 [67.5, 79.5]	75.0 [70.0, 81.5]
Metabolic markers∞						
TC (<5 mmol/L)-	4.5 [4.1, 5.0]*	4.9 [4.2, 5.6]	5.0 [4.6, 5.8]+++	4.7 [4.1, 5.2]	4.7 [4.3, 5.4]	5.0 [4.4, 5.7]+
HDL-C (>1 mmol/L)-	1.4 [1.2, 1.6]***	1.6 [1.3, 1.8]^^	1.7 [1.5, 2.0]+++	1.7 [1.4, 2.0]*	1.5 [1.3, 1.8]	1.5 [1.3, 1.8]++
LDL-C (0-3.4 mmol/L)-	2.8 [2.4, 3.3]	3.1 [2.3, 3.5]	3.0 [2.6, 3.6]+	2.8 [2.4, 3.3]	3.0 [2.5, 3.4]	3.1 [2.6, 3.7]+
TGS (<2 mmol/L)-	1.0 [0.8, 1.2]	0.9 [0.7, 1.2]	0.8 [0.6, 1.1]++	0.8 [0.6, 1.1]	0.9 [0.7, 1.1]	1.0 [0.8, 1.3]+++
TC:HDL (<4 mmol/L)-	3.2 [2.8, 4.0]	3.1 [2.6, 3.7]	2.9 [2.5, 3.3]++	2.9 [2.4, 3.3]*	3.1 [2.7, 3.7]	3.3 [2.8, 4.0]+++
HbA1c (<40 mmol/mol)-#	33.0 [31.6, 35.0]***	31.4 [29.4, 33.7]	30.7 [29.6, 32.4]+++	31.3 [30.0, 33.1]	32.2 [30.3, 34.8]^	31.6 [29.4, 33.4]
Glucose (3.5-5.4 mmol/L)-	5.4 [5.1, 5.8]	5.3 [5.0, 5.6]	5.2 [4.9, 5.6]++	5.3 [5.0, 5.6]	5.3 [5.1, 5.5]	5.3 [5.0, 5.6]
Insulin (3-25 uU/mL)-	16.2 [11.2, 25.1]***	10.0 [7.5, 14.8]^^^	8.2 [5.7, 11.7]+++	9.1 [6.7, 14.9]*	11.4 [7.2, 20.4]	11.8 [8.5, 17.4]++
HOMA-IR	3.8 [2.5, 6.3]***	2.4 [1.7, 3.5]^^^	1.9 [1.2, 2.7]+++	2.1 [1.5, 3.6]*	2.7 [1.6, 4.8]	2.8 [2.0, 3.9]+
NZDep2013	8 [6, 9]***	5 [2, 7]	4 [2, 6]+++	6 [3, 9]	6 [4, 9]^	5 [3, 7]

All values are reported as median [25th, 75th]. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. T1: Tertile 1; *n* = 95, T2: Tertile 2; *n* = 96, T3: Tertile 3; *n* = 96. T1= negatively adhering to the dietary pattern, T3 = positively adhering to the dietary pattern. Differences between tertiles within a dietary pattern were tested by Mann-Whitney test. * Significant difference between tertiles 1 and 2, $P < 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***. + Significant difference between tertiles 1 and 3, $P < 0.05$ +, $P \leq 0.01$++, $P \leq 0.001$+++, ^Significant difference between tertiles 2 and 3, $P < 0.05$ ^, $P \leq 0.01$^^, $P \leq 0.001$ ^^^. **€** NZE women (*n* = 2), ∞ Pacific woman (*n* = 1), and #Pacific women (*n* = 2) have not been included in analysis due to missing data. NZE: New Zealand European, BMI: body mass index. TC: Total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TGS: Triglycerides, HOMA-IR: Homeostasis model assessment index for insulin resistance (HOMA-IR), HbA1c: glycosylated haemoglobin, DXA used to measure all body fat content. -Normal healthy ranges for metabolic markers (53)

5.4.5. Characteristics of pattern 3

The “fruit, starchy vegetables, and nuts” explained 5.7 % of the variance and was characterised by high loadings of ‘healthy’ foods. This pattern positively correlated with “apple, banana & orange”, “other fruit”, “milk alternatives”, “coconut products” and “starchy vegetables”, and negatively with the following food groups; “crumbed and deep-fried”, “white bread”, “fast food burgers”, “high fat cheese” and “full fat milk” (Supplementary Table 5.2). At the nutrient level, this pattern positively correlated with dietary fibre (g/day), vitamin C (mg/day), magnesium (mg/day), and beta-carotene (µg/day); and negatively with saturated fat (g/day), sodium mg/day), and retinol (µg/day) (Supplementary Table 5.3). Higher adherence to the “fruit, starchy vegetables, and nuts” pattern was associated with lower BF% (see Table 5.3). Women in the highest tertile of adherence to this pattern were characterised by lower BF%, including VAT, AF% and GF% (Table 5.5).

5.4.6. Characteristics of pattern 4

The “animal meat and fat” pattern explained the smallest amount of variance (5.4 %) and was characterised by high loadings of animal protein and fat rich foods. This pattern correlated positively with the following food groups: “processed meat”, “animal fats”, “red meat”, “white meat” and “high fat milk”, and negatively with; “sweetened cereals”, “milk alternatives”, “legumes and meat alternatives” and “crumbed and deep fried” (Supplementary Table 5.2). At the nutrient level, this pattern was positively correlated with total cholesterol (g/day), protein (g/day), vitamin B12 (µg/day) and zinc (mg/day), and negatively correlated with total carbohydrate (including sugar and starch (g/day)), and total dietary fibre (g/day) (Supplementary Table 5.3). Higher adherence to the “animal meat and fat” pattern was associated with higher BF% (see Table 5.3). Women in the highest tertile of adherence to this pattern were older and had higher blood pressure (systolic and diastolic), BMI, and BF% including VAT and AF% (see Table 5.5).

Table 5.5. Participant characteristics stratified by tertiles of adherence to energy adjusted dietary pattern 3 and 4

	Pattern 3: Fruit, starchy vegetables and nuts			Pattern 4: Animal meat and fat		
	T1	**T2**	**T3**	**T1**	**T2**	**T3**
Age (years)	26 [22, 35]	28 [21, 36]	28 [23, 35]	25 [21, 32]	28 [22, 35]	30 [24, 37]++
Ethnicity *n* (%)						
Pacific	38 (30 %)	50 (40 %)	38 (30 %)	45 (36 %)	38 (30 %)	43 (34 %)
NZE	57 (35 %)	46 (29 %)	58 (36 %)	50 (31 %)	58 (36 %)	53 (33 %)
Body composition						
Weight (kg)	79.5 [65.8, 98.3]	78.0 [67.5, 95.0]	74.0 [63.9, 96.0]	72.9 [64.4, 87.9]	77.6 [66.2, 95.8]	85.1 [66.5, 104.8]++
Height (cm)	166.8[163.7, 171.9]	167.7 [163.3, 171.7]	168.4 [164.5, 172.6]	166.8 [164.4, 171.6]	168.2 [163.6, 171.4]	168.6 [162.9, 172.9]
BMI (kg/m^2)	29.4 [23.5, 34.9]	29.1 [23.4, 33.2]	26.0 [22.6, 33.1]	25.6 [22.6, 32.2]	28.5 [23.0, 33.1]	31.1 [23.4, 36.9]++
Total body fat %	35.5 [29.5, 41.1]	34.4 [29.4, 39.4]	32.4 [27.4, 37.5]++	33.2 [28.7, 36.9]	34.4 [29.2, 38.8]	36.5 [28.7, 42.0]+
Visceral fat %	33.5 [24.5, 41.3]	33.7 [25.2, 38.1]	30.0 [21.4, 38.1]+	30.6 [23.3, 36.8]	32.0 [24.5, 40.3]	35.0 [22.8, 42.6]+
Android fat %	35.6 [27.2, 42.9]	35.3 [27.4, 39.7]	31.5 [24.1, 39.3]+	32.5 [25.7, 38.3]	34.4 [26.8, 41.4]	37.4 [26.0, 44.5]+
Gynoid fat %	38.2 [35.6, 42.0]	37.2 [33.8, 42.2]	36.4 [33.2, 40.7]+	37.0 [34.6, 40.5]	36.6 [33.5, 41.0]	38.5 [34.1, 43.2]
Blood pressure €						
Systolic (<120 mmHg)	117.0[107.5, 125.5]	115.5 [109.5, 121.5]	111.3 [106.0, 121.0]	112.5 [106.0, 120.0]	115.0 [108.5, 122.5]	117.5[108.5, 126.5]++
Diastolic (<80 mmHg)	75.5 [68.5, 83.0]	73.5 [67.0, 79.5]	72.0 [67.8, 79.0]	71.5 [66.5, 78.5]	74.0 [68.0, 80.0]	74.5 [68.5, 83.5]+
Metabolic markers∞						
TC (<5 mmol/L)-	4.8 [4.5, 5.4]	4.9 [4.3, 5.6]	4.7 [4.1, 5.4]	4.8 [4.4, 5.4]	4.8 [4.2, 5.4]	4.8 [4.4, 5.7]
HDL-C (>1 mmol/L)-	1.5 [1.3, 1.8]	1.6 [1.3, 1.9]	1.6 [1.3, 1.8]	1.5 [1.3, 1.9]	1.6 [1.3, 2.0]	1.6 [1.3, 1.8]
LDL-C (0-3.4 mmol/L)-	3.0 [2.6, 3.5]	3.0 [2.4, 3.6]	2.9 [2.3, 3.4]	3.0 [2.4, 3.6]	2.9 [2.4, 3.3]	3.0 [2.4, 3.6]
TGS (<2 mmol/L)-	0.9 [0.7, 1.2]	0.9 [0.7, 1.1]	0.9 [0.6, 1.1]	0.9 [0.7, 1.1]	0.9 [0.7, 1.2]	0.9 [0.7, 1.2]
TC:HDL (<4 mmol/L)-	3.1 [2.7, 3.6]	3.1 [2.5, 3.8]	2.9 [2.5, 3.5]	3.0 [2.5, 3.8]	3.0 [2.5, 3.5]	3.3 [2.7, 4.0]
HbA1c (<40 mmol/mol)-#	31.4 [29.7, 33.6]	32.1 [30.2, 34.1]	31.5 [30.2, 33.1]	31.7 [30.0, 33.3]	31.9 [29.7, 34.1]	31.6 [30.2, 33.6]
Glucose (3.5-5.4 mmol/L)-	5.3 [5.0, 5.6]	5.3 [5.1, 5.7]	5.2 [5.0, 5.5]	5.3 [5.0, 5.5]	5.3 [5.1, 5.6]	5.3 [5.0, 5.7]
Insulin (3-25 uU/mL)-	11.9 [7.8, 18.2]	11.1 [7.7, 17.3]	10.2 [6.5, 15.0]	10.9 [7.4, 14.1]	10.8 [7.2, 17.3]	11.1 [6.9, 20.6]
HOMA-IR	2.8 [1.8, 4.5]	2.6 [1.8, 4.1]	2.3 [1.5, 3.8]	2.6 [1.8, 3.4]	2.6 [1.7, 4.0]	2.5 [1.6, 5.0]
NZDep2013	6 [3, 8]	5 [3, 9]	5 [3, 8]	6 [4, 9]	5 [3, 8]	5 [3, 8]

All values are reported as median [25th, 75th]. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. T1: Tertile 1; *n* = 95, T2: Tertile 2; *n* = 96, T3: Tertile 3; *n* = 96. T1= negatively adhering to the dietary pattern, T3 = positively adhering to the dietary pattern. Differences between tertiles within a dietary pattern were tested by Mann-Whitney test. * Significant difference between tertiles 1 and 2, $P < 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***. + Significant difference between tertiles 1 and 3, $P < 0.05$ +, $P \leq 0.01$++, $P \leq 0.001$+++, ^Significant difference between tertiles 2 and 3, $P < 0.05$ ^, $P \leq 0.01$^^, $P \leq 0.001$ ^^^. **€** NZE women (*n* = 2), ∞ Pacific woman (*n* = 1), and #Pacific women (*n* = 2) have not been included in analysis due to missing data. NZE: New Zealand European, BMI: body mass index. TC: Total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TGS: Triglycerides, HOMA-IR: Homeostasis model assessment index for insulin resistance (HOMA-IR), HbA1c: glycosylated haemoglobin, DXA used to measure all body fat content. -Normal healthy ranges for metabolic markers (53)

5.4.7. Percentage of total energy consumed from macronutrients

Protein intake

There was a positive association with higher adherence to the "colourful vegetables, plant protein, and dairy" and the "animal meat and fat" pattern, and the percentage of total energy consumed from protein. In comparison to the Acceptable Macronutrient Distribution Ranges (AMDRs) (41) the median percentage of protein intake was within the AMDRs (15-25 % of total energy intake) for all four dietary patterns (Table 5.6, Supplementary Figure 5.1, Supplementary Figure 5.2, Supplementary Figure 5.3, Supplementary Figure 5.4).

Total and saturated fat intake

Higher adherence to the "colourful vegetables, plant protein, and dairy", "sweet and fat rich carbohydrate", and "animal meat and fat" patterns were all positively associated with percentage of energy consumed from total and saturated fat. However, the "fruit, starchy vegetables, and nuts" pattern was inversely associated with the percentage of energy consumed from saturated fat. For all dietary patterns, the percentage of energy intake from total and saturated fat was above the AMDRs (20-35 % and <10 % of total energy intake respectively) (Table 5.6, Supplementary Figure 5.1, Supplementary Figure 5.2, Supplementary Figure 5.3, Supplementary Figure 5.4).

Total carbohydrate and dietary fibre intake

Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was inversely associated with the percentage of energy consumed from total carbohydrate. In contrasts the "sweet and fat rich carbohydrate" pattern was positively associated with the percentage of energy consumed from total carbohydrate. In comparison to the AMDRs, the percentage of energy consumed from total carbohydrate was below the recommended 45-65 % percentage of total energy intake, for all four dietary patterns (Table 5.6, Supplementary Figures 1-4). Higher adherence to the "colourful vegetables, plant protein, and dairy" and "fruit, starchy vegetables, and nuts" patterns was positively associated with the percentage of energy consumed from dietary fibre intake (Table 5.6, Supplementary Figure 5.1, Supplementary Figure 5.2, Supplementary Figure 5.3, Supplementary Figure 5.4).

Table 5.6. Acceptable Macronutrient Distribution Ranges (AMDRs) of the four habitual dietary patterns stratified by tertiles of adherence

Nutrient	Pattern 1: Colourful vegetables, plant protein, and dairy			Pattern 2: Sweet and fat rich carbohydrate			Pattern 3: Fruit, starchy vegetables, and nuts			Pattern 4: Animal meat and fat		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
	(% of EI)			(% of EI)			(% of EI)			(% of EI)		
Energy kJ/day	8542.8 [8064.0, 8855.9]	8524.8 [8108.6, 8857.0]	8583.1 [8187.6, 8816.2]	8586.9 [8192.1, 8854.7]	8572.7 [8198.0, 8823.9]	8398.3 [8091.0, 8871.7]	8561.2 [8096.9, 8848.8]	8570.2 [8174.4, 8895.3]	8498.6 [8144.4, 8811.1]	8699.5 [8347.4, 8938.0]	8554.7 [8191.8, 8851.7]^	8324.2 [8038.3, 8766.8]+++
Protein 15-25 % [a]	15.1 [13.3, 17.3]**	16.7 [14.9, 18.1]	16.8 [15.4, 18.2]+++	16.6 [13.9, 18.2]	15.9 [14.2, 17.4]	16.5 [15.0, 18.3]	16.3 [14.2, 18.4]	16.1 [14.3, 17.4]	16.6 [14.9, 18.2]	15.2 [13.6, 16.7]***	15.9 [14.4, 17.4]^^	17.6 [16.3, 19.0]+++
Total fat 20-30 % [a]	38.0 [31.5, 43.2]	39.1 [32.8, 45.1]^^	42.0 [36.4, 49.1]+++	39.3 [32.6, 46.6]	38.5 [33.7, 43.0]^	40.6 [36.4, 46.2]	40.1 [35.7, 45.9]	39.0 [34.8, 44.6]	39.4 [33.0, 46.5]	39.1 [32.6, 46.6]	38.3 [33.3, 45.1]^	42.0 [36.1, 46.4]
SFA <10 % [a]	14.8 [12.1, 16.8]	14.4 [12.2, 16.6]^^	15.6 [13.9, 19.1]++	14.3 [11.7, 17.6]	14.7 [12.6, 16.7]^^	15.9 [14.2, 18.1]++	16.3 [13.7, 18.3]	14.9 [12.8, 17.2]	14.3 [12.1, 16.3]+++	14.4 [11.7, 16.8]	14.6 [12.6, 16.8]^	15.8 [14.0, 18.2]++
CHO 45-65 % [a]	43.4 [34.1, 47.2]*	39.1 [33.5, 43.7]^^^	34.0 [29.4, 38.1]+++	32.4 [27.8, 38.5]***	38.0 [33.0, 43.4]^^^	42.0 [37.7, 46.4]+++	38.3 [33.0, 45.6]	38.0 [31.2, 44.2]	36.7 [32.6, 42.8]	38.7 [33.8, 45.6]	37.4 [31.7, 43.6]	36.6 [31.2, 44.0]
Dietary fibre	1.7 [1.3, 1.9]***	2.0 [1.7, 2.4]^^	2.3 [1.9, 2.7]+++	1.9 [1.6, 2.5]	1.9 [1.6, 2.4]	2.0 [1.8, 2.3]	1.8 [1.5, 2.1]	1.9 [1.6, 2.1]^^^	2.3 [1.9, 2.8]+++	2.0 [1.7, 2.5]	2.0 [1.7, 2.4]	1.9 [1.6, 2.20]

All values are reported as medians [25th, 75th percentiles]. Habitual nutrient intake calculated with the National Cancer Institute method (36). Daily nutrient intakes estimated for a period of one month. Total *n* = 287 in each dietary pattern; Pacific *n* = 126 and NZE *n* = 161. T1: Tertile 1; *n* = 95, T2: Tertile 2; *n* = 96, T3: Tertile 3; *n* = 96. T1= negatively adhering to the dietary pattern, T3 = positively adhering to the dietary pattern. Differences between tertiles within a dietary pattern were assessed with Mann-Whitney test. * Significant difference between tertiles 1 and 2, $P < 0.05$ *, $P \leq 0.01$ **, $P \leq 0.001$ ***. + Significant difference between tertiles 1 and 3, $P < 0.05$ +, $P \leq 0.01$++, $P \leq 0.001$+++, ^Significant difference between tertiles 2 and 3, $P < 0.05$ ^, $P \leq 0.01$^^, $P \leq 0.001$ ^^^. Total fat includes all saturated and unsaturated fats. SFA = Saturated fat, CHO = total carbohydrate. %EI: Percentage of energy intake (42). [a]AMDR: Acceptable Macronutrient Distribution Ranges based on New Zealand and Australian reference guidelines (41).

5.5. Discussion

The overall objective of the present study was to explore the association between dietary patterns, body fat content, and metabolic health markers in a population of healthy Pacific and NZE women with different metabolic disease risk and body fat profiles. The patterns were derived from two independent dietary intake assessment methods (5DFR and NZWFFQ), which captured multiple days of actual and habitual intake (respectively), and were used to estimate habitual dietary intake employing the NCI method. We identified four dietary patterns which explained 30.9 % of the variance in the Pacific and NZE women's habitual diet. Higher adherence to dietary patterns, characterised by core foods, the "colourful vegetables, plant protein, and dairy" and "fruit, starchy vegetables, and nuts" patterns were inversely associated with BF%. In contrast, higher adherence to the "sweet and fat rich carbohydrate" pattern characterised by higher discretionary food intake, and the "animal meat and fat" pattern characterised by less diversity of core foods, were positively associated with BF% for both Pacific and NZE women. Adjusting for ethnicity, age, energy intake, or socio-economic deprivation did not significantly influence the observed associations with adiposity. Suggesting, despite Pacific and NZE women having different metabolic disease risk, the impact of habitual food choice and subsequent risk of obesity is comparable. In addition, across all dietary patterns, the percentage of energy intake from total and saturated fat intake exceeded the AMDRs, whereas the percentage of total carbohydrate intake was below the AMDRs. The characteristics of the different dietary patterns are discussed below.

Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was associated with lower adiposity (BF%, VAT, and AF%) including lower fasting insulin and HOMA-IR. In contrast, the opposite association was observed with higher adherence to the "sweet and fat rich carbohydrate" pattern characterised by higher loadings of 'discretionary' food groups. Women who adhered to the "sweet and fat rich carbohydrate" pattern had higher levels of adiposity (BF%, VAT, and AF%) and showed clear evidence of hyperinsulinemia. This is an important observation because higher abdominal fat mass is associated with increased metabolic disease risk (43,44), and hyperinsulinemia can precede hyperglycaemia (45) on the path to developing reduced insulin sensitivity or insulin resistance. In comparison to Europeans, previous studies have found that Pacific people have more

abdominal fat (46), which is suggested to contribute to the differences in the prevalence of diabetes between Pacific and European populations (47,48). Despite Pacific women having higher fasting insulin, glucose, and HbA1c concentrations in comparison to NZE women, both populations were within the normal healthy ranges (49). Importantly, the observations from the present study suggest, that higher adherence to the "sweet and fat rich carbohydrate" pattern (which was characterised by both ethnicities) is linked to hyperinsulinemia and a long-term elevation of blood glucose concentrations, although within the normal range (49). Therefore, despite having different metabolic disease risks, habitual diet is associated with the same metabolic disease risks for these women.

Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was characterised by higher loadings of 'core' food groups, and older NZE women with lower levels of socio-economic deprivation and adiposity. Beck et al. (16) observed a similar trend at the population level for NZ, where higher adherence the 'healthy' pattern (characterised by breakfast cereal, fruit, low fat milk, yoghurt, and soups and stocks) was associated with older NZE women with lower socio-economic deprivation levels and BMI. Higher quality diets are consistently associated with higher socioeconomic status (SES) (50). Although information regarding the cost of food was not assessed, these observations align with global and local trends, where healthier eating patterns are associated with older women with higher SES (10,16,50), who also have a reduced metabolic disease risk. In contrast, higher 'discretionary' food intake is associated with an increased metabolic disease risk (3,4,10,15,51) and, in the present study, women with higher adherence to the "sweet fat and carbohydrate" pattern (characterised by discretionary food intake) had higher multifactorial metabolic disease risks.

Higher socio-economic deprivation is associated with consuming more 'discretionary' inexpensive energy dense and nutrient poor foods to save costs (52). However, we observed no difference in the level of socio-economic deprivation with adherence to the "sweet fat and carbohydrate" pattern. It has been proposed that globally the relative cheapness of discretionary foods is driving the increasing rates of obesity in high income countries (53,54). Adding to this, at the population level, recent global sales of ultraprocessed foods and beverages were positively associated with BMI trajectories (55). Ultraprocessed foods and beverages are defined as ready to consume products which are formulated from substances

derived from foods and additives, but contain minimal intact food (e.g., hydrogenated oils, fats, and highly processed starches and sugars) (54,56). Taken together, higher SES is associated with consuming more core foods (50), which in turn is associated with lower metabolic disease risks. In contrast, lower SES is associated with higher discretionary food intake (52), and, consequently, increased metabolic disease risks. However, the relative cheapness and availability of discretionary processed foods is likely driving global consumption irrespective of SES (53–55), and thus the burden of NCDs (e.g., obesity, T2DM) associated with higher discretionary food intake.

The recent 2017 Global Burden of Disease (GBD) (3) study emphasised the impact of 'core' versus 'discretionary' food and nutrient intake on health outcomes in 195 countries worldwide. Suboptimal diet (defined as higher intake of discretionary foods and nutrients and inadequate intake of core or 'healthy' foods and nutrients) was associated with increased risk of premature death globally (3). Higher 'discretionary' food intake is also a hallmark characteristic of the following dietary patterns: 'Western' (57,58), 'sweet-traditional' (59), 'meat and fried food' (60), 'meat and soda' (61), and 'meat' (62), all of which have been associated with increased risk of weight gain and/or obesity across global populations. In contrast, 'core' foods consistently characterise dietary patterns which are associated with a reduced risk of higher adiposity such as 'healthier' (61), 'vegetables and fruit' (60), 'prudent' (57,58,62) and 'green' (59) patterns. Thus, the observations of higher adherence to patterns characterised by discretionary foods and limited diversity of core foods being associated with higher adiposity, while higher adherence to dietary patterns characterised by core foods being associated with lower adiposity, aligns with global trends. Furthermore, the observations of this study emphasise that habitual intake of these foods are associated with the same metabolic disease risks for populations who have different metabolic disease risk and body fat profiles.

In the present study, higher adherence to dietary patterns characterised by higher loadings of core foods and lower loadings of discretionary foods was associated with lower adiposity (BF%, weight, BMI) and TC:HDL ratio. Opposed to dietary patterns which were characterised by higher loadings of discretionary foods and lower diversity of core food intake, which were associated with higher adiposity (BF%, weight, BMI) and TC:HDL ratio. Globally, adherence to

patterns characterised by 'core' and 'discretionary' foods are associated with lower and higher adiposity respectively (10,12,13,57–63). Highlighting, that although the assigned names and combinations of the food groups which characterise the patterns in the present study differ to those reported in the literature (likely reflective of the nutrition practices and food systems in NZ), overall the observations of this study are in agreement with the health associations reported in the literature.

A series of systematic reviews conducted by The United States Department of Agriculture (USDA) further emphasised there is not one dietary pattern more favourably associated with health outcomes (13). The USDA highlighted that certain foods consistently characterise patterns associated with similar health outcomes (13), which is in agreement with Liese et al.'s (15), observations from the DPMP. For example, across populations, higher adherence to dietary patterns characterised by fruit, vegetables, wholegrains, legumes, and moderate intake of alcohol and dairy are associated with a more favourable weight and a reduced risk of obesity (13). In contrast, higher adherence to dietary patterns characterised by meat, saturated fat, sugar sweetened foods and beverages are associated with an increased risk of obesity (13). Therefore, these observations suggest that the common drivers of dietary quality (and subsequent metabolic health risks) identified across different global populations, are relevant to NZ, and populations with different metabolic disease risk and body fat profiles.

An unexpected observation across all four dietary patterns was the percentage of total energy intake consumed from total and saturated fat found to be in excess of the AMDRs, coupled with a lower percentage of total energy intake consumed from total carbohydrate in comparison to the AMDRs. Typically, higher intake of one macronutrient displaces the intake of another, and a similar trend of a lower percentage of energy intake consumed from carbohydrate, coupled with a higher saturated fat intake, has been observed in another population of NZE women (20). The low carbohydrate intake may, in part, be explained by how a low carbohydrate diet is consistently publicised as a diet to lose or maintain weight (64–66), and women are more likely to attempt to control their weight through dietary manipulation (9,23). The exclusion criteria for the present study encompassed restrictive or extreme diets (25). However, during the data collection period of the PROMISE study there was substantial discourse on social media platforms, media outlets, and publications within

NZ discussing the perceived optimal 'balance' of total carbohydrate and total fat (67,68) for health. We cannot, therefore, exclude the potential for this widespread dietary discourse to have influenced the dietary trends of the study population.

From a public health perspective, high saturated fat intake is associated with increased risk of CVD (69,70), and thus current recommendations are to limit and replace saturated fat with unsaturated fats (13,69,70). Dietary patterns rich in foods containing mono- and poly-unsaturated fats (nuts and seeds, fish etc.), such as the Mediterranean diet, are associated with a relatively high "healthy" fat intake and consequently a reduced risk of cardiometabolic diseases (71). Within the present study, the dietary patterns with high loadings of nuts and seeds (vegetables and 'core' foods) were associated with lower adiposity (BF%, VAT, and AF%), while consuming a high percentage of energy from total fat intake. Therefore, it is important to explore the quality of the food source and combinations of the nutrients consumed together in relation to health outcomes, highlighting the benefits of dietary pattern analysis.

Further emphasising the impact of the quality of the food source consumed, only the "colourful vegetable, plant protein, and dairy" and "fruit, starchy vegetable, and nuts" patterns were characterised by a positive association with the percentage of energy consumed from dietary fibre intake. Plant based foods (vegetables, fruit, and wholegrains) are rich sources of indigestible polysaccharides (cellulose, hemicellulose, and resistant starch) which are collectively defined as dietary fibre (72). Dietary fibre is almost completely broken down in a process called fermentation by the gut microbiota in the large bowel (73). The gut microbiota is suggested to play an important role in human health through the digestion of fibre (74), and the health promoting benefits of higher habitual dietary fibre intake (notably from wholegrains and cereals) have been well documented (75,76). A recent systematic review identified a dose-response relationship where higher intakes of dietary fibre or wholegrains were considered cardioprotective reducing the risk of NCDs and all-cause mortality (75). In addition, higher habitual dietary fibre intake has been associated with a lower weight and BF% (63,76,77). Therefore, within the present study, dietary fibre intake could be considered a marker of diet quality, because higher adherence to patterns positively associated with dietary fibre intake were associated with lower overall adiposity. However, these observations also emphasise the importance of exploring the food sources of the energy

(and nutrients) and how these are consumed together. Across all dietary patterns there was not a substantial difference in the proportion of energy intake consumed from macronutrients, but there was in the types of foods which contributed to nutrient intakes ('core' versus 'discretionary') and subsequent metabolic disease risks (decreased and increased respectively).

The present study has several strengths; namely, the detailed dietary data. The research personnel went to great lengths to reduce the likelihood of misreporting throughout the data collection and entering processes (78). Most dietary pattern studies use retrospective reporting of dietary intake with FFQs (17–22,57,59–61) which rely on memory and thus are prone to overestimation (9,23,24). In comparison to a FFQ, utilising a prospective food record can reduce the likelihood of misreporting of dietary intake (24), while enabling culturally specific foods to be captured that a FFQ might fail to capture (9,23). In addition, FFQs are unable to consider person specific effects (e.g., covariates), which can influence habitual intake. However, the NCI method specifically addressed the intra individual day-to-day variation inherent in self-reported dietary, and the inclusion of covariates in the NCI modelling process improves the estimation of episodically consumed foods (79,80).

In addition, the *a posteriori* dietary patterns identified in this study had similar characteristics to other patterns reported in the literature and similar associations in relation to the risk of obesity (57,61,62). The four dietary patterns explained 30.9 % of the observed variance in habitual intake and we were able to correlate the nutrient and food group intake associated with adherence to the patterns, giving substantial logical depth to the dietary patterns we identified. Furthermore, this study was unique in its selective recruitment of healthy premenopausal Pacific and NZE women whom have different metabolic disease risk and body fat profiles. We also assessed several measures of body composition (BF%, VAT, and AF%), and metabolic health markers (e.g., fasting insulin, glucose, TC, TC:HDL) to further explore the associations between metabolic health and dietary patterns. Stratifying the population by BF% enabled a more objective assessment of metabolic health, as individuals with the same BMI can have different body compositions and metabolic disease risks (39,40).

There are limitations of the present study which need to be considered. Self-reported dietary data were used as the primary source of dietary information to calculate habitual dietary intake. Although, self-reported data is prone to under and over-reporting (9,23) it offers insight into the complexities of how people are eating that no current set of biomarkers currently are able to do (81). The energy reporting cut offs of <2100 kJ/day and >27000 kJ/day employed within this study were adapted and higher than the parameters recommended by Willett (9); however, higher cut-offs have similarly been used by other studies within ethnic minority groups (82,83). In addition, all dietary intakes were intensively checked and considered as plausible values for the study population (31–33). Further, the NCI method specifically addresses individual variation in day-to-day intake in its modelling process (79). Using *a posteriori* method to identify dietary patterns does mean they are specific to the present study population, and that observations cannot be generalised to other populations. Furthermore, due to the study design we cannot confirm whether these dietary patterns are the main driver of BF%, as there are further factors besides diet that influence metabolic health risk, such as a healthy lifestyle in general. However, there was no significant difference in BF% between Pacific and NZE women (data not shown), but there was when stratifying the population by tertiles of adherence to dietary patterns.

5.6. Conclusion

This study provides greater understanding of adherence to dietary patterns, food choice, and metabolic health risks in a population of women who have different metabolic disease risks. Higher adherence to the dietary patterns characterised by 'core' food groups ("colourful vegetables, plant protein, and dairy" and "fruit, starchy vegetables, and nuts") were associated with lower BF% and metabolic health risks. In contrast, higher adherence to patterns characterised by 'discretionary' foods ("sweet and fat rich carbohydrate") and less diversity of core foods ("animal meat and fat") were associated with higher BF%, for both Pacific and NZE women who have different metabolic disease risk and body fat profiles. The observed trend for the entire population to consume high fat and low carbohydrate intake highlights the importance of exploring the food source of the nutrients, and the combination of the foods consumed together in relation to health outcomes. Similar to global trends, we observed that dietary patterns characterised by 'core' foods were associated with lower

metabolic health risk, substantiating that the core features of a healthy diet can be achieved through many different patterns.

In the context of NZ having the third highest rate of obesity in the OECD, the potentially detrimental effect of higher discretionary food intake highlights a public health priority for NZ, as well as for the global population. Notably, when reflecting on the ease of access and relative cheapness of discretionary foods, and the observations of higher multifactorial metabolic risks factors (irrespective of ethnicity, SES, or BMI) associated with higher intake of discretionary foods. Combined, it emphasises the need for local and global food systems, industries and environments, to support the healthy choice being the easy choice. Future longitudinal research taking a habitual food-based approach with multiple assessment points would be required to establish whether adherence to such patterns perpetuate negative metabolic associations over time. However, current evidence suggests to reduce the global burden of diseases attributable to suboptimal dietary intakes, we need to habitually eat more 'core' foods. Dietary patterns are a practical starting point to explore the synergies of different combinations of foods and their health effect, as nutrients are not consumed in isolation. Dietary pattern research can guide future research directions to establish causality and develop therapeutic targets, to ultimately start addressing one of the key modifiable risk factors in the obesity epidemic, which is our habitual diet.

5.7. References

1. World Health Organization. Obesity and overweight [Internet]. 2018 [cited 2019 Apr 12].
2. Willett WC, Stampfer MJ. Current evidence on healthy eating. Annu Rev Public Health [Internet]. 2013 Mar 18;34(1):77–95. doi:10.1146/annurev-publhealth-031811-124646
3. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8
4. Wang DD, Li Y, Afshin A, Springmann M, Mozaffarian D, Stampfer MJ, et al. Global improvement in dietary quality could lead to substantial reduction in premature death. J Nutr [Internet]. 2019;149(6):1065–74. doi:10.1093/jn/nxz010
5. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].
6. Organization for Economic Co-operation and Development (OECD). OECD Obesity update 2017. Paris, France; 2017.
7. Hu FB. Dietary pattern analysis: A new direction in nutritional epidemiology. Curr Opin Lipidol [Internet]. 2002 Feb;13(1):3–9. doi:10.1097/00041433-200202000-00002
8. Tapsell LC, Neale EP, Satija A, Hu FB. Foods, nutrients, and dietary patterns: interconnections and implications for dietary guidelines. Adv Nutr. 2016;7(3):445–54. doi:10.3945/an.115.011718
9. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001
10. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. Nutr Rev. 2004;62(5):177–203. doi:10.1301/nr.2004.may.177
11. Moeller SM, Reedy J, Millen AE, Dixon LB, Newby PK, Tucker KL, et al. Dietary patterns: Challenges and opportunities in dietary patterns research. J Am Diet Assoc. 2007;107(7):1233–9. doi:10.1016/j.jada.2007.03.014
12. Schwingshackl L, Hoffmann G. Diet quality as assessed by the healthy eating index, the alternate healthy eating index, the dietary approaches to stop hypertension score, and health outcomes: A systematic review and meta-analysis of cohort studies. J Acad Nutr Diet [Internet]. 2015;115(5):780-800.e5. doi:10.1016/j.jand.2014.12.009
13. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.
14. Tucker KL. Dietary patterns, approaches, and multicultural perspective. Appl Physiol Nutr Metab. 2010;35:211–8. doi:10.1139/H10-010
15. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhouser ML, et al. The dietary patterns methods project: synthesis of findings across cohorts and relevance to

dietary guidance. J Nutr [Internet]. 2015;145(3):393–402. doi:10.3945/jn.114.205336

16. Beck KL, Jones B, Ullah I, McNaughton SA, Haslett SJ, Stonehouse W. Associations between dietary patterns, socio-demographic factors and anthropometric measurements in adult New Zealanders: An analysis of data from the 2008/09 New Zealand Adult Nutrition Survey. Eur J Nutr. 2018;57(4):1421–33. doi:10.1007/s00394-017-1421-3

17. Beck KL, Kruger R, Conlon CA, Heath ALM, Matthys C, Coad J, et al. Suboptimal iron status and associated dietary patterns and practices in premenopausal women living in Auckland, New Zealand. Eur J Nutr. 2013;52(2):467–76. doi:10.1007/s00394-012-0348-y

18. Wall CR, Gammon CS, Bandara DK, Grant CC, Atatoa Carr PE, Morton SMB. Dietary patterns in pregnancy in New Zealand-influence of maternal socio-demographic, health and lifestyle factors. Nutrients. 2016;8(5). doi:10.3390/nu8050300

19. Thompson JMD, Wall C, Becroft DMO, Robinson E, Wild CJ, Mitchell EA. Maternal dietary patterns in pregnancy and the association with small-for-gestational-age infants. Br J Nutr. 2010;103(11):1665–73. doi:10.1017/S0007114509993606

20. Schrijvers JK, McNaughton SA, Beck KL, Kruger R. Exploring the dietary patterns of young New Zealand women and associations with BMI and body fat. Nutrients. 2016;8(8):1–15. doi:10.3390/nu8080450

21. Saeedi P, Black KE, Haszard JJ, Skeaff S, Stoner L, Davidson B, et al. Dietary patterns, cardiorespiratory and muscular fitness in 9-11-year-old children from Dunedin, New Zealand. Nutrients. 2018;10(7):1–13. doi:10.3390/nu10070887

22. Howe AS, Black KE, Wong JE, Parnell WR, Skidmore PM. Dieting status influences associations between dietary patterns and body composition in adolescents: A cross-sectional study. Nutr J. 2013;12(1):1–10. doi:10.1186/1475-2891-12-51

23. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.

24. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

25. Kindleysides S, Kruger R, Douwes J, Tannock G, Renall N, Slater J, et al. PRedictors linking Obesity and the gut MIcrobiomE (The PROMISE study): Protocol and recruitment strategy of a cross-sectional study on pathways that affect the gut microbiome and its impact on obesity. JMIR Res Protoc. 2019;8:1–16. doi:10.2196/14529

26. Stewart A, Marfell-Jones M, Olds T, Hans De Ridder J. International standards for anthropometric assessment. Lower Hutt, New Zealand: International Society for the Advancement of Kinanthropometry; 2011.

27. Garrow J. Quetelet's index (W/H2) as a measure of fatness. J Obes. 1985;9(2):147–53.

28. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: Insulin resistance and β-cell function from fasting plasma glucose

and insulin concentrations in man. Diabetologia. 1985;28(7):412–9. doi:10.1007/BF00280883

29. Atkinson J, Salmond C, Crampton P. NZDep2013 index of deprivation. NZDep2013 Index of Deprivation. Wellington; 2014.

30. Beck KL, Houston ZL, McNaughton SA, Kruger R. Development and evaluation of a food frequency questionnaire to assess nutrient intakes of adult women in New Zealand. Nutr Diet. 2018;1–7. doi:10.1111/1747-0080.12472

31. Gemming L, Jiang Y, Swinburn B, Utter J, Mhurchu CN. Under-reporting remains a key limitation of self-reported dietary intake: An analysis of the 2008/09 New Zealand Adult Nutrition Survey. Eur J Clin Nutr [Internet]. 2014;68(2):259–64. doi:10.1038/ejcn.2013.242

32. Smith C, Gray AR, Mainvil LA, Fleming EA, Parnell WR. Secular changes in intakes of foods among New Zealand adults from 1997 to 2008/09. Public Health Nutr [Internet]. 2015 Dec 10;18(18):3249–59. doi:10.1017/S1368980015000890

33. Metcalf PA, Scragg RRK, Schaaf D, Dyall L, Black PN, Jackson R. Dietary intakes of European, Māori, Pacific and Asian adults living in Auckland: The Diabetes, Heart and Health Study. Aust N Z J Public Health [Internet]. 2008;32(5):454–60. doi:10.1111/j.1753-6405.2008.00279.x

34. National Health and Medical Research Council. Australian Dietary Guidelines. Canberra: National Health and Medical Council; 2013.

35. Ministry of Health. Eating and activity guidelines for New Zealand adults [Internet]. Wellington; 2015.

36. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

37. Field A. Discovering statistics using SPSS. 3rd ed. London: SAGE Publications Ltd; 2009.

38. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr [Internet]. 1997 Apr 1;65(4):1220S-1228S. doi:10.1093/ajcn/65.4.1220S

39. Oliveros E, Somers VK, Sochor O, Goel K, Lopez-Jimenez F. The concept of normal weight obesity. Prog Cardiovasc Dis. 2014;56(4):426–33. doi:10.1016/j.pcad.2013.10.003

40. Kramer C, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions? A systematic review and meta-analysis. Ann Intern Med. 2013;159(11):758–69. doi:10.7326/0003-4819-159-11-201312030-00008

41. National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. Nutrient reference values for Australia and New Zealand. Nutrition. Canberra; 2006.

42. Food Standards Australia New Zealand (FSANZ). Australia New Zealand Food Standards Code-Section 11-Calculation of values for nutrition information panel [Internet]. F2018C00953. 2018 [cited 2019 Aug 20].

43. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y, et al.

Abdominal visceral and subcutaneous adipose tissue compartments. Circulation. 2007;116(1):39–48. doi:10.1161/circulationaha.106.675355

44. Després J-P. Body fat distribution and risk of cardiovascular disease. Circulation. 2012;126(10):1301–13. doi:10.1161/circulationaha.111.067264

45. Dickey RA, Bartuska DG, Bray GW, Callaway CW, Davidson ET, Feld S, et al. AACE/ACE position statement on the prevention, diagnosis, and treatment of obesity (1998 Revision) AACE/ACE Obesity Task Force. Endocr Pract. 1998;4(5):297–350.

46. Rush EC, Freitas I, Plank LD. Body size, body composition and fat distribution: Comparative analysis of European, Maori, Pacific Island and Asian Indian adults. Br J Nutr [Internet]. 2009 Aug 10;102(04):632. doi:10.1017/S0007114508207221

47. Ministry of Health. Quality standards for diabetes care toolkit. Wellington; 2014.

48. Coppell KJ, Mann JI, Williams SM, Jo E, Drury PL, Miller J, et al. Prevalence of diagnosed and undiagnosed diabetes and prediabetes in New Zealand: Findings from the 2008/09 adult nutrition survey. N Z Med J. 2013;126(1370):23–42.

49. New Zealand Guidelines Group. New Zealand primary care handbook 2012 [Internet]. 3rd ed. New Zealand Primary Care Handbook. Wellington: Ministry of Health; 2012. 45–64 p.

50. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr [Internet]. 2008 May 1;87(5):1107–17. doi:10.1093/ajcn/87.5.1107

51. 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 [Internet]. 2018;360:j5644. doi:10.1136/bmj.j5644

52. Darmon N, Briend A, Drewnowski A. Energy-dense diets are associated with lower diet costs: A community study of French adults. Public Health Nutr. 2004;7(1):21–7. doi:10.1079/PHN2003512

53. Headey DD, Alderman HH. The relative caloric prices of healthy and unhealthy foods differ systematically across income levels and continents. J Nutr [Internet]. 2019 Nov 1;149(11):2020–33. doi:10.1093/jn/nxz158

54. Monteiro CA, Moubarac J, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. Obes Rev. 2013;14:21–8. doi:10.1111/obr.12107

55. Vandevijvere S, Jaacks LM, Monteiro CA, Moubarac J, Girling-Butcher M, Lee AC, et al. Global trends in ultraprocessed food and drink product sales and their association with adult body mass index trajectories. Obes Rev [Internet]. 2019 Nov 17;20(S2):10–9. doi:10.1111/obr.12860

56. Monteiro CA, Cannon G, Moubarac JC, Levy RB, Louzada MLC, Jaime PC. The UN decade of nutrition, the NOVA food classification and the trouble with ultra-processing. Public Health Nutr. 2018;21(1):5–17. doi:10.1017/S1368980017000234

57. Schulze MB, Fung TT, Manson JE, Willett WC, Hu FB, Matthias B, et al. Dietary patterns and changes in body weight in women. Obesity. 2006;14(8):1444–53.

doi:10.1038/oby.2006.164

58. Duffey KJ, Steffen LM, Van Horn L, Jacobs DR, Popkin BM. Dietary patterns matter: Diet beverages and cardiometabolic risks in the longitudinal coronary artery risk development in young adults (CARDIA) study. Am J Clin Nutr [Internet]. 2012 Apr 1;95(4):909–15. doi:10.3945/ajcn.111.026682

59. Togo P, Osler M, Sørensen TIA, Heitmann BL. A longitudinal study of food intake patterns and obesity in adult Danish men and women. Int J Obes. 2004;28:583–93. doi:10.1038/sj.ijo.0802598

60. Boggs DA, Palmer JR, Spiegelman D, Stampfer MJ, Adams-campbell LL, Rosenberg L. Dietary patterns and 14-y weight gain in African American women. Am J Clin Nutr. 2011;94(1):86–94. doi:10.3945/ajcn.111.013482.INTRODUCTION

61. Pachucki MA. Food pattern analysis over time: unhealthful eating trajectories predict obesity. Int J Obes. 2012;36:686–94. doi:10.1038/ijo.2011.133

62. Tucker LA, Tucker JM, Bailey BW, Lecheminant JD. Dietary patterns as predictors of body fat and BMI in women: A factor analytic study. Am J Heal Promot. 2015;29(4):136–47. doi:10.4278/ajhp.130327-QUAN-129

63. Schlesinger S, Neuenschwander M, Schwedhelm C, Hoffmann G, Bechthold A, Boeing H, et al. Food groups and risk of overweight, obesity, and weight gain: A systematic review and dose-response meta-analysis of prospective studies. Adv Nutr. 2019;10(2):205–18. doi:10.1093/advances/nmy092

64. Astrup A, Meinert Larsen T, Harper A. Atkins and other low-carbohydrate diets: Hoax or an effective tool for weight loss? Lancet. 2004;364:897–9. doi:10.1016/S0140-6736(04)16986-9

65. De Souza RJ, Bray GA, Carey VJ, Hall KD, LeBoff MS, Loria CM, et al. Effects of 4 weight-loss diets differing in fat, protein, and carbohydrate on fat mass, lean mass, visceral adipose tissue, and hepatic fat: Results from the POUNDS LOST trial. Am J Clin Nutr. 2012;95(3):614–25. doi:10.3945/ajcn.111.026328

66. Samaha FF, Iqbal N, Seshadri P, Chicano KL, Daily DA, McGrory J, et al. A low-carbohydrate as compared with a low-fat diet in severe obesity. N Engl J Med. 2003;348(21):2074–81. doi:10.1056/NEJMoa022637

67. Mann J, Mclean R, Skeaff M, Te Morenga L. Low carbohydrate diets: Going against the grain. Lancet. 2014 Oct 25;384:1479–80. doi:10.1016/S0140-6736(14)61413-6

68. Schofield G, Henderson G, Thornley S, Crofts C. Very low-carbohydrate diets in the management of diabetes revisited. N Z Med J. 2016;129(1432):67–73.

69. Schwab U, Lauritzen L, Tholstrup T, Haldorssoni T, Riserus U, Uusitupa M, et al. Effect of the amount and type of dietary fat on cardiometabolic risk factors and risk of developing type 2 diabetes, cardiovascular diseases, and cancer: A systematic review. Food Nutr Res. 2014;58. doi:10.3402/fnr.v58.25145

70. Astrup A, Bertram HCS, Bonjour JP, De Groot LCP, De Oliveira Otto MC, Feeney EL, et al. WHO draft guidelines on dietary saturated and trans fatty acids: Time for a new approach? BMJ [Internet]. 2019;366(July):1–6. doi:10.1136/bmj.l4137

71. Estruch R, Ros E, Salas-Salvadó J, Covas M-I, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a mediterranean diet supplemented with extra-virgin olive oil or nuts. N Engl J Med [Internet]. 2018;378(25):e34. doi:10.1056/NEJMoa1800389

72. Mann J, Cummings JH, Englyst HN, Key T, Liu S, Riccardi G, et al. FAO/WHO Scientific update on carbohydrates in human nutrition: Conclusions. Eur J Clin Nutr. 2007;61(S1):S132–7. doi:10.1038/sj.ejcn.1602943

73. Cummings JH, Stephen AM. The role of the dietary fibre in the human colon. Can Med Assoc J. 1980;123(11):1109–14.

74. Shanahan F, Van Sinderen D, O'Toole PW, Stanton C. Feeding the microbiota: Transducer of nutrient signals for the host. Gut. 2017;66:1709–17. doi:10.1136/gutjnl-2017-313872

75. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9

76. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

77. Tucker LA, Thomas KS. Increasing total fiber intake reduces risk of weight and fat gains in women. J Nutr. 2009;139(3):576–81. doi:10.3945/jn.108.096685

78. Cade JE, Warthon-Medina M, Albar S, Alwan NA, Ness A, Roe M, et al. DIET@NET: Best practice guidelines for dietary assessment in health research. BMC Med [Internet]. 2017 Dec 15;15(1):202. doi:10.1186/s12916-017-0962-x

79. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

80. Dodd KW, Guenther PM, Freedman LS, Subar AMYF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: A review of the theory. J Am Diet Assoc. 2006;106(10):1640–50. doi:10.1016/j.jada.2006.07.011

81. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45. doi:10.3945/jn.115.219634

82. Kolonel LN, Henderson BE, Hankin JH, Nomura AMY, Wilkens LR, Pike MC, et al. A multiethnic cohort in Hawaii and Los Angeles: Baseline characteristics. Am J Epidemiol. 2000;151(4):346–57. doi:10.1093/oxfordjournals.aje.a010213

83. Chacko George G, Milani TJ, Hanss-Nuss H, Kim M, Freeland-Graves JH. Development and validation of a semi-quantitative food frequency questionnaire for young adult women in the southwestern United States. Nutr Res [Internet]. 2004 Jan;24(1):29–43. doi:10.1016/j.nutres.2003.09.006

Supplementary Table 5.1. List of the 55 food groups

	Food Group	Food Items Included
1	Full fat dairy milk	Dark blue, purple, silver top milk, lactose free regular fat milk
2	Low fat dairy milk	Lite and trim milk (green or light blue), lactose free reduced fat milk
3	Milk alternatives, including sweetened	Soy, almond, rice, oat, coconut varieties, milk alternative based drinks from cafés
4	Sweetened dairy milk products	Flavoured milk, fermented or evaporated milk, breakfast drinks (e.g., Up & Go), Yakult fermented milk drink, hot chocolates, milk-based smoothies, milk-based drinks from cafés, coffee sachets, coffees made with syrup and cream (e.g., caramel macchiato)
5	Dairy yoghurt	All types of cows milk yoghurt. (Note: Soy yoghurt under soy products, coconut yoghurt under coconut fats)
6	High fat cheese	Cream cheese, goat cheese, haloumi, parmesan, cheddar, processed cheese, blue vein, mascarpone
7	Low fat cheese	Brie, bocconcini, edam, feta, mozzarella, camembert, cottage, ricotta, paneer
8	Apple, banana, orange	Apple, banana orange
9	Other fruit	All other fruit (fresh, canned, dried)
10	Tomatoes	Fresh, canned, cooked tomatoes
11	Dark yellow vegetables	Carrots, pumpkin, butternut squash
12	Green vegetables	Lettuce, spinach, cabbage, broccoli, watercress, green beans, brussel sprouts, courgette
13	Other non-starchy vegetables	Capsicum, onion, mushrooms, frozen mixed vegetables, beetroot, squash
14	Potatoes / potato dishes (excluding chips)	Potato (boiled, mashed, baked, salad, scalloped, roasted)
15	Starchy vegetables	Kumara, yam, parsnip, turnip, swedes (boiled, mashed, baked) Taro (flesh, roots, stalks), green banana, sweet corn kernels, breadfruit, cassava, green banana
16	White breads	Plain white bread, wraps, focaccia, bagels, pita bread, rēwena bread, doughboys, breadcrumbs, including gluten free options, naan (plain)
17	Discretionary breads	Crumpets, scone, savoury muffin, plain croissant, pancakes, waffles, iced bun, savoury pin wheels, garlic bread, fruit bread, roti, naan (garlic)
18	Crackers	All crackers made from grains, cream crackers, Cruskits, rice crackers
19	Whole grain products	Wholegrain breads (High fibre, wholemeal, wholegrain, including gluten free options), grains (Quinoa, buckwheat, bulgur wheat, brown rice, wholemeal pasta, wholegrain gluten free pasta e.g., brown rice)
20	Refined grains	White rice, white pasta, noodles (instant, egg, rice), canned spaghetti, cous cous, including gluten free pasta
21	Refined grain mixed dishes	Macaroni and cheese, carbonara, white rice salad, two-minute noodles
22	Oats	Porridge, rolled oats, oat bran
23	Sweetened cereal	Sultana bran, light and fruity cereal, chocolate-based cereals, Nutri- Grain, honey puffs, milo cereals, fruit loops, oat sachets, all muesli and granola
24	Unsweetened cereals	Weet-Bix, bran cereals, rice bubbles, cornflakes
25	Red meats	Beef, lamb, venison, mince, patties, (including red meat mixed dishes: Stir fry, curry, stew)
26	White meats	Chicken, pork, turkey, (including white meat mixed dishes: Stir fry, curry, stew, casserole)
27	Processed meats	Corned beef (canned), corned silverside, smoked chicken, smoked hock and salami, ham, sausages, frankfurters, bacon, chorizo, luncheon meat

28	Fish and seafood	Canned and fresh (including mixed dishes Oka ika mata, curry, stew) and processed fish products (e.g., fish balls)
29	Eggs	Whole eggs (boiled, poached, fried, plain omelette) including egg mixed dishes (Quiche, frittata, omelette with filling, egg and banana pancake)
30	Legumes and meat alternatives	Baked beans, black beans, dahl, canned or dried legumes, hummus, and legume based vegetarian meals and products (including meat alternatives and soy products edamame beans tofu, tempeh)
31	Peanut butter and peanuts	Peanut butter and peanuts
32	Nuts and seeds	Brazil nuts, walnuts, almond, cashew, pistachio, chia, linseed, pumpkin, sesame
33	Animal fats	Cream, sour cream, reduced cream, butter, lard, dripping, ghee
34	Coconut fats and products	Coconut oil, cream, milk, desiccated, yoghurt, fresh
35	Plant based fats	Avocado (whole fruit), canola, sunflower, olive, vegetable oil, cooking spray, oil-based salad dressings (French/Italian)
36	Margarine	All margarines
37	Creamy dressings and sauces	Creamy readymade meal-based sauces, dips, mayonnaise, aioli, tartare sauce, white sauce, cheese sauce
38	Savoury sauces and condiments	Curry pastes, herb and spices, vinegar, gravy and garlic sauce, oil based condiments such as sundried tomatoes / olives in oil, pesto, tomato, barbeque, mint, soy, gravy, mustard, chutney, miso, pasta sauce, tomato paste, sweet chilli sauce (including savoury spreads Vegemite, marmite)
39	Sweet spreads	Jam, honey, marmalade, syrup (maple, golden), Nutella, chocolate peanut butter, chocolate butter
40	Cake and biscuits	Slices, cakes, loaves, muffins, biscuits, doughnuts, sweet pies, pastries, tarts
41	Puddings and other desserts	Including all milk alternative ice creams, ice cream, custard, milk-based puddings (e.g., rice, instant, semolina, pavlova, sticky date, fruit pies and crumbles, jelly, ice blocks)
42	Sweet snack foods	Fruit and nut mixes, bliss balls, chocolate, lollies, muesli bars
43	Savoury snack foods	Popcorn, potato crisps, corn chips, Twisties, bhuja mix
44	Crumbed and deep fried	Hot chips/fries, hash browns, and packaged home baked chips, wontons, paraoa (fry) bread, schnitzel, nuggets, crumbed fish
45	Fast-food (burgers)	Pies, dumpling, burgers, pizzas, curries, noodle-based dishes, Nandos chicken, egg fu yong
46	Fast food (salads & sushi)	Salads, sandwiches, wraps, sushi, vegetable-based stir fry
47	Fruit and vegetable juice	Fruit and/or vegetable juice, fruit and/or vegetable smoothies
48	Soft drinks and other sugar sweetened beverages	All cordials, flavoured water, sports drinks, soft drinks, fruit drinks, iced tea, energy drinks
49	Diet drinks	All exclusively artificially sweetened beverages
50	Tea	Black, green, herbal, chai, kombucha
51	Coffee	Instant, brewed, espresso, pre-mixed sachet, filter, cold brew
52	Alcoholic beverages	Wine (standard and low alcohol), beer (standard and low alcohol), cider, spirits, RTDs, sherry, port, liqueurs, sake
53	Water	Water (unflavoured, soda, tap)
54	Sugar added to food and drink	All sugar added to food and drink
55	Soups and stock	All soups (instant, canned, packet) and stocks

Supplementary Table 5.2. Correlation of habitual food group intake with energy (kJ) adjusted dietary pattern scores

Food group	Median [25th, 75th]+ n=287	Pattern 1	P value	Pattern 2	P value	Pattern 3	P value	Pattern 4	P value
Full fat milk	19.4 [0.0, 60.8]	-0.023	p=0.698	-0.009	p=0.884	-0.172	p=0.003	0.273	p<0.001
Low fat milk	3.8 [0.0, 23.1]	0.092	p=0.119	0.314	p<0.001	0.209	p<0.001	-0.140	p=0.018
Milk alternatives	0.0 [0.0, 26.5]	0.116	p=0.049	-0.062	p=0.294	0.386	p<0.001	-0.299	p<0.001
Sweetened milk	4.6 [3.9, 12.3]	-0.185	p=0.002	0.456	p<0.001	-0.142	p=0.016	-0.191	p=0.001
Dairy yoghurt	3.9 [2.6, 8.9]	0.498	p<0.001	0.062	p=0.294	0.117	p=0.047	0.055	p=0.356
High fat cheese	5.2 [3.5, 9.0]	0.662	p<0.001	0.112	p=0.058	-0.201	p=0.001	0.099	p=0.093
Low fat cheese	4.7 [3.2, 7.9]	0.691	p<0.001	0.124	p=0.036	0.016	p=0.782	0.029	p=0.621
Apple, banana & oranges	29.6 [14.8, 53.2]	0.187	p=0.001	-0.095	p=0.108	0.491	p<0.001	-0.091	p=0.122
Other fruit	32.6 [21.2, 50.2]	0.272	p<0.001	-0.037	p=0.532	0.390	p<0.001	-0.073	p=0.218
Tomatoes	29.2 [20.0, 34.1]	0.775	p<0.001	0.004	p=0.952	0.104	p=0.079	0.049	p=0.405
Yellow vegetables	14.5 [9.8, 23.7]	0.596	p<0.001	0.004	p=0.948	0.341	p<0.001	-0.004	p=0.950
Green vegetables	26.2 [13.7, 45.9]	0.737	p<0.001	-0.130	p=0.027	0.309	p<0.001	0.108	p=0.067
Other non-starchy vegetables	20.0 [12.1, 31.1]	0.795	p<0.001	-0.080	p=0.176	0.217	p<0.001	0.070	p=0.239
Potatoes	37.1 [32.9, 42.9]	0.084	p=0.156	0.288	p<0.001	0.032	p=0.593	0.092	p=0.118
Starchy vegetables	19.3 [15.5, 27.2]	-0.227	p<0.001	-0.107	p=0.72	0.374	p<0.001	-0.090	p=0.128
White bread	8.7 [4.9, 15.9]	0.068	p=0.249	0.340	p<0.001	-0.220	p<0.001	0.021	p=0.720
Discretionary breads	4.8 [3.8, 7.3]	0.458	p<0.001	0.302	p<0.001	-0.032	p=0.584	-0.036	p=0.539
Crackers	1.5 [1.4, 2.7]	0.066	p=0.265	0.129	p=0.028	0.207	p<0.001	0.025	p=0.671
Wholegrain products	39.4 [29.6, 55.7]	0.219	p<0.001	0.125	p=0.034	0.285	p<0.001	0.030	p=0.610
Refined grain	14.9 [9.6, 25.0]	-0.128	p=0.030	0.084	p=0.156	0.146	p=0.013	0.041	p=0.492
Refined grains mixed	12.2 [0.0, 17.0]	-0.507	p<0.001	0.216	p<0.001	-0.165	p=0.005	0.072	p=0.222
Oats	0.8 [0.7, 0.9]	0.146	p=0.013	-0.049	P=0.413	0.246	p<0.001	-0.093	p=0.115
Sweetened cereals	9.1 [8.3, 11.7]	-0.075	p=0.206	0.020	p=0.736	0.211	p<0.001	-0.324	p<0.001
Unsweetened cereals	2.6 [0.0, 3.2]	-0.568	p<0.001	0.113	p=0.056	0.119	p=0.044	0.018	p=0.757
Red meat	19.0 [13.5, 24.4]	0.432	p<0.001	0.032	p=0.586	0.049	p=0.411	0.445	p<0.001
White meat	37.3 [27.5, 49.3]	0.011	p=0.857	-0.094	p=0.111	0.152	p=0.010	0.390	p<0.001
Processed meat	10.3 [7.2, 15.2]	0.205	p<0.001	0.120	p=0.042	-0.104	p=0.077	0.541	p<0.001
Fish and seafood	9.8 [7.2, 14.0]	-0.383	p<0.001	-0.128	p=0.030	0.318	p<0.001	0.113	p=0.057
Egg and egg products	24.1 [17.3, 33.8]	0.620	p<0.001	-0.110	p=0.063	0.067	p=0.257	0.156	p=0.008
Legumes and meat alternatives	8.3 [3.8, 15.4]	0.679	p<0.001	0.040	p=0.505	0.204	p<0.001	-0.242	p<0.001
Peanuts	0.9 [0.8, 1.9]	0.451	p<0.001	-0.074	p=0.214	0.339	p<0.001	-0.043	p=0.468

Food group	Median [25th, 75th]+ n=287	Pattern 1	P value	Pattern 2	P value	Pattern 3	P value	Pattern 4	P value
Nuts and seeds	1.4 [1.2, 4.2]	0.629	p<0.001	-0.168	p=0.004	0.365	p<0.001	-0.050	p=0.397
Animal fats	4.3 [2.8, 7.0]	0.248	p<0.001	0.042	p=0.478	-0.065	p=0.269	0.474	p<0.001
Coconut products	2.7 [2.3, 3.7]	-0.038	p=0.518	-0.139	p=0.019	0.377	p<0.001	0.058	p=0.332
Plant based fats	2.1 [1.0, 5.3]	0.583	p<0.001	-0.143	p=0.015	0.242	p<0.001	0.048	p=0.422
Margarine	1.7 [0.0, 2.4]	-0.294	p<0.001	0.341	p<0.001	0.100	p=0.092	-0.169	p=0.004
Creamy based sauces and dressings	1.9 [1.4, 3.3]	0.355	p<0.001	0.430	p<0.001	-0.053	p=0.374	0.059	p=0.318
Savoury sauces and condiments	14.6 [9.1, 21.7]	0.640	p<0.001	0.123	p=0.037	0.081	p=0.172	0.108	p=0.068
Sweet spreads	1.1 [0.9, 1.9]	0.036	p=0.540	0.236	p<0.001	0.109	p=0.066	-0.082	p=0.168
Cake and biscuits	23.8 [18.8, 32.2]	-0.146	p=0.013	0.369	p<0.001	0.029	p=0.620	-0.099	p=0.096
Puddings and desserts	9.4 [7.4, 15.4]	0.011	p=0.848	0.347	p<0.001	-0.060	p=0.308	-0.049	p=0.404
Sweet snacks	11.6 [6.5, 19.8]	0.363	p<0.001	0.151	p=0.010	-0.111	p=0.059	-0.119	p=0.045
Savoury snacks	2.6 [1.8, 6.0]	-0.136	p=0.021	0.258	p<0.001	-0.168	p=0.004	-0.047	p=0.429
Crumbed and deep fried	15.4 [10.9, 23.5]	-0.476	p<0.001	0.093	p=0.117	-0.243	p<0.001	-0.193	p=0.001
Fast food burgers	72.7 [41.1, 118.1]	-0.679	p<0.001	0.139	p=0.018	-0.205	p<0.001	-0.060	p=0.310
Fast food salad and sushi	9.8 [8.7, 14.6]	-0.164	p=0.005	0.321	p<0.001	-0.068	p=0.253	0.042	p=0.478
Fruit and vegetable juice	4.4 [3.0, 9.5]	-0.229	p<0.001	0.080	p=0.177	0.052	p=0.379	-0.092	p=0.122
Sugar sweetened beverages	232.7 [149.5, 304.7]	-0.718	p<0.001	0.163	p=0.006	-0.114	p=0.054	-0.128	p=0.030
Diet drinks	7.5 [0.0, 9.7]	-0.146	p=0.013	0.273	p<0.001	-0.140	p=0.017	0.052	p=0.383
Tea	27.7 [12.8, 169.1]	0.392	p<0.001	-0.128	p=0.030	0.300	p<0.001	-0.042	p=0.479
Coffee	2.2 [1.3, 15.3]	0.511	p<0.001	0.027	p=0.652	0.141	p=0.017	0.182	p=0.182
Alcoholic beverages	18.0 [6.3, 33.2]	0.521	p<0.001	0.001	p=0.984	-0.059	p=0.316	0.160	p=0.006
Water	766.1 [438.3, 1295.9]	0.241	p<0.001	-0.141	p=0.017	0.128	p=0.031	0.166	p=0.005
Added sugar to food and beverages	0.5 [0.0, 1.8]	-0.347	p<0.001	0.155	p=0.009	-0.072	p=0.226	0.094	p=0.114
Soups and stocks	27.6 [0.0, 40.7]	0.191	p=0.001	0.149	p=0.012	0.324	p<0.001	-0.014	p=0.817

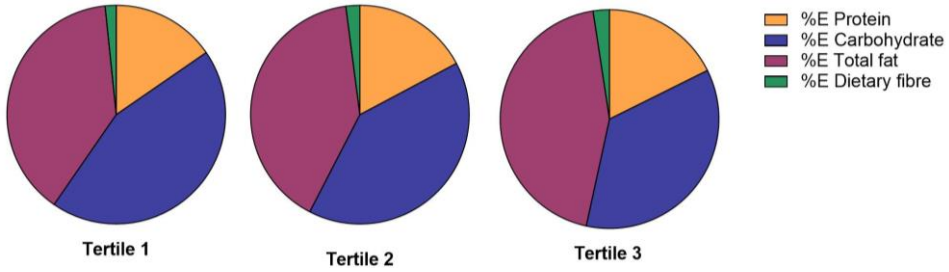
+All values are reported as medians [25th, 75th percentiles]. Daily habitual food groups intake estimated with the NCI method for a period of one month (36). Total *n* = 287 in each pattern; Pacific *n* = 126 and NZE *n* = 161. Pattern 1: "colourful vegetable, plant protein, and dairy", Pattern 2: "sweet and fat rich carbohydrate", Pattern 3: "fruit, starchy vegetables, and nuts", Pattern 4: "animal meat and fat". A Spearman's rho correlation was run to determine the relationship between an individual's energy (kJ) adjusted dietary pattern factor score and food group intake. All food groups in g/day.

Supplementary Table 5.3. Correlation of habitual nutrient intake with energy (kJ) adjusted dietary pattern scores

Nutrient	Median [25th, 75th]+ n=287	Pattern 1	P value	Pattern 2	P value	Pattern 3	P value	Pattern 4	P value
Weight (g/d)	2493.7 [2042.1, 3135.7]	0.416	p<0.001	-0.077	p=0.192	0.189	p=0.001	0.124	p=0.036
Protein (g/d)	82.4 [73.2, 91.6]	0.303	p<0.001	0.056	p=0.347	0.051	p=0.392	0.342	p<0.001
Total fat (g/d)	89.3 [75.7, 102.3]	0.281	p<0.001	0.047	p=0.424	-0.022	p=0.716	0.112	p=0.059
SFA (g/d)	33.7 [28.7, 39.9]	0.218	p<0.001	0.173	p=0.003	-0.182	p=0.002	0.188	p=0.001
PUFA (g/d)	12.3 [9.9, 14.3]	0.373	p<0.001	-0.040	p=0.495	0.206	p<0.001	-0.050	p=0.398
MUFA (g/d)	33.7 [28.3, 38.6]	0.252	p<0.001	-0.028	p=0.639	0.026	p=0.657	0.101	p=0.088
Cholesterol (g/d)	294.6 [241.9, 363.2]	0.381	p<0.001	-0.056	p=0.342	-0.019	p=0.745	0.433	p<0.001
CHO (g/d)	194.3 [163.0, 227.0]	-0.350	p<0.001	0.431	p<0.001	-0.057	p=0.334	-0.184	p=0.002
Sugar (g/d)	80.7 [67.4, 95.3]	-0.110	p=0.062	0.377	p<0.001	0.019	p=0.746	-0.231	p<0.001
Starch (g/d)	110.2 [91.5, 132.2]	-0.401	p<0.001	0.370	p<0.001	-0.065	p=0.276	-0.140	p=0.017
Dietary Fibre (g/d)	20.6 [17.5, 24.9]	0.520	p<0.001	0.055	p=0.350	0.355	p<0.001	-0.170	p=0.004
Thiamin (mg/d)	1.2 [1.0, 1.5]	0.120	p=0.042	0.242	p<0.001	0.155	p=0.009	-0.008	p=0.899
Riboflavin (mg/d)	1.8 [1.5, 2.1]	0.430	p<0.001	0.173	p=0.003	0.036	p=0.543	0.092	p=0.122
Niacin (mg/d)	18.4 [16.0, 21.4]	0.025	p=0.675	-0.047	p=0.431	0.142	p=0.016	0.213	p<0.001
Niacin eq (mg/d)	34.2 [30.9, 39.0]	0.139	p=0.019	-0.030	p=0.611	0.107	p=0.070	0.300	p<0.001
Vitamin C (mg/d)	68.2 [50.6, 88.3]	0.460	p<0.001	-0.081	p=0.171	0.355	p<0.001	-0.076	p=0.198
Vitamin E (mg/d)	9.3 [7.4, 10.9]	0.502	p<0.001	0.007	p=0.900	0.268	p<0.001	-0.079	p=0.181
Vitamin B6 (mg/d)	2.2 [1.8, 2.7]	0.166	p=0.005	0.043	p=0.469	0.270	p<0.001	0.065	p=0.276
Vitamin B12 (µg/d)	3.6 [3.1, 4.2]	0.097	p=0.100	0.079	p=0.182	-0.092	p=0.120	0.343	p<0.001
Total Folate (µg/d)	309.5 [244.7, 373.6]	0.583	p<0.001	0.072	p=0.224	0.255	p<0.001	-0.073	p=0.218
Total Vitamin A (µg/d)	731.4 [550.5, 905.8]	0.776	p<0.001	0.038	p=0.525	0.171	p=0.004	0.086	p=0.146
Retinol (µg/d)	318.1 [239.5, 389.1]	0.543	p<0.001	0.204	p=0.001	-0.132	p=0.025	0.274	p<0.001
Beta carotene (µg/d)	2263.5 [1413.6, 3231.5]	0.759	p<0.001	-0.025	p=0.676	0.298	p<0.001	0.012	p=0.845
Sodium (mg/d)	2660.7 [2331.6, 3084.2]	-0.044	p=0.455	0.353	p<0.001	-0.165	p=0.005	0.154	p=0.009
Potassium (mg/d)	2860.6 [2591.7, 3092.6]	0.780	p<0.001	-0.074	p=0.212	0.139	p=0.018	-0.022	p=0.715
Magnesium (mg/d)	297.3 [256.3, 355.8]	0.581	p<0.001	-0.005	p=0.928	0.301	p<0.001	-0.053	p=0.370
Calcium (mg/d)	773.4 [623.6, 902.7]	0.641	p<0.001	0.221	p<0.001	0.030	p=0.608	-0.072	p=0.226
Phosphorous (mg/d)	1348.1 [1226.8, 1450.7]	0.774	p<0.001	0.031	p=0.598	0.033	p=0.579	0.001	p=0.992
Iron (mg/d)	11.4 [9.9, 13.1]	0.251	p<0.001	0.053	p=0.368	0.197	p=0.001	-0.023	p=0.701

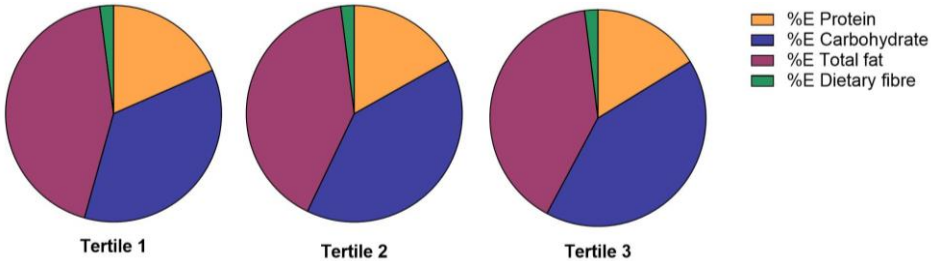
Nutrient	Median [25th, 75th]+ n=287	Pattern 1	P value	Pattern 2	P value	Pattern 3	P value	Pattern 4	P value
Zinc (mg/d)	10.0 [9.0, 11.1]	0.291	p<0.001	0.052	p=0.378	0.022	p=0.715	0.289	p<0.001
Selenium (µg/d)	61.2 [52.4, 71.2]	0.310	p<0.001	-0.091	p=0.124	0.124	p=0.035	0.208	p<0.001
Iodine (µg/d)	100.1 [81.0, 120.1]	0.486	p<0.001	0.178	p=0.003	-0.008	p=0.887	-0.034	p=0.562
Caffeine (mg/d)	82.2 [31.3, 179.4]	0.506	p<0.001	0.028	p=0.632	0.039	p=0.511	0.142	p=0.016
Alcohol (g/d)	0.3 [0.1, 0.9]	0.558	p<0.001	-0.005	p=0.927	-0.018	p=0.761	0.177	p=0.003
Water (g/d)	2082.5 [1620.0, 2658.2]	0.428	p<0.001	-0.114	p=0.054	0.194	p=0.001	0.133	p=0.024

+All values are reported as medians [25th, 75th percentiles]. Daily habitual nutrient intake estimated with the NCI method (36) for a period of one month. Total *n* = 287 in each pattern; Pacific *n* = 126 and NZE *n* = 161. Pattern 1: "colourful vegetable, plant protein, and dairy", Pattern 2: "sweet and fat rich carbohydrate", Pattern 3: "fruit, starchy vegetables, and nuts", Pattern 4: "animal meat and fat". A Spearman's rho correlation was run to determine the relationship between an individual's energy (kJ) adjusted dietary pattern factor score and nutrient intake. SFA = Saturated fat, PUFA = Polyunsaturated fat, MUFA = Monounsaturated fat, CHO = Total carbohydrate (including starch and total sugar), Total Vitamin A: Total Vitamin A equivalents.



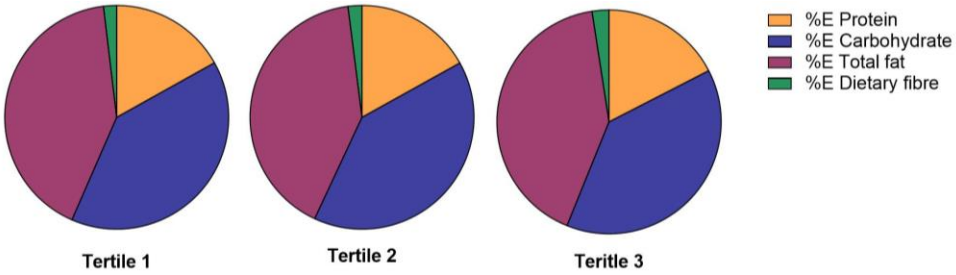
Supplementary Figure 5.1. Tertiles of adherence to "colourful vegetable, plant protein, and dairy" dietary pattern and the percentage of energy consumed from macronutrients

Percentage of energy (%E) consumed from macronutrients (41,42). %E from alcohol not included. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. Tertile 1; *n* = 95, Tertile 2; *n* = 96, Tertile 3; *n* = 96. Tertile 1= negatively adhering to the dietary pattern, Tertile 3 = positively adhering to the dietary pattern



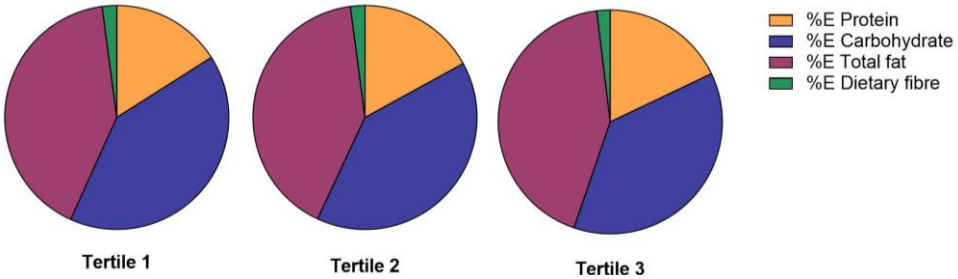
Supplementary Figure 5.2. Tertiles of adherence to "sweet and fat rich carbohydrate" dietary pattern and the percentage of energy consumed from macronutrients

Percentage of energy (%E) consumed from macronutrients (41,42). %E from alcohol not included. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. Tertile 1; *n* = 95, Tertile 2; *n* = 96, Tertile 3; *n* = 96. Tertile 1= negatively adhering to the dietary pattern, Tertile 3 = positively adhering to the dietary pattern



Supplementary Figure 5.3. Tertiles of adherence to "fruit, starchy vegetable, and nuts" dietary pattern and the percentage of energy consumed from macronutrients

Percentage of energy (%E) consumed from macronutrients (41,42). %E from alcohol not included. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. Tertile 1; *n* = 95, Tertile 2; *n* = 96, Tertile 3; *n* = 96. Tertile 1= negatively adhering to the dietary pattern, Tertile 3 = positively adhering to the dietary pattern



Supplementary Figure 5.4. Tertiles of adherence to "animal meat and fat" dietary pattern and the percentage of energy consumed from macronutrients

Percentage of energy (%E) consumed from macronutrients (41,42). %E from alcohol not included. Total *n* = 287 in each factor; Pacific *n* = 126 and NZE *n* = 161. Tertile 1; *n* = 95, Tertile 2; *n* = 96, Tertile 3; *n* = 96. Tertile 1= negatively adhering to the dietary pattern, Tertile 3 = positively adhering to the dietary pattern

6. Chapter 6. Habitual dietary intake is associated with bacterial enterotypes characterising the faecal microbiotas of lean and obese New Zealand European and Pacific women

6.1. Abstract

Introduction: Diet is considered a key driver of the obesity epidemic and habitual diet can influence the composition of the gut microbiota. Therefore, the microbiota may play a role in this multifaceted disease. The aim of this study was to explore the characteristics of microbiota composition, in relation to habitual diet, body fat content, and metabolic markers, in a population of Pacific and New Zealand European (NZE) women.

Methods: Pacific (n=125) and NZE (n=161) women aged 18-45 years were recruited based on body mass index (BMI, lean versus obese) and stratified as low (<35 %) or high (≥35 %) body fat percentage (BF%). Dietary intake was assessed with a semi-quantitative food frequency questionnaire and a 5-day estimated food record. Habitual diet was estimated with the National Cancer Institute (NCI) method. Principal component analysis was performed to derive habitual dietary patterns. BF% was assessed by dual-energy x-ray absorptiometry (DXA). Fasting blood samples were analysed for metabolic markers (e.g., fasting glucose, insulin and lipid profile). The DNA was extracted from faecal samples and the metagenomic sequences associated with the microbiota were analysed using MetaPhlAN and QIIME2 software. Enterotypes were predicted in R and the species that defined enterotypes were determined with STAMP software. Adjusted multivariate regression analysis was conducted to explore the association between microbiota composition and habitual diet.

Results: Four habitual dietary patterns were identified which explained 30.9 % of the observed dietary variance. Three enterotypes were identified: enterotype 1 was characterised by Pacific and NZE women (n=146) and the abundance of *Eubacterium rectale* and *Faecalibacterium prausnitzii*. Enterotype 2 (n=70) was characterised by Pacific women, *Bifidobacterium adolescentis, Bifidobacterium bifidum, and Lactobacillus ruminis;* and by higher BF%, visceral adipose tissue (VAT), and concentrations of fasting insulin. Enterotype 3 (n=70) was predominately found in older NZE women with lower socioeconomic deprivation, and characterised by *Subdoligranulum species, Akkermansia muciniphila, Ruminococcus bromii,* and *Methanobrevibacter smithii*. Adherence to the "colourful vegetables, plant protein, and dairy" dietary pattern was positively associated with enterotypes 1 and 3, and negatively with enterotype 2.

Conclusion: Adherence to one dietary pattern was associated with all three enterotypes, and women characterised by enterotype 2 had higher adiposity and metabolic disease risks. This study highlights the association between habitual diet, microbiota composition and metabolic health. We cannot confirm whether the microbiota is a cause or consequence of metabolic health. However, these findings warrant further investigation.

6.2. Introduction

The gut microbiota (the community of bacteria residing in the bowel) is of interest because of its influences on human physiology and metabolic disease risks. Microbial metabolites can affect numerous metabolic functions, which can influence the phenotype and health of the host. Animal models have established a causal role of the microbiota in the development of diet-induced obesity; for example, germ free mice (mice with no microbiota) are protected from diet-induced obesity (1), and the obese phenotype can be transmitted to lean hosts via microbiota transplants (2,3). Signatures of an obese microbiota composition have been proposed to be a higher ratio of the phyla *Firmicutes* to *Bacteroidetes* (4,5) or reduced microbiota diversity (number and distribution of species present) (6). However, the translation of insights from animal models to humans is equivocal; the complexity and discrepancies in the scientific literature do not allow clear conclusions to be made. Despite this, higher microbiota diversity has been proposed to be a marker of health status in humans, as lower diversity has been associated with obesity (6–8), inflammatory bowel disease (IBD) (9), and T2DM (10), in comparison to healthy participants. Although diversity does not differentiate between inter-individual differences in microbiota composition, this may not be important; higher diversity may be more beneficial for the host than community structure, because a diverse microbiota is more resilient to external influences (i.e., has a greater capacity to do different things in response to perturbations).

The composition of the diet (e.g., protein, fat, and carbohydrate content/ratio) can modulate the microbiota composition and function. The microbiota are driven by the amount and type of non-digestible dietary components consumed, such as plant polysaccharides (cellulose, hemicellulose, and resistant starch) collectively defined as dietary fibre (11). These host indigestible substrates pass through the gut to the large bowel, where the microbiota (equipped with carbohydrate degrading enzymes) metabolise and ferment the complex carbohydrate structures, as well as enhancing the bioavailability of phytochemicals in plant cell walls (12–14). A product of microbial mediated carbohydrate fermentation are short chain fatty acids (SCFAs: acetate, propionate, and butyrate), which can influence metabolic functions of the host; for example, appetite regulation, including blood glucose and lipid control (13–15). The microbiota can also metabolise dietary sources of L-carnitine and choline

(e.g., eggs, beef, pork, and fish) to trimethylamine, which is further oxidised to trimethylamine *N*-oxide (TMAO) by the liver (16). Higher TMAO is associated with increased cardiovascular disease (CVD) risk (17). Thus, the composition of the diet directly influences the substrates available for the microbiota which, in turn, can influence the functions of the microbiota and thus the biological function and health of the host.

Higher diversity in the habitual diet is associated with better health outcomes for the host and appears to be positively associated with microbiota diversity (18). Based on the evidence, the quality, and quantity of food sources habitually consumed can impact health outcomes either positively or negatively (19–22). For example, dietary patterns rich in core foods such as fruits, vegetables, complex carbohydrates, lean proteins, and unsaturated fats (e.g., Mediterranean diet), reduce the risk of obesity, cardiometabolic diseases, and premature mortality (20–23). Such dietary patterns are rich in substrates that are the primary food source of the microbiota, suggesting that the beneficial effects of plant-based and Mediterranean diets might be at least partly mediated by the functions of the microbiota. In contrast, dietary patterns characterised by animal protein and fat, refined and sugar sweetened 'discretionary' foods (so called Western diet), have been associated with reduced microbial diversity (24), increased risk of obesity and diabetes (22), including inducing endotoxemia and inflammation in the gut (25).

The evidence suggests an individual's habitual diet (i.e., what they usually eat) can exert the strongest selective pressure on the composition of their microbiota (26,27), which in turn can influence the functions of the microbiota (with the potential to influence disease risk). Despite this, there is a paucity of convincing research linking habitual dietary intake to health outcomes and the microbiota in one setting. Observational studies exploring habitual diet in relation to microbiota composition tend to focus on specific hypotheses; for example, differences between participants' adherence to specific diets (e.g., Mediterranean diet (28–30), omnivores versus vegetarians/vegans (31–33)), or nutrient intakes (26,34,35), or are conducted in particular populations such as elderly (18,36,37), and convenience samples from larger cohorts (e.g., Twins UK (34), hospital based studies (35,36)). In addition, most studies are limited in their dietary assessment. For instance, assessing habitual diet only with Food Frequency Questionnaires (FFQs) (18,26,27,29,30,33,34,38,39) which, are prone to overestimate dietary intake (40–42) yet, are often favoured in larger studies to reduce

participant burden. Further, the handling of dietary data are not always accurately described (e.g., only a few studies reported accounting for misreporting (29,35) or adjusting for energy intake (28,32,34,35,37)) and adjustment for covariates (e.g., age, sex, body mass index (BMI)) which could either influence dietary intake or microbiota composition is rarely done (18,27,37–39). Moreover, most studies assess the composition of the faecal microbiota by analysis of 16S rRNA gene sequences (18,26–29,33–35,37,38). Utilising 16S rRNA gene sequencing provides confident identification to the family level, however greater analytical depth is required to explore bacterial species abundances accurately. Whole genome metagenomic shotgun sequencing enables greater analytical depth, however it is more expensive so is not widely used. Essentially it is challenging to draw clear conclusions, secondary to the heterogeneity between previous studies, and, at times, the insufficiently described dietary assessment and analyses.

Well-designed large observational human studies with in depth assessment of microbiota composition and robust dietary assessment (e.g., multiple assessment methods capturing multiple days of dietary intake) are lacking. Filling this gap will advance understanding of the diet-microbiota relationship. Exploring habitual dietary patterns will enable the assessment of the broader diet (reflecting the combinations of foods and nutrients consumed together), and thus the association of the whole diet and microbiota composition. Assessing the composition of the microbiota with metagenomic shotgun sequencing would enable greater analytical depth. Researching this association in a healthy free-living population and adjusting for factors which can influence habitual diet and microbiota composition (e.g., age, socioeconomic status, energy intake, faecal water content), will contribute to the advancement of knowledge. The aim of this part of the PhD research was to explore the characteristics of microbiota composition in relation to habitual diet (dietary patterns, foods, and nutrients), body fat content, and metabolic health markers, in a population of healthy NZ women (Pacific and NZ European) with different metabolic disease risk and body fat profiles (lean and obese).

6.3. Materials and Methods

6.3.1. Study design

The participants were part of the cross-sectional "The PRedictors linking Obesity and gut MIcrobiomE": PROMISE study, which was conducted between July 2016 and September 2017, at the Human Nutrition Research Unit at Massey University in Albany, NZ. This PhD research aimed to further explore the microbiota characteristics of these women in relation to their dietary intake. Participants were Pacific and NZE women, free from any chronic disease, aged 18-45 years (post-menarche and premenopausal) from the Auckland region. Details of the study procedures and recruitment strategies have been published elsewhere (37). Briefly, participants attended two study visits and completed at home data collection between study visits. Only details relevant to this study are presented.

The study was approved by the Southern Health Disability Ethics Committee (16/STH/32) and conducted according to the guidelines of the declaration of Helsinki. The trial was registered at anzctr.org.au (ACTRN12618000432213). All participants were informed in detail about the procedures and measurements and gave written informed consent to participate in the study.

6.3.2. Demographic and anthropometric information and blood sampling

At the first study visit, trained research staff followed standard operating procedures to conduct one-on-one interviews with participants to capture a range of demographic information (e.g., occupation, personal/household income, address). Level one International Society for the Advancement of Kinanthropometry (ISAK) accredited research staff also collected anthropometric measurements (e.g., weight, stretched height) according to ISAK protocols (44). The Quetelet index was used to calculate BMI kg/m^2 (weight (kg)/height (m^2)) (45). The New Zealand Deprivation 2013 index (NZDep2013), an area-based measure of socioeconomic deprivation, was used to assign a socioeconomic deprivation score ranging between decile one "least deprived" to decile ten "most deprived" to each participant (46); details of the NZDep2013 have been presented previously (see chapter three).

At the second study visit, body composition was assessed with a whole-body scan using Dual-energy X-ray Absorptiometry (DXA) (Hologic QDR Discovery A, Hologic Inc, Bedford, MA with APEX V. 3.2 software). Total body fat percentage (BF%), visceral fat (VAT), android fat (AF%), and gynoid fat (GF%) percentage were all assessed with DXA. A trained phlebotomist collected fasting blood samples following an overnight fasting of at least 10 hours, and blood metabolic markers (e.g., plasma insulin, glycosylated haemoglobin (HbA1c), and lipid profile) were assessed (43). Blood collection processes are described in detail in the PROMISE protocol paper (43), and processing of metabolic markers has been described previously (see chapter three). Homeostasis model assessment (HOMA-IR) index for insulin resistance was calculated (fasting blood glucose [mmol/L] x fasting blood insulin [µU/mL]/22.5) (47).

6.3.3. Dietary assessment

Between the first and second study visit participants completed a 5-day non-consecutive estimated food record (5DFR) at home. Dietary assessment methodology has previously been described in detail (see chapter three). During the second study visit, each food record was reviewed by a dietitian and with the participant, to inform accurate dietary data entry. Participants also completed a validated semi-quantitative NZ Women's Food Frequency Questionnaire (NZWFFQ) regarding the previous 30 days intake (48). The 220-item NZWFFQ was hosted by SurveyMonkey© software (SurveyMonkey Inc, San Mateo, California, USA), which allowed research staff to monitor live progress as participants completed it online.

6.3.4. Dietary data processing

Energy, macro- and micro-nutrient analysis of the 5DFR and NZWFFQ were completed using FoodWorks9 (Xyris Software (Australia) Pty Ltd, Queensland, Australia) nutrition analysis software, which hosts multiple Australian food composition databases as well as New Zealand database (FOODFiles 2016), developed by the NZ Institute for Plant & Food Research and the NZ Ministry of Health. Dietary data processing has been presented in detail previously (see chapter three). In brief, to ensure standardised and consistent data entry, trained research staff entered the 5DFR and NZWFFQ into the software following a standard operating procedure. Energy, nutrient, and food intake data were exported from FoodWorks9 in

standard units, and extensively reviewed using the raw dietary data to verify the accuracy of the processed dietary data. Cut-offs of >2100 kJ/day and <27000 kJ/day, were considered to indicate valid completion of the 5DFR and NZWFFQ; all others were excluded from further analyses. All reported foods items from the 5DFR (n>2850) and the NZWFFQ were then summarised into 55 food groups, based on similar nutritional composition and characteristics for further analysis (Supplementary Table 6.1). The process of collapsing all reported food items into the 55 food groups has been presented previously (see chapter three and five). For descriptive purposes, the 55 food groups were further classified as either "core" or "discretionary" based on NZ and Australian dietary guidelines (49,50). Core or 'healthy' foods are considered nutrient-rich foods from the four core food groups (e.g., fruit, vegetables, dairy) (49). In contrast, 'discretionary' foods do not fit into the core or healthy food groups and are, characteristically, processed, energy dense and nutrient poor foods (e.g., potato chips/crisps, confectionary) (49,50).

Total energy (reported as kilojoules (kJ)) includes the energy contribution from all the macronutrients as well as total dietary fibre. Total dietary fibre includes all non-starch polysaccharides (cellulose, hemicellulose, and resistant starches; and NZFOODFILES uses AOAC Prosky method to analyse dietary fibre). Total carbohydrate (reported in grams (g)) includes free sugars, dextrins, starch, and glycogen. Total fat (reported in grams (g)) includes all saturated and unsaturated fats.

Calculation of habitual dietary data

The National Cancer Institute (NCI; USA) method (51,52) was used to calculate individual habitual dietary intake for each participant with valid dietary data (within the specified kJ/day cut-offs), which has been presented in detail previously (see chapter three and four). In the current analysis, the 5DFR was used as the primary dietary data, and the covariates age, ethnicity, BMI (kg/m^2), season (summer, autumn, winter, spring), weekend (weekday = Monday - Thursday, weekend = Friday - Sunday), and FFQ information (in standard units/day) were considered (see chapter three and four). The average daily intake of 36 nutrients (in standard units/day) and 55 food groups (g/day) consumed within the last month was calculated for each participant.

Dietary pattern analysis

Principle component analysis (PCA) was conducted to identify dietary patterns for use in subsequent analysis (see below). The procedure has been presented in detail elsewhere (see chapter three and five). Dietary pattern scores were adjusted for total energy intake, and energy-adjusted dietary patterns were used for all subsequent analyses.

Faecal sample collection

Faecal samples were collected after completion of the food record and were stored in the participants' home freezers 11 to 14 days prior to delivery to the research unit. Subsequent storage was at -80°C until laboratory analysis. Samples will be kept for 10 years and will not be analysed for anything that participants have not provided consent for.

Faecal water content

Faecal water content was used in this study as a proxy for colonic transit time. To determine faecal water content, approximately 200mg of each faecal sample was placed in a pre-weighed microfuge tube, the weight recorded, and the tube with cap open placed in a 37°C incubator. The tubes were dried until a constant dry weight was obtained, and percentage water content was then calculated.

6.3.5. Microbiota analysis

DNA extraction and sequencing

Emeritus Professor Gerald Tannock and Dr Blair Lawley conducted the DNA extraction, analyses of the microbiota with MetaPhlan 2.0 and QIIME2, and the prediction of enterotypes (Microbiome Otago, University of Otago) as follows. DNA was extracted from 250 mg faeces according to the kit protocol provided by the manufacturer (PowerSoil DNA isolation kit, Mo Bio, Carlsbad, CA, USA), with the following modification. Faecal samples were suspended in 1mL of TN150 buffer (containing 10 mM TRIS-CL pH 8.0, 150 mM NaCl). The suspension was centrifuged at 14,600×*g* (3 min, 5°C) and then suspended in 700 µl solution from the PowerBead Tubes, from the PowerSoil DNA isolation kit. The suspension was added back to the PowerBead Tubes and the standard protocol followed. DNA was eluted in 100 µl of elution buffer (warmed to 70°C) and stored at -80°C.

Quality and quantity of genomic DNA was checked on a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and on a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) prior to sending the cleaned DNA to New Zealand Genomics Ltd. (NZGL) for shotgun metagenome sequencing. NZGL prepared 384 Thruplex DNA libraries and carried out 2×125 bp paired-end sequencing across 24 lanes on an Illumina HiSeq 2500 instrument (Illumina, San Diego, CA, USA). Libraries were sequenced across a minimum of six HiSeq lanes, and multiple libraries were prepared for several samples to test for library preparation and sequence run bias. An average of 13,150,561 (range 7,6940,894-17,081,755) reads were recovered for each sample.

For quality control BBDuk (https://sourceforge.net/projects/bbmap/) was used to trim adapters, remove low quality reads and remove reads <100 bp after trimming. KneadData (http://huttenhower.sph.harvard.edu/kneaddata) was used as quality control to remove human genome reads (i.e., DNA) from bacterial reads, implementing the hg19 database. Sequence data used in this study will be deposited with the short-read archive (SRA).

Bioinformatic analysis

Microbiota taxonomic profiles were created from DNA sequences using MetaPhlan 2.0 (version 2.6.0) according to default parameters (53).

Microbiota composition and diversity was further analysed with QIIME2 (version 2018.8, Bolyen E et al., (54) (https://qiime2.org/) using converted output tables from MetaPhlan 2.0. Beta diversity group significance for each metric (Bray-Curtis Dissimilarity index, and Jacard similarity matrix) was measured with PERMANOVA (55) and group dispersion was measured with PERMDISP (56).

Enterotypes were predicted in R using the approach described in Arumugum et al. (57) and following the tutorial provided by EMBL (http://enterotyping.embl.de). Differential abundance testing to determine which species were driving enterotypes was carried out with Statistical Analysis of Taxonomic and Functional Profiles (STAMP) (58). Each enterotype was

compared to all other samples using Welch's *t*-test using Benjamini-Hochberg for multiple testing correction.

6.3.6. Statistical analysis

Participants were selectively recruited based on BMI and stratified into high and low BF% groups using the median as the cut point, that is, low-BF% (<35 %) versus high-BF% (≥35 %) as individuals with the same BMI can have different body compositions and metabolic disease risk (59,60). The NZDep2013 index was collapsed into quintiles (e.g., quintile one = decile one and two, quintile two = decile three and four) which were used for subsequent analysis. For each participant the *Firmicutes* to *Bacteriodetes* ratio was calculated by dividing the relative abundance of the phyla *Firmicutes* by the relative abundance of *Bacteriodetes.*

To assess the association between habitual diet and microbiota composition (expressed as enterotypes and relative abundance) we measured associations with energy-adjusted dietary patterns (determined with PCA as described previously (see chapter three and five)), the 55 food groups, and 36 nutrients. Logistic regression analyses were conducted for dichotomous outcomes (e.g., relative abundance of bacteria species) and multinomial regression for (outcome) variables that were categorical (e.g., enterotype) but not dichotomous. For species that were either present or absent in the study population, "presence" versus "absence" was used as the outcome variable. Species that were relatively abundant in all participants were stratified as either "lower" or "higher" relative abundance (using the median as the cut point). Univariate analysis was conducted followed by multivariate analyses controlling for potential confounders including age, NZDep2013 quintiles, faecal water content, and energy intake. Due to the design of the study (e.g., participants were selectively recruited based on ethnicity and BMI) we also adjusted for ethnicity and BF% groups. Measurements that did not follow a normal distribution were reported as medians [25^{th}, 75^{th}] and non-parametric tests were conducted to assess differences between groups. Analysis was conducted using SAS Enterprise Guide version 7.1 (SAS institute, Cary, NC, USA). All independent variables were assessed for collinearity by assessing tolerance and the variance inflation factor (VIF). No collinearity was detected. Analyses were conducted separately for NZE and Pacific participants, as well as for both groups combined. *P*-values <0.05 were considered statistically significant.

6.4. Results

6.4.1. Characteristics of the study population

A total of 351 women were eligible and enrolled to participate in the wider PROMISE study, and 304 participants completed all aspects of the study (43). For this study 17 participants were excluded prior to calculation of the habitual dietary data because of misreporting (>27000 kJ/day) their energy intake on the NZWFFQ. One subject was excluded from analysis based on not having any DNA sequence data (Figure 6.1). We had complete data for 286 participants: 125 Pacific (44 %) and 161 NZE (56 %) women, with a median age of 23 and 32 years respectively (see Table 6.1). There were no differences in BF% between Pacific and NZE women. Pacific women were younger, had a higher BMI and VAT including lower total and high-density lipoprotein (HDL) cholesterol in comparison to NZE women. In comparison to Pacific women, NZE women had lower HbA1c, fasting plasma concentrations of insulin, and HOMA-IR.

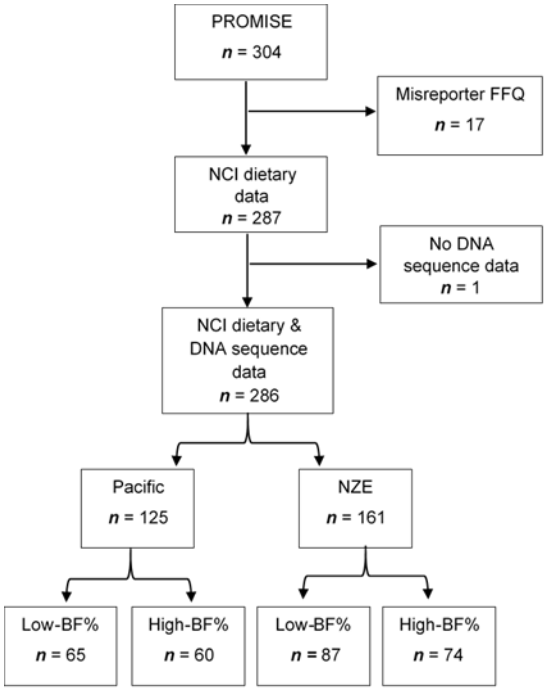


Figure 6.1. Overview of participant's included in the study

Table 6.1. Demographic, microbiota, and body composition characteristics of participants stratified by ethnicity

	Total population n= 286	**Pacific n= 125**	**NZE n=161**
Age (years)	28 [22, 35]	23 [20, 29]	32 [25, 37]*
Weight (kg)	77.6 [65.6, 96.0]	82.3 [72.3, 98.8]	70.7 [61.4, 92.5]*
Height (cm)	167.8 [163.7, 172.2]	168.6 [163.8, 174.1]	167.2 [163.7, 171.2]
BMI (kg/m^2)	28.1 [23.0, 33.4]	29.5 [24.7, 34.8]	25.0 [22.2, 32.9]*
Total body fat (%)	34.5 [28.8, 39.6]	34.6 [29.5, 39.3]	34.3 [27.8, 39.9]
Visceral fat (%)	32.3 [23.7, 38.9]	34.3 [26, 39.8]	30.0 [21.4, 38.3]*
Body fat groups n (%)			
<35 %	152 (53 %)	65 (52 %)	87 (54 %)
≥35 %	134 (47 %)	60 (48 %)	74 (46 %)
Metabolic biomarkers∞			
TC (<5 mmol/L)-	4.8 [4.3, 5.4]	4.5 [4.2, 5.1]	5.0 [4.5, 5.8]*
HDL-C (>1 mmol/L)-	1.5 [1.3, 1.9]	1.4 [1.3, 1.7]	1.7 [1.4, 2.0]*
LDL-C (0-3.4 mmol/L)-	3.0 [2.4, 3.5]	2.9 [2.4, 3.3]	3.1 [2.4, 3.6]*
TG (<2 mmol/L)-	0.9 [0.7, 1.2]	0.9 [0.7, 1.2]	0.9 [0.7, 1.2]
TC:HDL (<4 mmol/L)-	3.1 [2.6, 3.7]	3.1 [2.6, 3.7]	3.0 [2.5, 3.7]
Glucose (3.5-5.4 mmol/L)-	5.3 [5.0, 5.6]	5.3 [5.1, 5.7]	5.3 [5.0, 5.6]
Insulin (3-25 uU/mL)-	10.9 [7.3, 17.1]	14.6 [9.8, 23.2]	8.7 [6.4, 12.6]*
HOMA-IR	2.6 [1.7, 4.0]	3.4 [2.3, 5.9]	2.1 [1.5, 3.1]*
HbA1c (<40 mmol/mol)-#	31.7 [30.0, 33.6]	32.9 [31.4, 35.4]	30.7 [29.3, 32.6]*
Deprivation Index#	6 [3, 8]	8 [6, 9]	4 [2, 6]*
Microbiota characteristics			
Predicted species abundance	73 [67, 79]	75 [67, 81]	72 [67, 77]
Pielou's Evenness	0.70 [0.66, 0.73]	0.69 [0.65, 0.72]	0.71 [0.67, 0.74]*
Shannon index	4.3 [4.1, 4.6]	4.3 [3.9, 4.5]	4.3 [4.1, 4.6]
Firmicutes/Bacteroides	6.7 [3.3, 16.0]	7.8 [3.8, 19.0]	6.0 [2.9, 13.2]
Faecal water content (%)≈	69 [61, 74]	68 [60, 74]	69 [63, 73]

All values are reported as medians [25th, 75th percentiles]. Mann Whitney statistical test used to identify a significant difference ($p < 0.05$). *Statistically significant difference between ethnicities. ≈NZE women (n=1), € NZE women (n=2), ∞Pacific woman (n=1), #Pacific women (n=2) and πPacific women (n=3) have not been included in analysis due to missing data. Body fat % assessed with DXA. NZE: New Zealand European, BMI: body mass index, TC: Total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TG: Triglycerides, TC:HDL: Total cholesterol/ high density lipoprotein cholesterol, HbA1c: glycosylated haemoglobin. New Zealand Deprivation Index (NZDep2013). Firmicutes/Bacteroides ratio, Faecal water content (%): percentage of water content in the faecal sample analysed. -Normal healthy ranges for metabolic biomarkers (61).

6.4.2. Microbiota characteristics

To investigate differences in taxa comprising the participants' microbiotas, the relative abundance at the phylum, family, genus and species level was explored (Figure 6.2).

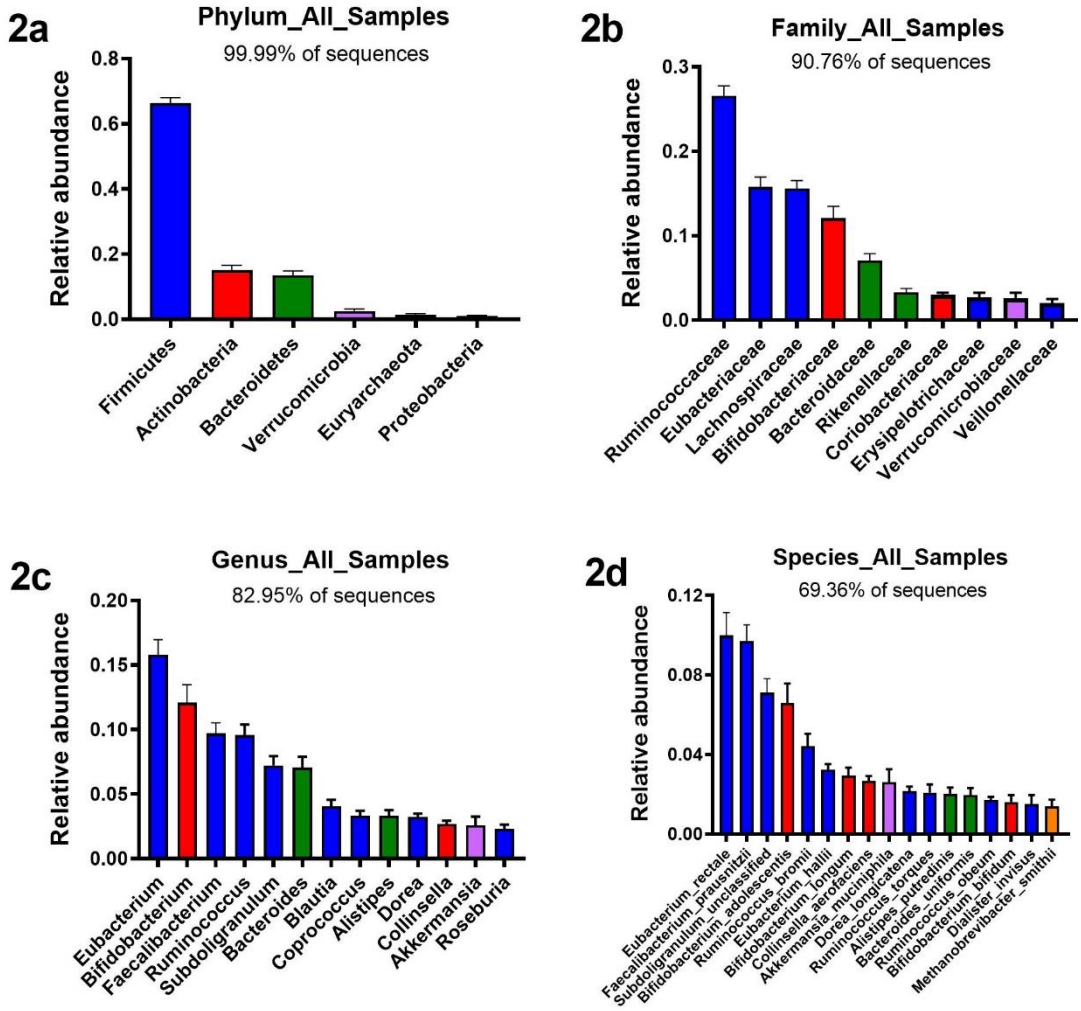


Figure 6.2. Phylogenetic characteristics of the microbiota for all participants

DNA was extracted from faeces and analysed with metagenomic shotgun sequencing. All participant's (n=286, Pacific n=125, NZ European n=161). Figure 2a. Phylum level. 2b. Family level. 2c. Genus level. 2d. Species level.

Utilising both QIIME2 (54) and STAMP (58), stratifying the population by BF% groups did not reveal any significant differences in phylogenetic characteristics (Figure 6.3). However, stratifying the population by ethnicity highlighted significant differences in phylogenetic characteristics between groups (see Figure 6.4). Thus, characterisation of the composition of the microbiota in terms of enterotypes was performed.

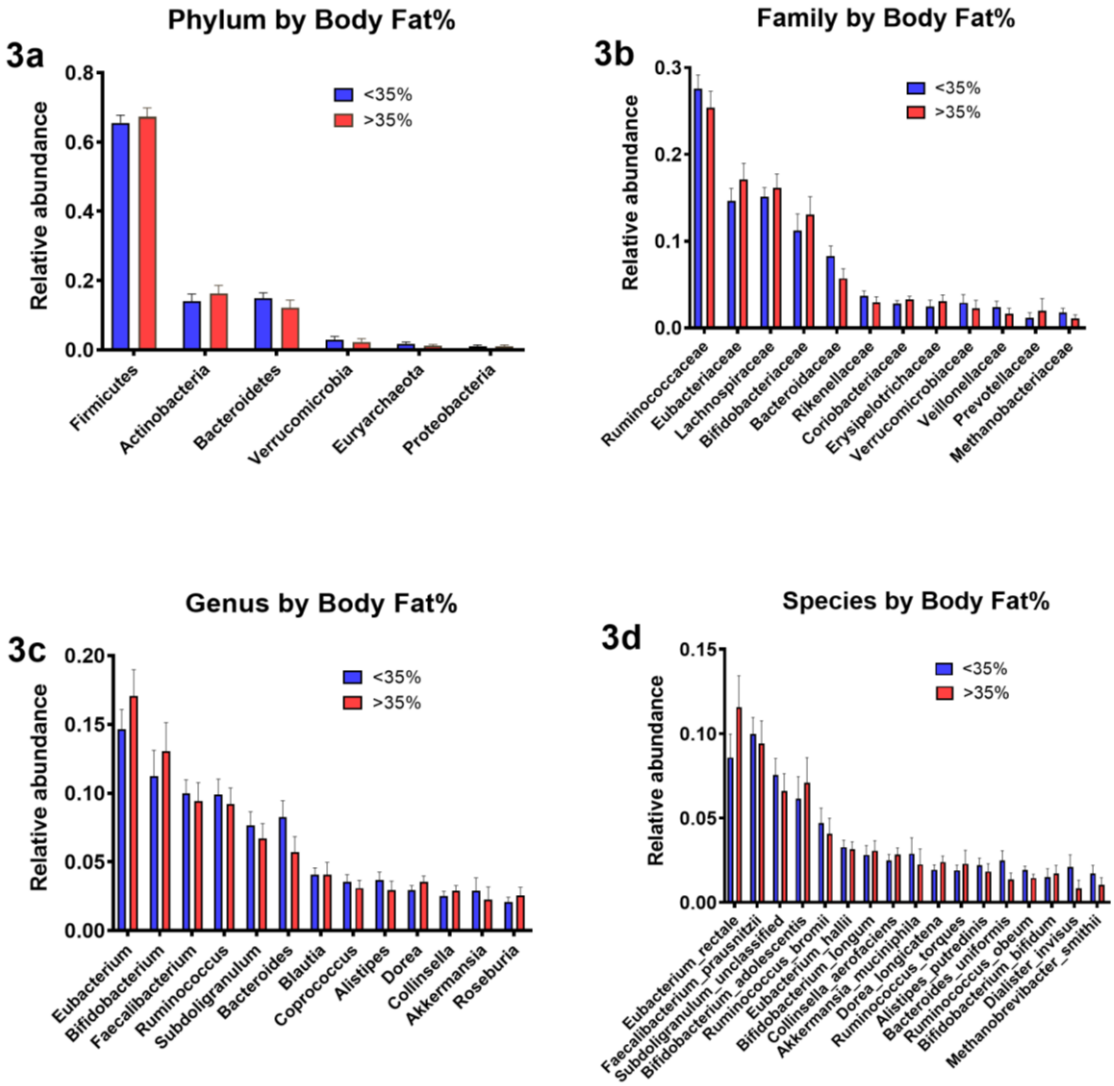


Figure 6.3. Phylogenetic characteristics of microbiota stratified by body fat groups
DNA was extracted from faeces and analysed with metagenomic shotgun sequencing for all participant's (n=286). All figures stratified by total body fat percentage (BF% assessed with DXA), BF% <35 % n=152 (blue bars), BF% ≥35 % n=134 (red bars). *indicating significant differential abundance between BF% groups following Benjamini-Hochberg adjusted Welch's t test. Figure 3a. Phylum level. 3b. Family level. 3c. Genus level. 3d. Species level

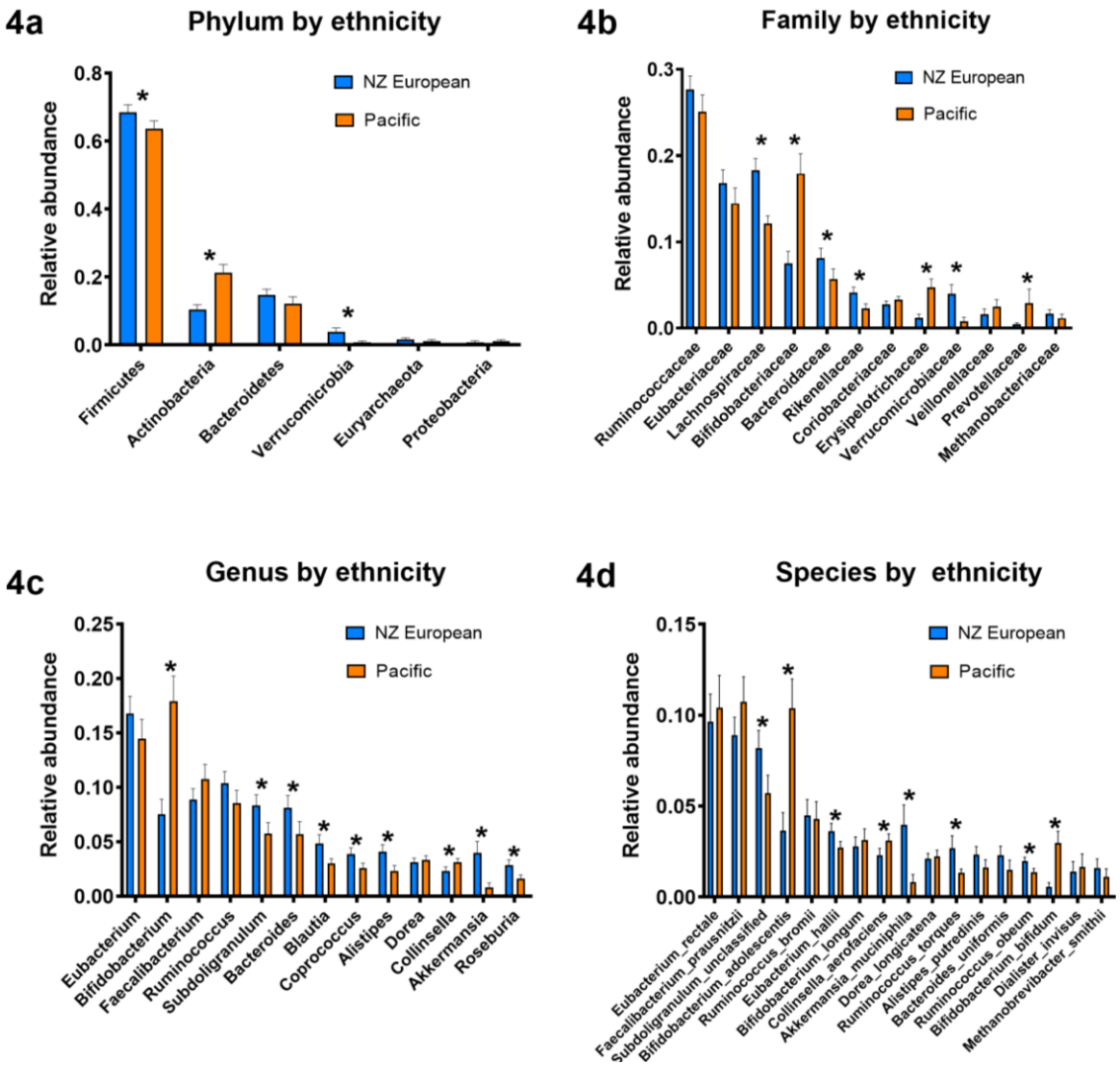


Figure 6.4. Phylogenetic characteristics of microbiota stratified by ethnicity

DNA was extracted from faeces and analysed with metagenomic shotgun sequencing for all participant's (n=286). All figures stratified by ethnicity, Pacific n=125 (orange bars), NZ European n=161 (blue bars). *indicating significant differential abundance between ethnicities following Benjamini-Hochberg adjusted Welch's t test. Figure 4a. Phylum level. 4b. Family level. 4c. Genus level. 4d. Species level

Enterotypes detected in the faecal microbiota of Pacific and NZE women

Three enterotypes were identified (Figure 6.5) and were comparable when stratified by ethnicity (data not shown). Using STAMP (58), it was determined that enterotype 1 was characterised by the abundance of butyrate-producing bacterial species, *Eubacterium rectale* and *Faecalibacterium prausnitzii* (see Figure 6.6). Enterotype 2 was characterised by the abundance of lactic acid-producing bacterial species *Bifidobacterium adolescentis, Bifidobacterium bifidum*, and *Lactobacillus ruminis* (see Figure 6.7). Enterotype 3 was characterised by the abundance of *Subdoligranulum species, Akkermansia muciniphila, Ruminococcus bromii,* and *Methanobrevibacter smithii* (see Figure 6.8). The microbiota of 146 participants was characterised by the presence of enterotype 1, which included both Pacific and NZE women. Enterotype 2 (n=70) was predominately found in Pacific women, and Enterotype 3 (n=70) predominately in NZE women (Table 6.2).

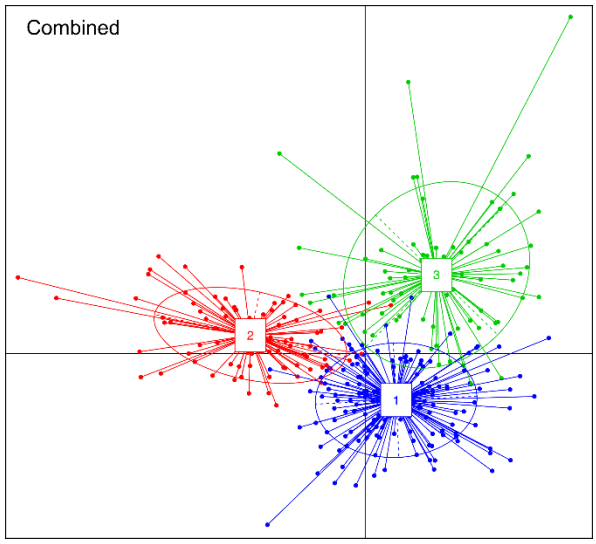


Figure 6.5. Enterotype PCoA plot

Showing clustering of samples for all participant's (n=286) in three enterotypes. Blue= enterotype 1, n=146 (NZE n=89, Pacific n=57). Red= enterotype 2, n=70 (NZE n=11, Pacific n=59), Green= enterotype 3, n=70 (NZE n=61, Pacific n=9)

Women characterised by enterotype 2 were younger and had a higher BMI and VAT, including higher fasting insulin, HbA1c concentrations and a higher HOMA-IR index, in comparison to women with enterotypes 1 and 3. Women characterised by enterotype 3 were older, had a lower deprivation index, total cholesterol: high density lipoprotein cholesterol ratio (TC:HDL), and had higher HDL cholesterol compared to enterotypes 1 and 2 (Table 6.2).

Table 6.2. Demographic, microbiota, and body composition characteristics of participants stratified by enterotypes

	Enterotype 1 n= 146	**Enterotype 2 n= 70**	**Enterotype 3 n= 70**
Ethnicity n (%)			
Pacific	57 (39 %)	59 (84 %)	9 (13 %)
NZ European	89 (61 %)	11 (16 %)	61 (87 %)
Age (years)	29 [23, 35]^	22 [20, 27]+	31 [24, 38]~
Weight (kg)	75.4 [64.5, 95.0]^	88.3 [68.5, 100.8]+	73.1 [63.9, 90.2]
Height (cm)	168.5 [163.9, 172.7]	168.4 [163.7, 172.6]	166.5 [164.0, 170.2]
BMI (kg/m^2)	26.3 [22.8, 33.1]^	31.9 [24.6, 35.9]+	25.2 [22.7, 32.4]
Total body fat (%)	34.0 [28.5, 39.3]^	35.5 [31.9, 40.1]	32.5 [27.9, 40.2]
Visceral fat (%)	30.8 [23.1, 37.9]^	35.4 [29.1, 41.3]+	29.1 [20.4, 38.3]
Body fat groups n (%)			
<35 %	83 (57 %)	29 (41 %)	40 (57 %)
≥35 %	63 (43 %)	41 (59 %)	30 (43 %)
Metabolic biomarkers∞			
TC (<5 mmol/L)-	4.8 [4.4, 5.5]^	4.6 [4.2, 5.1]	4.8 [4.2, 5.5]
HDL-C (>1 mmol/L)-	1.5 [1.3, 1.9]^	1.4 [1.2, 1.6]+	1.7 [1.4, 2.0]~
LDL-C (0-3.4 mmol/L)-	3.0 [2.5, 3.5]	2.9 [2.5, 3.4]	2.8 [2.4, 3.5]
TG (<2 mmol/L)-	0.9 [0.7, 1.1]^	1.0 [0.8, 1.4]+	0.7 [0.6, 1.2]
TC:HDL (<4 mmol/L)-	3.1 [2.7, 3.7]^	3.3 [2.7, 4.0]+	2.8 [2.4, 3.3]~
Glucose (3.5-5.4 mmol/L)-	5.3 [5.0, 5.6]^	5.4 [5.1, 5.8]	5.3 [4.9, 5.6]
Insulin (3-25 uU/mL)-	9.9 [6.9, 14.5]^	16.1 [10.7, 26.9]+	8.7 [6.4, 14.0]
HOMA-IR	2.3 [1.5, 3.4]^	3.7 [2.4, 6.3]+	2.2 [1.6, 3.4]
HbA1c (<40 mmol/mol)-#	31.6 [29.6, 33.6]^	33.4 [31.7, 35.5]+	30.7 [29.3, 31.9]~
Deprivation Index#	5 [3, 8]^	8 [6, 9]+	4 [3, 6]~
Microbiota characteristics			
Predicted species abundance	72 [64, 78]	73 [67, 78]	76 [69, 81]~
Pielou's Evenness	0.71 [0.66, 0.74]^	0.68 [0.64, 0.71]+	0.70 [0.67, 0.74]
Shannon index	4.4 [4.1, 4.6]^	4.2 [3.9, 4.4]+	4.4 [4.2, 4.6]
Firmicutes/Bacteroides	5.2 [2.6, 10.6]^	15.2 [7.0, 28.4]+	5.0 [2.8, 11.3]
Faecal water content (%)≈	69 [62, 74]	70 [61, 77]+	65 [60, 69]~

All values are reported as medians [25th, 75th percentiles]. Mann Whitney statistical test used to identify a significant difference ($p < 0.05$). ^Statistically significant difference between enterotype 1 and 2. ~Statistically significant difference between enterotype 1 and 3. +Statistically significant difference between enterotype 2 and 3. ≈NZE women (n=1), **€** NZE women (n=2), ∞PI woman (n=1), #PI women (n=2) and πPI women (n=3) have not been included in analysis due to missing data. Body fat % assessed with DXA. BMI: body mass index, TC: Total cholesterol, HDL-C: high density lipoprotein cholesterol, LDL-C: low density lipoprotein cholesterol, TG: Triglycerides, TC:HDL: Total cholesterol/ high density lipoprotein cholesterol, HbA1c: glycosylated haemoglobin. New Zealand Deprivation Index (NZDEP2013). Firmicutes/Bacteroides ratio, Faecal water content (%): percentage of water content in the faecal sample analysed. -Normal healthy ranges for metabolic biomarkers (61).

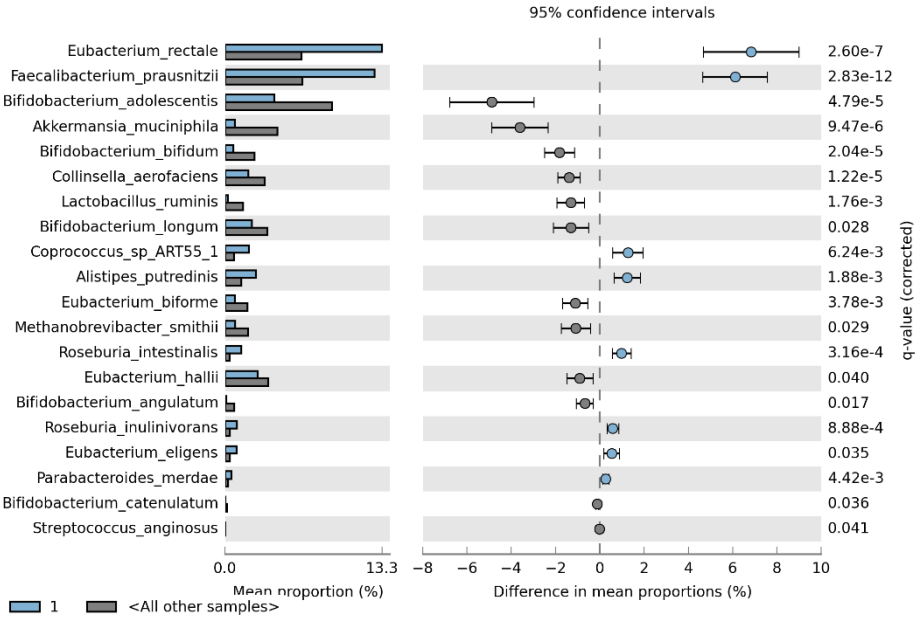


Figure 6.6. Bacterial species that define enterotype 1.

This extended error bar plot (generated in STAMP) depicts species level features within enterotype 1 (blue circles) with significant differential abundance in comparison to enterotypes 2 and 3 (grey circles), following Benjamini-Hochberg adjusted Welch's t test.

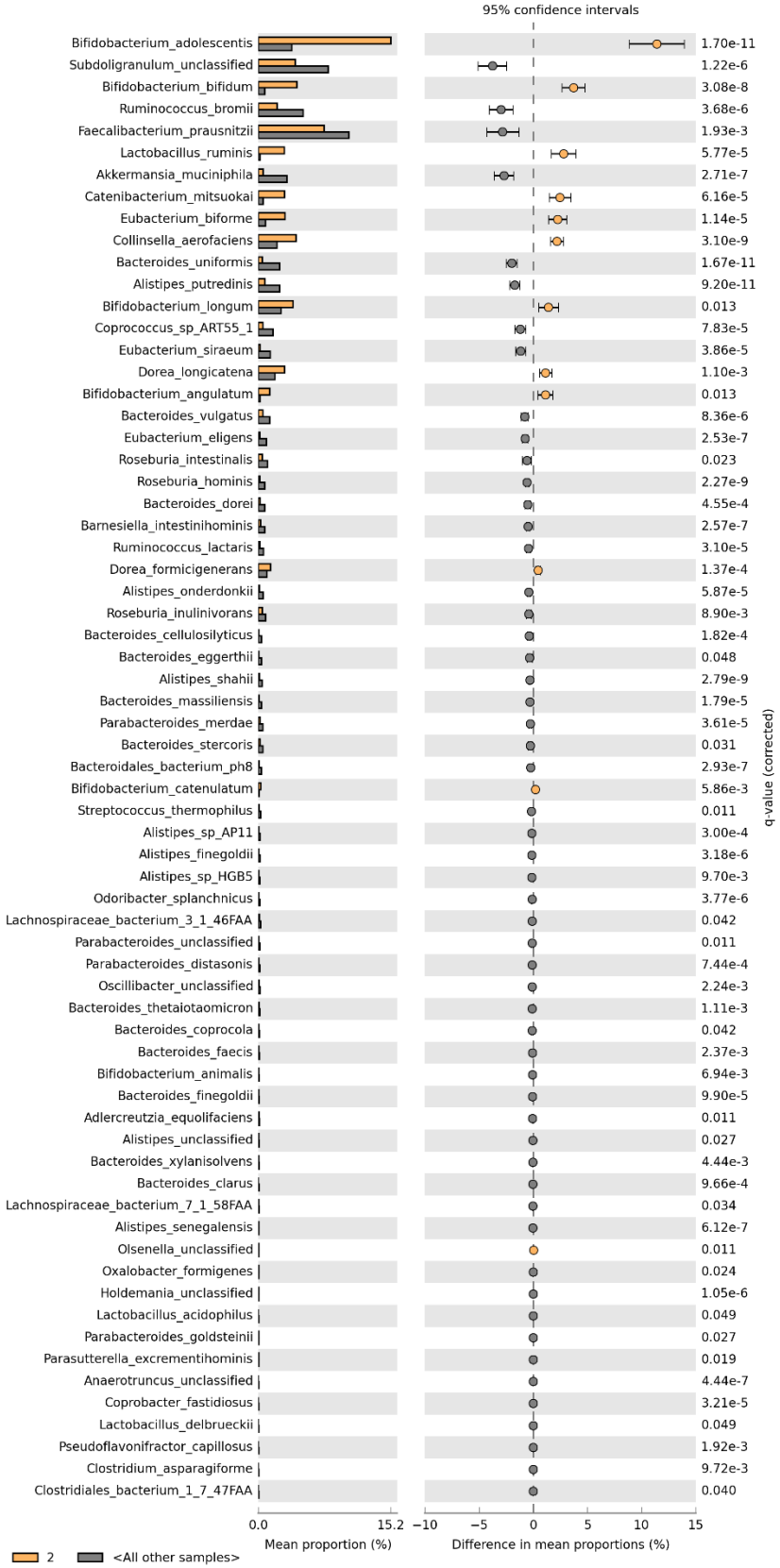


Figure 6.7. Bacterial species that define enterotype 2

This extended error bar plot (generated in STAMP) depicts species level features within enterotype 2 (orange circles) with significant differential abundance in comparison to enterotypes 1 and 3 (grey circles), following Benjamini-Hochberg adjusted Welch's t test.

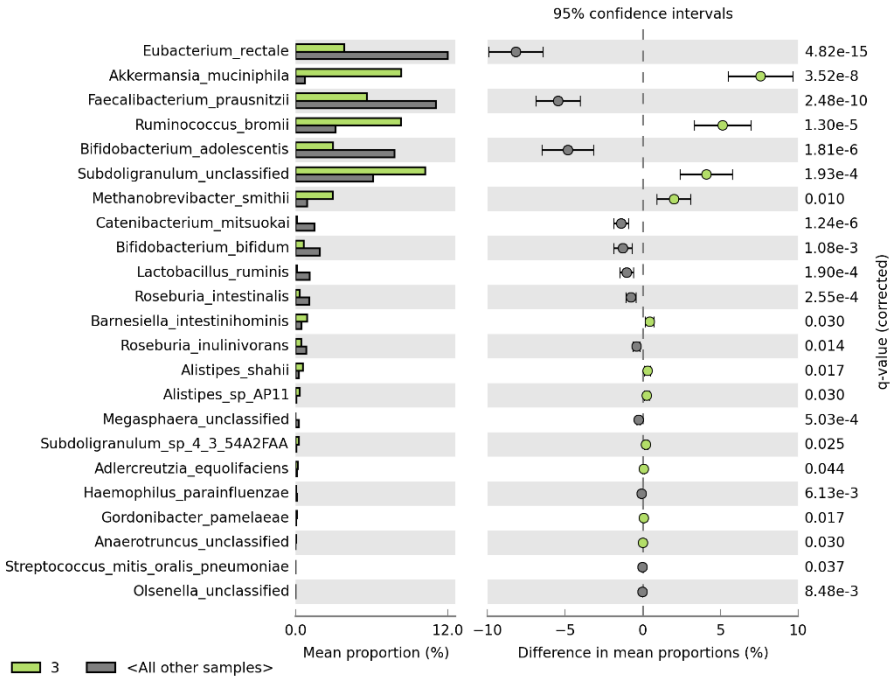


Figure 6.8. Bacterial species that define enterotype 3

This extended error bar plot (generated in STAMP) depicts species level features within enterotype 3 (green circles) with significant differential abundance in comparison to enterotypes 1 and 2 (grey circles), following Benjamini-Hochberg adjusted Welch's t test.

Alpha diversity

The average number of predicted species per participant was 73 (Table 6.1), and there were no significant differences between the predicted species abundance in Pacific and NZE women. However, the microbiota of NZE women was characterised by significantly higher evenness (distribution of species present) compared to that of Pacific women. There was no difference in Shannon index (richness and evenness). Bacterial richness (number of microbiota present) was greater in women characterised by enterotype 3 than enterotype 1. Alpha diversity (evenness and Shannon index) was lower in women characterised by enterotype 2 compared to enterotypes 1 and 3 (Table 6.2).

Beta diversity

Jaccard Similarity Matrix and Bray-Curtis Dissimilarity indexes showed between-group diversity, and that the composition of Pacific women's microbiota was different to that of NZE women (see Figure 6.9). NZE women's community structures were most similar to each other when assessed with the Jaccard Similarity Matrix index (presence vs. absence of taxa). However, this trend changed when the relative abundance of the species was considered with the Bray-Curtis Dissimilarity index, and Pacific women's community structure were more similar to each other in comparison to NZE women. Differences were also observed between enterotypes. Enterotype 2 community structures were most similar to each other when assessed with Bray-Curtis Dissimilarity index, and least similar to each other when assessed with Jaccard Similarity Matrix index. In contrast, enterotype 3 community structures were most similar to each other when assessed with the Jaccard Similarity Matrix index, and more dissimilar to each other when assessed with the Bray-Curtis Dissimilarity index. Enterotype 1 had less similar community structures when assessed with Jaccard Similarity Matrix and Bray-Curtis Dissimilarity indexes (see Figure 6.9).

***Firmicutes* to *Bacteriodetes* (F:B) ratio**

There were no significant differences in F:B ratio between Pacific and NZE women (Table 6.1); but, participants with a low-BF% had a lower F:B ratio (5.2, 95 %CL 3.1, 9.5) compared to those with high-BF% (9.8, 95 %CL 4.4, 21.3). Enterotype 2 was associated with a high-BF% and a greater F:B ratio in comparison to enterotypes 1 and 3. There was no difference between enterotype 1 and 3 (Table 6.2).

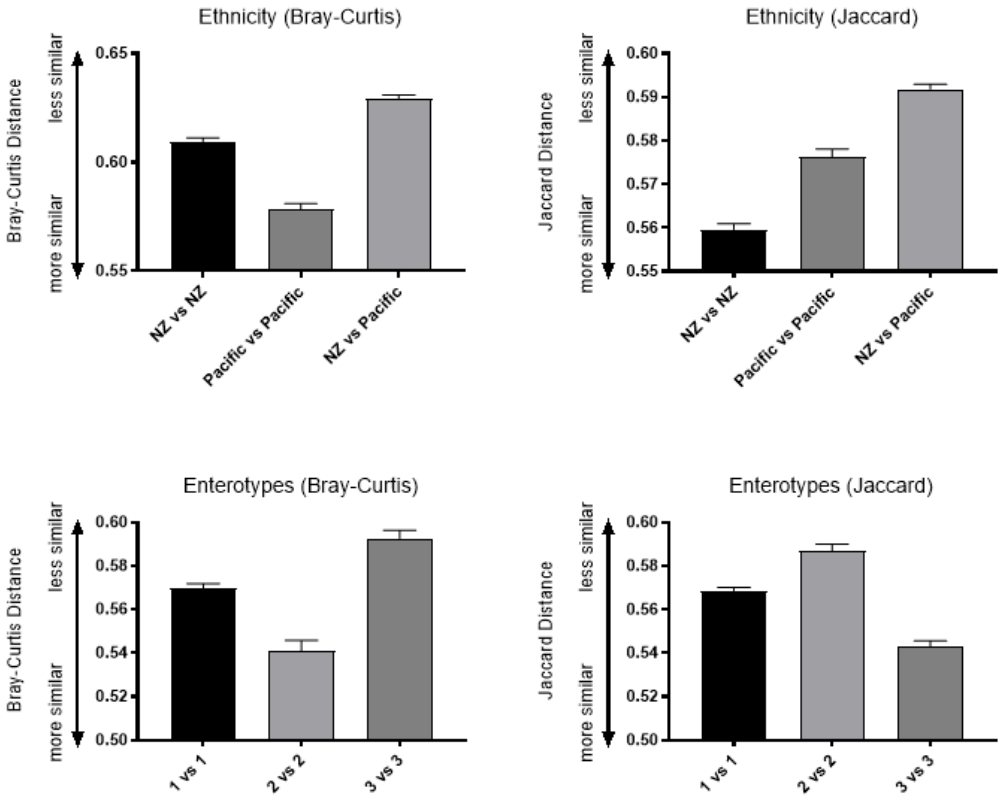


Figure 6.9. Beta diversity stratified by ethnicity and enterotypes

Bray-Curtis dissimilarity index and Jaccard distance matrix were applied by using QIIME2. Participants (n=286) were stratified by ethnicity (Pacific n=125, NZ European n=161) and enterotypes: 1: enterotype 1, n=146 (NZE n=89, Pacific n=57), 2: enterotype 2, n=70 (NZE n=11, Pacific n=59), 3: enterotype 3, n=70 (NZE n=61, Pacific n=9). Significance for each beta diversity metric was measured with PERMANOVA and group dispersion.

Faecal water content

There were no significant differences in faecal water content between Pacific and NZE women. However, faecal water content was significantly lower in women characterised by enterotype 3 (Table 6.2). In addition, the species that characterised enterotype 3 were all inversely associated with faecal water content: *Akkermansia muciniphila* (r = -0.20, p=0.001), *Ruminococcus bromii* (r = -0.19, p=0.001), *Subdoligranulum spp.* (r = -0.23, p<0.001), and *Methanobrevibacter smithii* (r = -0.14, p=0.022).

Dietary patterns

Results were similar for both Pacific and NZE women (data not shown) therefore all participants with available habitual dietary data (within the misreporting cut-offs) were included in the PCA. Four habitual dietary patterns were identified, and combined they explained 30.9 % of the total variance of the habitual dietary data. The dietary patterns were named according to the nutritional composition of the food groups which loaded highly (>0.4) onto the components. Colourful vegetables, dairy products, plant proteins, coffee and alcohol characterised the "colourful vegetables, plant protein, and dairy" pattern which explained the largest amount of variance (13.7 %) (Supplementary Table 6.2). Adherence to the "colourful vegetables, plant protein, and dairy" pattern was the only pattern significantly associated with all three enterotypes (data for other patterns not shown, and the characteristics of these dietary patterns have been presented previously see chapter five). Subsequent analyses focussed on this pattern.

6.4.3. Association of diet with enterotypes

There was no difference in energy intake between enterotypes (p=0.294). Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was positively associated with enterotypes 1 and 3, and negatively with enterotype 2 (see Table 6.3). Adjustment for BF% categories did not substantially alter the results (data not shown). For each enterotype, the significant associations between habitual intake of food groups and nutrients, following adjusted analyses, will be highlighted. The relative abundance of bacterial species which characterise each enterotype, and the association with food group intake, will also be highlighted (see Table 6.4 and Table 6.5).

Table 6.3. Adherence to the energy-adjusted "colourful vegetable, plant protein, and dairy products" dietary pattern with the odds of being in enterotypes 1 or 3

	OR **95 % CI**
Enterotype 1	2.72 [1.38, 5.35]**
Enterotype 3	2.98 [1.29, 6.86]**

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. Enterotype 1: n=145, Enterotype 2: n=68, Enterotype 3: n=69. OR: odds ratio. *P value <0.05, **P value ≤0.01, ***P value ≤0.001. Model OR: adjusted for age, ethnicity, faecal water content (%), and NZDep2013 index.

6.4.4. Association between dietary variables and enterotype 1

The food groups “green vegetables”, “nuts and seeds”, “sweetened cereals” (e.g., flavoured cereals, mueslis) and “legumes and meat alternatives” (e.g., beans, soy products, and plant-based proteins) were positively associated with enterotype 1 (Supplementary Table 6.3). At the nutrient level, total folate, total vitamin A (including beta carotene), calcium, iodine, total fat (including polyunsaturated and monounsaturated fat), vitamin E, and total dietary fibre were all positively associated with enterotype 1 (Supplementary Table 6.4). In contrast, “refined grains” (e.g., white rice and pasta and instant noodles) and “sugar sweetened beverages” (SSB) were negatively associated with enterotype 1. The relative abundance of *Faecalibacterium prausnitzii* was also positively associated “green vegetables” and “nuts and seeds” (Table 6.4, Supplementary Table 6.5). In contrast, “white bread” was negatively associated with the relative abundance of *Faecalibacterium prausnitzii*. Adjustment for BF% groups did not affect these results (data not shown).

6.4.5. Association between dietary variables and enterotype 2

In comparison to the other enterotypes, participants characterised by enterotype 2 consumed more of the following food groups: “fast food burgers”, “refined grains mixed dishes” (e.g., macaroni and cheese, white rice salad), “unsweetened cereals” (e.g., Weet-Bix, cornflakes), “fish and seafood” (e.g., canned, fresh), “crumbed and deep fried” (e.g., hot chips/fries, chicken nuggets). “SSB”, “added sugar to foods and beverages”, and “full fat milk” (see Supplementary Table 6.6). Enterotype 2 was negatively associated with the “colourful vegetables, plant protein, and dairy” pattern. In addition, certain food groups considered key components of this pattern were negatively associated with the relative abundance of species which characterised enterotype 2. For example, “nuts and seeds” were negatively associated to the relative abundance of *Lactobacillus ruminis* and *Bifidobacterium bifidum* (Table 6.5, Supplementary Table 6.7 and Supplementary Table 6.8). *Lactobacillus ruminis* was also negatively associated with “high fat cheese” and positively with “margarine”. The food groups “potatoes”, “starchy vegetables” (e.g., kumara, taro, yam) and “refined grains” were positively associated with the relative abundance of *Bifidobacterium adolescentis* (Table 6.5, Supplementary Table 6.9).

Table 6.4. Habitual food group intake significantly associated with the relative abundance of bacterial species

	Faecalibacterium prausnitzii OR1 95 % CI	*Subdoligranulum species* OR2 95 % CI
Low fat milk	1.00 [0.99, 1.00]*	
Green vegetables	1.01 [1.00, 1.02]**	
White bread	0.97 [0.95, 0.99]**	
Wholegrain products		1.01 [1.00, 1.03]*
Egg and egg products		1.02 [1.00, 1.04]*
Nuts and seeds	1.06 [1.02, 1.10]**	
Sweet snacks		0.98 [0.95, 1.00]*
Sugar sweetened beverages		1.00 [0.99, 1.00]*

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups are in units of g/day. Lower relative abundance is the reference category. OR: odds ratio. All models adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *P value <0.05, **P value ≤0.01, ***P value ≤0.001. OR1: Lower relative abundance: n = 140, higher relative abundance: n = 142, OR2: Lower relative abundance: n = 140, higher relative abundance: n = 142 (cut point is the median)

6.4.6. Association between dietary variables and enterotype 3

The food groups “egg and egg products” (e.g., whole eggs, quiche, omelette), “sweetened cereals” and “legumes and meat alternatives”, including the nutrients cholesterol, retinol, total folate, total vitamin A (including beta carotene), calcium, and iodine were all positively associated with enterotype 3 (see Supplementary Table 6.4)**.** The food groups “savoury snacks” (e.g., potato chips/crisps, popcorn), “refined grains” and “SSB” were negatively associated with enterotype 3 (see Supplementary Table 6.3)**.** The food groups “wholegrain products” (e.g., wholegrain bread and grains) and “egg and egg products” were positively associated with the relative abundance of *Subdoligranulum species;* whereas “sweet snacks” (e.g., chocolate, sweets, and muesli bars) and “SSB” were negatively associated (Table 6.4, Supplementary Table 6.10). The relative abundance of *Methanobrevibacter smithii* was negatively associated with the food groups “sweet snacks” and “other fruit” (Table 6.5, Supplementary Table 6.11). The food groups “other non-starchy vegetables” and “unsweetened cereals” were positively associated with the relative abundance of *Akkermansia muciniphila* (Table 6.5, Supplementary Table 6.12) and *Ruminococcus bromii* was positively associated with “green vegetables” and “margarine” (Table 6.5, Supplementary Table 6.13). Adjustment for BF% categories did not alter the results (data not shown).

Table 6.5. Habitual food group intake significantly associated with the presence or absence of bacterial species

	Lactobacillus ruminis **OR 95 % CI**	***Bifidobacterium bifidum*** **OR 95 % CI**	***Bifidobacterium adolescentis*** **OR 95 % CI**	***Methanobrevibacter smithii*** **OR 95 % CI**	***Akkermansia muciniphila*** **OR 95 % CI**	***Ruminococcus bromii*** **OR 95 % CI**
Women with presence of species	**n=63**	**n=116**	**n=220**	**n=114**	**n=174**	**n=195**
Low fat milk		1.00 [1.00, 1.01]*				
Milk alternatives		0.99 [0.98, 1.00]*				
Sweetened milk			0.99 [0.99, 1.00]*			
Dairy yoghurt						
High fat cheese	0.86 [0.76, 0.98]*					
Other fruit				0.99 [0.98, 1.00]*		
Green vegetables						1.02 [1.01, 1.03]**
Other non-starchy vegetables		0.97 [0.95, 1.00]*			1.03 [1.01, 1.06]*	
Potatoes			1.05 [1.02, 1.08]***			
Starchy vegetables			1.04 [1.00, 1.08]*			
Refined grain			1.04 [1.01, 1.08]*			
Unsweetened cereals					1.06 [1.01, 1.12]*	
Legumes and meat alternatives	1.02 [1.00, 1.05]*					
Peanuts		0.83 [0.70, 0.98]*				
Nuts and seeds	0.78 [0.62, 0.98]*	0.87 [0.80, 0.96]**				
Plant based fats		0.93 [0.87, 0.99]*				
Margarine	1.20 [1.03, 1.41]*					1.14 [1.00, 1.30]*
Sweet snacks				0.97 [0.94, 0.99]*		
Sugar sweetened beverages	1.01 [1.00, 1.01]*					

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups are in units of g/day. All models: Absence of species is the reference category and all models are adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *P value <0.05, **P value ≤0.01, ***P value ≤0.001. Only food groups with significant associations are reported in the table.

6.5. Discussion

This study provides greater understanding of the association between habitual diet and characteristics of the microbiota in relation to metabolic health markers. Metagenomic shotgun sequencing was used to characterise the microbiota of a large population of NZ women, who have different metabolic disease risk and body fat profiles. High quality dietary data was used to identify four habitual dietary patterns, and adherence to one of these patterns was associated with microbiota composition and metabolic health profiles (e.g., body composition and metabolic health markers). Further, following adjusted analyses, we associated the habitual intake of specific food groups with the relative abundance of bacterial species which characterised the enterotypes.

The participants in this study could be categorised into three groups by characterising their faecal microbiota composition. Characterisation of the microbiotas in terms of enterotypes gave three groupings; enterotype 2 was predominately found in Pacific women and enterotype 3 was predominated in NZE women. Enterotype 1 did not differ in the prevalence between the ethnicities. Dietary differences observed between enterotypes were reflected in their metabolic health profiles. Higher adherence to the "colourful vegetables, plant protein, and dairy" pattern and greater consumption of 'core' foods (e.g., vegetables, cheese, nuts and seeds, plant-based fats and coffee) was associated with a more favourable metabolic health profile (e.g., lower BMI, VAT, and fasting plasma insulin concentrations). In contrast, higher consumption of 'discretionary' foods was associated with a more detrimental metabolic health profile (e.g., higher BMI, VAT, fasting insulin concentrations). Further, we observed significant differences in age, ethnicity, deprivation, BMI, VAT, and BF% between enterotypes.

The degree of adherence to the "colourful vegetables, plant protein, and dairy" dietary pattern was associated with all three enterotypes. Higher adherence to this dietary pattern was positively associated with enterotypes 1 and 3, and negatively with enterotype 2. Therefore, the degree of adherence to this pattern was significantly associated with the composition of the microbiota. To the best of our knowledge, no other studies have associated dietary patterns with enterotypes. In previous observational studies, the number of enterotypes identified varies from two (18,26,62,63) to three (27,31,32) discrete clusters, and

they tend to be characterised by members of the genera *Bacteroides*, *Prevotella,* and the family *Ruminococcaceae.* The *Prevotella* genus was present in some participants, despite this there was no evidence of a *Prevotella* enterotype in this study population. In Western populations, *Prevotella* and *Bacteroides* are considered the driving taxa which explain the inter-individual differences in microbiota composition (64). However, the analytical depth of microbiota analysis will determine the quality of enterotyping, and most other enterotyping studies have not used metagenomic sequencing. Further, the association of these enterotypes (characterised in most studies) with particular diets is less clear.

Wu et al. (26), associated long term dietary intake (assessed with FFQ) with the *Bacteroides* enterotype (e.g., animal protein and fat intake) and *Prevotella* enterotype (e.g., carbohydrate intake). However, recent studies did not observe an association between dietary intake and enterotypes (31,33). These studies employed self-reported adherence to diet categories (e.g., vegetarian/vegan versus omnivore) as a classification measure. Such qualitative assessment of dietary intake fails to capture the actual composition of the diet, which can differ substantially between individuals within self-reported diet categories, due to factors such as age, sex, SES, culture etc. This may explain why they did not observe any associations between enterotypes and dietary intake. Losasso et al. (33), concluded that using such broad self-reported classifications of dietary intake are not useful to assess the composition of diet in relation to the microbiota composition, as it can lead to the oversimplification of results.

Recent studies have suggested that the microbiota influences host health outcomes.
Tang et al. (62), identified two enterotypes in a healthy population of males and females (n=136) and their observations were in agreement with Wu et al. (26), where long term diet (assessed with a semi quantitative FFQ) was more strongly associated with microbiota composition in comparison to short term diet (assessed with a three day food record) (62). Despite this, there was limited differences in dietary intake between individuals within the enterotypes, yet there were significant differences in plasma and stool metabolites between enterotypes (62). The authors proposed that the microbial enterotypes mediate the diet-host health relationship through the production of metabolites. A recent study in Danish adults (n=181) also proposed that enterotypes have the potential to influence host health outcomes, as individuals with a *Prevotella* enterotype lost more weight and BF% on a high fibre diet in

comparison to individuals with a *Bacteroides* enterotype (65,66). Together these studies highlight the importance of understanding the diet-microbiota relationship, and emphasise that enterotypes have the potential to influence host health outcomes.

Only a few observational studies have characterised the role of the overall diet and thus dietary patterns and their association with the microbiota composition and host health profiles. A recent study with German older adults (n=225) explored dietary patterns in relation to bacterial patterns which were both derived with treelet transformation. However, none of the five dietary patterns identified in that study predicted any of the seven bacterial patterns (37). In another study with older American men (n=517) there was no association with either the 'Western' (e.g., processed meat, refined grains, eggs, potatoes, and snacks) or 'prudent' pattern (e.g., fruit, vegetables, nuts, and white meat) and alpha diversity. However, the 'Western' pattern was negatively associated with the *Faecalibacterium* genera and positively with BMI, where the 'prudent' pattern showed the opposite association (36).

In this PhD research higher adherence to the "colourful vegetables, plant protein, and dairy" pattern was positively associated with enterotypes 1 and 3, which although not a 'prudent' pattern per se, it was characterised by higher loadings of core foods and was associated with lower BF% and BMI. Furthermore, enterotype 1 was characterised by higher relative abundance of the butyrate-producing bacterial species *Faecalibacterium prausnitzii* and *Eubacterium rectale*, as well as higher alpha diversity, and was positively associated with dietary fibre intake. A lower fibre diet is associated with lower levels of butyrate, because there are less available substrates for the microbiota to ferment to produce SCFA (67), suggesting greater fibre intake from core foods is associated with higher relative abundance of butyrate-producing bacterial species (*Eubacterium rectale* and *Faecalibacterium prausnitzii*) in these participants. Adding to this, the food groups "nuts and seeds" and "green vegetables" (rich in dietary fibre) were positively associated with the relative abundance of *Faecalibacterium prausnitzii*. This is in agreement with previous observations where higher intake of plant foods (68), dietary fibre from beans (35), and adherence to the Mediterranean diet (30) were associated with higher relative abundance of *Faecalibacterium prausnitzii*.

In a longitudinal follow up of adult female twins (n=1632) higher habitual dietary fibre intake was positively associated with alpha diversity (which was associated with lower weight gain over time), and lower BMI was positively associated with *Faecalibacterium prausnitzii* (34). Moreover, higher intake of 'junk food' and 'snacks' (i.e., lower fibre intake) has been negatively associated with the relative abundance of *Faecalibacterium prausnitzii* (29). Therefore, the observations of the present study are in agreement with diet-microbiota signatures previously identified across other studies, such as higher habitual fibre intake from core foods is associated with higher alpha diversity, including higher relative abundance of butyrate producing species and lower metabolic disease risk.

Higher adherence to the "colourful vegetable, plant protein, and dairy" pattern significantly reduced the likelihood of being in enterotype 2, and adjustment for BF% categories did not alter these results. There was no difference in energy intake between the enterotypes; however, women classified as enterotype 2 consumed more discretionary foods which, due to their energy density and nutrient poor characteristics, when consumed in excess are considered risk factors for higher adiposity (19,22,69). These women had a higher BMI and VAT in comparison to women classified as enterotypes 1 and 3, including higher BF% in comparison to enterotype 1. This is an important observation as higher VAT is associated with increased metabolic disease risks (70,71). Further, it is important to highlight that there was no difference in BF% between Pacific and NZE women, but there is a difference in BF% between enterotypes 1 and 2, and enterotype 1 is characterised by both ethnicities. In addition, despite Pacific women having higher circulating HbA1c, fasting insulin, and glucose concentrations in comparison to NZE women, the parameters for all participants were within the normal ranges (61). Although, enterotype 2 was associated with higher fasting plasma insulin concentrations, and hyperinsulinemia is a risk factor for metabolic disease as it precludes hyperglycaemia (61). Therefore, these results suggest that habitual diet is associated microbiota composition, food choice is associated with adiposity, and women classified as enterotype 2 have a higher multifactorial metabolic disease risk.

Despite all analyses being adjusted for deprivation, women in enterotype 2 were characterised by higher levels of deprivation in comparison to enterotypes 1 and 3. It is important to highlight that living in a deprived area in NZ increases the risk of being obese by

1.6 times (adjusted for age, sex, and ethnicity) (72). Thus, although these women were characterised by higher discretionary food intake (which would increase their risk of weight gain), the evidence suggests people with limited financial resources will often consume inexpensive energy dense 'discretionary' foods to save costs (73). There was no difference in energy intake between these enterotypes, however there was in the foods which contributed to these energy intakes. The evidence suggests food choice is associated with obesity and deprivation, and despite these women having different metabolic disease risk, consuming more nutrient rich core foods was associated with lower adiposity, metabolic health risks, and socioeconomic deprivation.

It is well established that a lower fibre diet is considered a risk factor for weight gain (34), including lower microbiota diversity (74). In the present study women characterised as enterotype 2 had lower alpha diversity (Shannon index) and higher BMI and BF%. Enterotype 2 was negatively associated with the "colourful vegetables, plant protein, and dairy" pattern and positively associated with a greater intake of carbohydrate based discretionary foods (i.e., lower fibre intake). These observed diet-microbiota associations also aligned with the functional capacity of the microbiota species which characterised enterotype 2. Enterotype 2 was characterised by higher relative abundance of lactic acid-producing bacterial species, *Bifidobacterium adolescentis*, *Bifidobacterium bifidum*, and *Lactobacillus ruminis.* Higher intake of chemically modified resistant starch (RS4) has been associated with enriching the relative abundance of *Bifidobacterium adolescentis* (75)*,* and the food group "refined grains" was positively associated with the relative abundance of *Bifidobacterium adolescentis.* Although we cannot confirm the resistant starch content of this food group, we can speculate that chemically modified RS4 would be present, and the likelihood that discretionary foods would also contain RS4. In comparison, higher intake of type 2 resistant starch (RS2: intrinsically resistant to digestion e.g., found in fruit, corn, potatoes) promotes the relative abundance of *Ruminococcus bromii* and *Eubacterium rectale* (75)*,* as well as *Bifidobacterium adolescentis* for some individuals (76)*.* In the PhD research, higher intake of starchy food groups "potatoes" and "starchy vegetables", were positively associated with the relative abundance of *Bifidobacterium adolescentis*, which is a starch degrading bacteria. Therefore, these observations clearly highlight the association between habitual dietary intake and microbiota composition. Including, associations with metabolic health, because higher

discretionary food intake (and lower fibre intake) was associated with higher adiposity, and lower microbiota diversity, in a population of women with different metabolic disease risk and body fat profiles.

In general, core foods were positively associated with the relative abundance of the species which characterised enterotype 3, where discretionary foods were negatively associated. For example, the fibre rich food groups “other non-starchy vegetables” and “unsweetened cereals” were positively associated with the relative abundance of *Akkermansia muciniphila,* where “wholegrain products” and “egg and egg products” were positively associated with *Subdoligranulum species,* and “sweet snacks” and “SSB” were negatively associated. Higher relative abundance of *Akkermansia muciniphila* has been associated with a healthier metabolic status and improved insulin sensitivity (77). In contrast, higher SSB intake has been negatively associated with alpha diversity (38,39) and lower relative abundance of *Akkermansia muciniphila* (29).

Women characterised by enterotype 3 were significantly older and had lower levels of deprivation in comparison to women in either enterotypes 1 and 2. Older women are more likely to attempt to manage their weight (78), and higher quality diets (i.e., consuming a diverse range of core foods) are consistently associated with higher SES (78). Thus, these observations may be a reflection of such trends. However, higher habitual dietary fibre intake, notably from core or ‘non-discretionary’ food groups would provide adequate substrates for the microbiota, which in turn could promote SCFA production with the potential to improve insulin sensitivity and metabolic health outcomes for the host (13,14,79). In addition, habitually consuming foods rich in dietary fibre, whilst limiting the intake of refined and processed foods is associated with lower BF% (80), a reduced risk of obesity (19,22,69) and improved metabolic health outcomes (20,34,79). Collectively, the observations of the present study highlight diet-microbiota-host associations, as higher adherence to the “colourful vegetables, plant protein, and dairy” pattern was associated with enterotypes 1 and 3 and a more favourable metabolic health profile. In contrast, higher discretionary food intake was positively associated with enterotype 2 and potentially higher metabolic disease risks.

The observations of the current study have highly supported microbiota associations reported in the literature (as mentioned above). It has also been suggested that methane delays colonic transit time by slowing down intestinal motility (81); the removal of hydrogen through methane production is proposed to alter gut fermentation which, in turn, could affect colonic peristalsis (82). Women characterised by enterotype 3 had a significantly lower faecal water content in comparison to enterotype 1 and 2, and the species that characterised enterotype 3 were all inversely associated with faecal water content. This suggests, that these species are associated with firmer stool consistencies and may have a fitness to grow in conditions of slower colonic transit. In turn, through methane production, *Methanobrevibacter smithii* may influence the colonic environment which favours the growth of these species. In addition, lower alpha diversity (observed in enterotype 2) was associated with higher VAT (8), and higher BMI and adiposity (BF%) were associated with a higher F:B ratio. This F:B ratio was significantly higher in enterotype 2 (whom had a higher percentage of women with BF% ≥35 %). We cannot confirm whether higher F:B ratio is an indicator of higher body weight/obesity, as it might only be an indicator of belonging to enterotype 2. However, the observed differences in BF%, F:B ratio, and Shannon index between enterotypes, but not ethnicities, highlights that although the enterotypes are specific to the study population, enterotypes are a powerful tool to identify metabolic disease risk in a healthy population.

There are several strengths in the present study. Firstly, the large sample size of free-living disease-free adults. This study was unique in its selective recruitment of Pacific and NZE women with different metabolic disease risk and body fat profiles. We also assessed several measures of body composition (BMI, BF%, VAT) and metabolic health markers (fasting insulin, glucose and HOMA-IR) to characterise metabolic health profiles associated with habitual diet and microbiota profiles (enterotypes). Secondly, the high quality of the underlying dietary data, and thirdly, using metagenomic shotgun sequencing enabled greater analytical depth compared to other relevant studies. Investigating dietary patterns is more reflective with regard to how people eat. Utilising such methodologies might highlight the synergistic effect of foods (which deliver nutrients) consumed together and their effect on the microbiota ecosystem. Most microbiota studies that investigate associations with habitual diet use self-reported FFQs which rely on memory, are prone to overestimation, and only a few studies adjust for energy intake or consider person-specific effects. Participant's of the present study

completed a prospective 5DFR before they completed their faecal sample, and a semi-quantitative NZWFFQ (which overlapped with their 5DFR and faecal sample). Both dietary tools (measuring actual and habitual dietary intake) were included in the modelled habitual dietary data (considering person specific effects). Further, the NCI modelling process specifically addressed intra-individual variation inherent in dietary data, and including the NZWFFQ information improved the precision of the estimation of episodically consumed foods (83). Furthermore, the dietary intake of participants was characterised at the nutrient, food group and dietary pattern level which gave substantial logical depth to the observed results. Where possible, to improve the clarity of results all analyses were adjusted for covariates which could either affect dietary intake or microbiota composition.

There are several limitations of the present study. Self-reported dietary data was the primary source of dietary information, despite self-reported dietary data being prone to under- and over-estimation of intake (40,41) it offers insight into the complexities of what individuals eat that no set of biomarkers are currently able to provide (84). In addition, the energy reporting cut-offs of <2100 kJ/day and >27000 kJ/day employed within this study are higher than the parameters recommended by Willett (28). However, all energy intakes that were used in the analysis were intensively checked and considered realistic values for the population (85–87). The generalisability of these observations are limited to the study population. Additionally, causality cannot be inferred due to the cross-sectional study design, and there are other factors besides diet which can influence metabolic health profiles, such as healthy lifestyle in general. Furthermore, the use of enterotypes as a classification measure identifies clusters of bacterial species based on their presence and relative abundances. Thus, we cannot infer anything about their integrated function. However, the enterotypes we have characterised are defined by species with known attributes such as butyrate produces, starch degraders, and mucin degraders which further sets each enterotype apart. Therefore, the observations of the current study are in agreement at many levels of stratification and analyses, including with the literature.

6.6. Conclusion

In a population of Pacific and NZE women with different metabolic disease risk, their microbiota was characterised into three enterotypes based on the presence of distinct species. Four dietary patterns were identified; and one pattern, the "colourful vegetables, plant protein, and dairy", was associated with all three enterotypes. This pattern was associated with higher 'core' food groups which was positively associated with enterotype 1 (characterised by Pacific and NZE women) and enterotype 3 (characterised by NZE women). Enterotype 2 was negatively associated with the "colourful vegetables, plant protein, and dairy" pattern and characterised by higher 'discretionary' food intake, and women with increased metabolic disease risks. We cannot confirm whether the microbiota is a cause or consequence of metabolic health. However, habitual diet is associated with metabolic health outcomes, and this study clearly highlights habitual diet-microbiota associations, in relation to metabolic disease risks.

From a public health perspective, similar diet-microbiota-host associations were observed across ethnicities, suggesting these findings might potentially translate to other ethnicities and cultures in future research. To advance our understanding of the association between diet and the microbiota, and to work towards developing therapeutic targets, we need more well-designed longitudinal and intervention studies with repeated measures (within individuals) and in different populations. Collecting data at multiple timepoints would enable the assessment of the stability of the microbiota, and the association of dietary intake and metabolic health biomarkers overtime. Enterotyping and analyses of habitual diet with the NCI method are potential strategies to standardised diet-microbiota methodology. Future studies should consider dietary pattern assessment to explore the whole diet in relation to the microbiota. Together this could guide allocation to treatment in future studies, to determine the efficacy of interventions and potential causality.

6.7. References

1. Bäckhed F, Manchester JK, Semenkovich CF, Gordan JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. Proc Natl Acad Sci. 2007;104(3):979–84. doi:10.1007/BF00540675

2. 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

3. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science (80-) [Internet]. 2013 Sep 6;341(6150):1241214–1241214. doi:10.1126/science.1241214

4. Ley RE, Turnbaugh PJ, Klein S, Gordan JI, Gordon JI. Human gut microbes associated with obesity. Nature [Internet]. 2006;444(7122):1022–3. doi:10.1038/nature4441021a

5. Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci. 2005;102(31):11070–5. doi:10.1073/pnas.0504978102

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

7. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. Nature [Internet]. 2013;500(7464):541–6. doi:10.1038/nature12506

8. Beaumont M, Goodrich JK, Jackson MA, Yet I, Davenport ER, Vieira-Silva S, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Genome Biol [Internet]. 2016;17(1). doi:10.1186/s13059-016-1052-7

9. Manichanh C. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut [Internet]. 2006;55(2):205–11. doi:10.1136/gut.2005.073817

10. Lambeth SM, Carson T, Lowe J, Ramaraj T, Leff JW, Luo L, et al. Composition diversity and abundance of gut microbiome in prediabetes and type 2 diabetes. J Diabetes Obes. 2015 Dec 26;2(2):108–14. doi:10.15436/2376-0949.15.031

11. Mann J, Cummings JH, Englyst HN, Key T, Liu S, Riccardi G, et al. FAO/WHO Scientific update on carbohydrates in human nutrition: Conclusions. Eur J Clin Nutr. 2007;61(S1):S132–7. doi:10.1038/sj.ejcn.1602943

12. Cummings JH, Stephen AM. The role of the dietary fibre in the human colon. Can Med Assoc J. 1980;123(11):1109–14.

13. Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: Short-chain fatty acids as key bacterial metabolites. Cell. 2016;165(6):1332–45. doi:10.1016/j.cell.2016.05.041

14. Cummings JH, Macfarlane GT. The control and consequences of bacterial fermentation in the human colon. J Appl Bacteriol [Internet]. 1991 Jun;70(6):443–59.

doi:10.1111/j.1365-2672.1991.tb02739.x

15. Cani PD, Van Hul M, Lefort C, Depommier C, Rastelli M, Everard A. Microbial regulation of organismal energy homeostasis. Nat Metab [Internet]. 2019;1(1):34–46. doi:10.1038/s42255-018-0017-4

16. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med [Internet]. 2013;472:57–63. doi:10.1038/nm.3145

17. Tang W, Wang Z, Levison B, Koeth R, Britt E, Fu X, et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med. 2013;368(17):1575–84. doi:10.1056/NEJMoa1109400

18. Claesson M, Jeffery IB, Conde S, Power S, O'Connor E, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;(488):178–84. doi:10.1038/nature11319

19. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

20. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9

21. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8

22. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.

23. Wang DD, Li Y, Afshin A, Springmann M, Mozaffarian D, Stampfer MJ, et al. Global improvement in dietary quality could lead to substantial reduction in premature death. J Nutr [Internet]. 2019;149(6):1065–74. doi:10.1093/jn/nxz010

24. Albenberg LG, Wu GD. Diet and the intestinal microbiome: Associations, functions, and implications for health and disease. Gastroenterology [Internet]. 2014;146(6):1564–72. doi:10.1053/j.gastro.2014.01.058

25. Pendyala S, Walker JM, Holt PR. A high-fat diet is associated with endotoxemia that originates from the gut. Gastroenterology [Internet]. 2012;142:1100–1. doi:10.1053/j.gastro.2012.01.034

26. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science (80-) [Internet]. 2011 Oct 7;334(6052):105–8. doi:10.1126/science.1208344

27. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science (80-) [Internet]. 2016 Apr

29;352(6285):560–4. doi:10.1126/science.aad3503

28. De Filippis F, Pellegrini N, Vannini L, Jeffery IB, La Storia A, Laghi L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut. 2016;65(11):1812–21. doi:10.1136/gutjnl-2015-309957

29. Mitsou EK, Kakali A, Antonopoulou S, Mountzouris KC, Yannakoulia M, Panagiotakos DB, et al. Adherence to the Mediterranean diet is associated with the gut microbiota pattern and gastrointestinal characteristics in an adult population. Br J Nutr [Internet]. 2017;117(12):1645–55. doi:10.1017/s0007114517001593

30. Gutiérrez-Díaz I, Fernández-Navarro T, Salazar N, Bartolomé B, Moreno-Arribas MV, de Andres-Galiana EJ, et al. Adherence to a Mediterranean diet influences the fecal metabolic profile of microbial-derived phenolics in a Spanish cohort of middle-age and older people. J Agric Food Chem [Internet]. 2017 Jan 25;65(3):586–95. doi:10.1021/acs.jafc.6b04408

31. de Moraes ACF, Fernandes GR, da Silva IT, Almeida-Pititto B, Gomes EP, Pereira A da C, et al. Enterotype may drive the dietary-associated cardiometabolic risk factors. Front Cell Infect Microbiol. 2017;7(FEB):1–9. doi:10.3389/fcimb.2017.00047

32. Wu GD, Compher C, Chen EZ, Smith SA, Shah RD, Bittinger K, et al. Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production. BMJ. 2016;65(1):63–72. doi:10.1136/gutjnl-2014-308209

33. Losasso C, Eckert EM, Mastrorilli E, Villiger J, Barrucci F, Pernthaler J, et al. Assessing the influence of vegan, vegetarian and omnivore oriented westernized dietary styles on human gut microbiota: A cross sectional study. Front Microbiol. 2018;9(March):1–12. doi:10.3389/fmicb.2018.00317

34. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int J Obes. 2017;41(7):1099–105. doi:10.1038/ijo.2017.66

35. Lin D, Peters BA, Friedlander C, Freiman HJ, Goedert JJ, Sinha R, et al. Association of dietary fibre intake and gut microbiota in adults. Br J Nutr [Internet]. 2018;120(9):1014–22. doi:10.1017/s0007114518002465

36. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. Am J Clin Nutr [Internet]. 2019 Oct 1;110(4):1003–14. doi:10.1093/ajcn/nqz174

37. Oluwagbemigun K, Foerster J, Watkins C, Fouhy F, Stanton C, Bergmann MM, et al. Dietary patterns are associated with serum metabolite patterns and their association is influenced by gut bacteria among older german adults. J Nutr [Internet]. 2019 Sep 2;1–10. doi:10.1093/jn/nxz194

38. Partula V, Mondot S, Torres MJ, Kesse-Guyot E, Deschasaux M, Assmann K, et al. Associations between usual diet and gut microbiota composition: Results from the Milieu Intérieur cross-sectional study. Am J Clin Nutr. 2019;109(5):1472–83. doi:10.1093/ajcn/nqz029

39. 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 (80-) [Internet]. 2016 Apr 29;352(6285):565–9. doi:10.1126/science.aad3369

40. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001

41. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.

42. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

43. Kindleysides S, Kruger R, Douwes J, Tannock G, Renall N, Slater J, et al. PRedictors linking Obesity and the gut MIcrobiomE (The PROMISE study): Protocol and recruitment strategy of a cross-sectional study on pathways that affect the gut microbiome and its impact on obesity. JMIR Res Protoc. 2019;8:1–16. doi:10.2196/14529

44. Stewart A, Marfell-Jones M, Olds T, Hans De Ridder J. International standards for anthropometric assessment. Lower Hutt, New Zealand: International Society for the Advancement of Kinanthropometry; 2011.

45. Garrow J. Quetelet's index (W/H2) as a measure of fatness. J Obes. 1985;9(2):147–53.

46. Atkinson J, Salmond C, Crampton P. NZDep2013 index of deprivation. NZDep2013 Index of Deprivation. Wellington; 2014.

47. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: Insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia. 1985;28(7):412–9. doi:10.1007/BF00280883

48. Beck KL, Houston ZL, McNaughton SA, Kruger R. Development and evaluation of a food frequency questionnaire to assess nutrient intakes of adult women in New Zealand. Nutr Diet. 2018;1–7. doi:10.1111/1747-0080.12472

49. Ministry of Health. Eating and activity guidelines for New Zealand adults [Internet]. Wellington; 2015.

50. National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. Nutrient reference values for Australia and New Zealand. Nutrition. Canberra; 2006.

51. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

52. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

53. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes.

Nat Methods [Internet]. 2012;9(8):811–4. doi:10.1038/nmeth.2066

54. Bolyen E, Dillon M, Bokulich N, Abnet C, Al-Ghalith G, Alexander H, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Prepr. 2018; doi:10.7287/peerj.preprints.27295

55. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol [Internet]. 2001 Feb;26(1):32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

56. Anderson MJ. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics [Internet]. 2006;62(1):245–53. doi:10.1111/j.1541-0420.2005.00440.x

57. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature [Internet]. 2011;473(7346):174–80. doi:10.1038/nature09944

58. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics [Internet]. 2014;30(21):3123–4. doi:10.1093/bioinformatics/btu494

59. Oliveros E, Somers VK, Sochor O, Goel K, Lopez-Jimenez F. The concept of normal weight obesity. Prog Cardiovasc Dis. 2014;56(4):426–33. doi:10.1016/j.pcad.2013.10.003

60. Kramer C, Zinman B, Retnakaran R. Are metabolically healthy overweight and obesity benign conditions? A systematic review and meta-analysis. Ann Intern Med. 2013;159(11):758–69. doi:10.7326/0003-4819-159-11-201312030-00008

61. New Zealand Guidelines Group. New Zealand primary care handbook 2012 [Internet]. 3rd ed. New Zealand Primary Care Handbook. Wellington: Ministry of Health; 2012. 45–64 p.

62. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, et al. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. Front Genet [Internet]. 2019;10:454. doi:10.3389/fgene.2019.00454

63. Dhakan DB, Maji A, Sharma AK, Saxena R, Pulikkan J, Grace T, et al. The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. Gigascience. 2019;8(3):1–20. doi:10.1093/gigascience/giz004

64. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, et al. Enterotypes in the landscape of gut microbial community composition. Nat Microbiol [Internet]. 2018 Jan 18;3(1):8–16. doi:10.1038/s41564-017-0072-8

65. Hjorth MF, Roager HM, Larsen TM, Poulsen SK, Licht TR, Bahl MI, et al. Pre-treatment microbial Prevotella-to-Bacteroides ratio, determines body fat loss success during a 6-month randomized controlled diet intervention. Int J Obes [Internet]. 2018;42:580–3. doi:10.1038/ijo.2017.220

66. Hjorth MF, Blædel T, Bendtsen LQ, Lorenzen JK, Holm JB, Kiilerich P, et al. Prevotella-to-Bacteroides ratio predicts body weight and fat loss success on 24-week diets varying in macronutrient composition and dietary fi ber: Results from a post-hoc analysis. Int J

Obes. 2019;43:149–57. doi:10.1038/s41366-018-0093-2

67. Litvak Y, Byndloss MX, Bäumler AJ. Colonocyte metabolism shapes the gut microbiota. Science (80-) [Internet]. 2018 Nov 30;362(6418):eaat9076. doi:10.1126/science.aat9076

68. Mcdonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G, et al. American gut: an open platform for citizen science microbiome research. Am Soc Microbiol. 2018;3(3):1–28. doi:10.1128/mSystems.00031-18

69. Schlesinger S, Neuenschwander M, Schwedhelm C, Hoffmann G, Bechthold A, Boeing H, et al. Food groups and risk of overweight, obesity, and weight gain: A systematic review and dose-response meta-analysis of prospective studies. Adv Nutr. 2019;10(2):205–18. doi:10.1093/advances/nmy092

70. Fox CS, Massaro JM, Hoffmann U, Pou KM, Maurovich-Horvat P, Liu C-Y, et al. Abdominal visceral and subcutaneous adipose tissue compartments. Circulation. 2007;116(1):39–48. doi:10.1161/circulationaha.106.675355

71. Després J-P. Body fat distribution and risk of cardiovascular disease. Circulation. 2012;126(10):1301–13. doi:10.1161/circulationaha.111.067264

72. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].

73. Darmon N, Briend A, Drewnowski A. Energy-dense diets are associated with lower diet costs: A community study of French adults. Public Health Nutr. 2004;7(1):21–7. doi:10.1079/PHN2003512

74. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL. Diet-induced extinctions in the gut microbiota compound over generations. Nature [Internet]. 2016;529:212–5. doi:10.1038/nature16504

75. Martínez I, Kim J, Duffy PR, Schlegel VL, Walter J. Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. PLoS One [Internet]. 2010;5(11):e15046. doi:10.1371/journal.pone.0015046

76. Venkataraman A, Sieber JR, Schmidt AW, Waldron C, Theis KR, Schmidt TM. Variable responses of human microbiomes to dietary supplementation with resistant starch. Microbiome [Internet]. 2016 Dec 29;4(1):33. doi:10.1186/s40168-016-0178-x

77. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: Relationship with gut microbiome richness and ecology. Gut [Internet]. 2016;65(3):426–36. doi:10.1136/gutjnl-2014-308778

78. Darmon N, Drewnowski A. Does social class predict diet quality? Am J Clin Nutr [Internet]. 2008 May 1;87(5):1107–17. doi:10.1093/ajcn/87.5.1107

79. Zhao L, Zhang F, Ding X, Wu G, Lam YY, Wang X, et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science (80-) [Internet]. 2018 Mar 9;359(6380):1151–6. doi:10.1126/science.aao5774

80. Tucker LA, Thomas KS. Increasing total fiber intake reduces risk of weight and fat gains in women. J Nutr. 2009;139(3):576–81. doi:10.3945/jn.108.096685

81. Pimentel M, Lin HC, Enayati P, van den Burg B, Lee H-R, Chen JH, et al. Methane, a gas produced by enteric bacteria, slows intestinal transit and augments small intestinal contractile activity. Am J Physiol Liver Physiol [Internet]. 2006 Jun;290(6):G1089–95. doi:10.1152/ajpgi.00574.2004

82. Sahakian AB, Jee S-R, Pimentel M. Methane and the gastrointestinal tract. Dig Dis Sci [Internet]. 2010 Aug 15;55(8):2135–43. doi:10.1007/s10620-009-1012-0

83. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009;65(4):1003–10. doi:10.1111/j.1541-0420.2009.01223.x

84. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr. 2015;145(12):2639–45. doi:10.3945/jn.115.219634

85. Gemming L, Jiang Y, Swinburn B, Utter J, Mhurchu CN. Under-reporting remains a key limitation of self-reported dietary intake: An analysis of the 2008/09 New Zealand Adult Nutrition Survey. Eur J Clin Nutr [Internet]. 2014;68(2):259–64. doi:10.1038/ejcn.2013.242

86. Smith C, Gray AR, Mainvil LA, Fleming EA, Parnell WR. Secular changes in intakes of foods among New Zealand adults from 1997 to 2008/09. Public Health Nutr [Internet]. 2015 Dec 10;18(18):3249–59. doi:10.1017/S1368980015000890

87. Metcalf PA, Scragg RRK, Schaaf D, Dyall L, Black PN, Jackson R. Dietary intakes of European, Māori, Pacific and Asian adults living in Auckland: The Diabetes, Heart and Health Study. Aust N Z J Public Health [Internet]. 2008;32(5):454–60. doi:10.1111/j.1753-6405.2008.00279.x

Supplementary Table 6.1. List of 55 food groups

	Food Group	Food Items Included
1	Full fat dairy milk	Dark blue, purple, silver top milk, lactose free regular fat milk
2	Low fat dairy milk	Lite and trim milk (green or light blue), lactose free reduced fat milk
3	Milk alternatives, including sweetened	Soy, almond, rice, oat, coconut varieties, milk alternative based drinks from cafés
4	Sweetened dairy milk products	Flavoured milk, fermented or evaporated milk, breakfast drinks (e.g., Up & Go), Yakult fermented milk drink, hot chocolates, milk-based smoothies, milk-based drinks from cafés, coffee sachets, coffees made with syrup and cream (e.g., caramel macchiato)
5	Dairy yoghurt	All types of cows milk yoghurt. (Note: Soy yoghurt under soy products, coconut yoghurt under coconut fats)
6	High fat cheese	Cream cheese, goat cheese, haloumi, parmesan, cheddar, processed cheese, blue vein, mascarpone
7	Low fat cheese	Brie, bocconcini, edam, feta, mozzarella, camembert, cottage, ricotta, paneer
8	Apple, banana, orange	Apple, banana orange
9	Other fruit	All other fruit (fresh, canned, dried)
10	Tomatoes	Fresh, canned, cooked tomatoes
11	Dark yellow vegetables	Carrots, pumpkin, butternut squash
12	Green vegetables	Lettuce, spinach, cabbage, broccoli, watercress, green beans, brussel sprouts, courgette
13	Other non-starchy vegetables	Capsicum, onion, mushrooms, frozen mixed vegetables, beetroot, squash
14	Potatoes / potato dishes (excluding chips)	Potato (boiled, mashed, baked, salad, scalloped, roasted)
15	Starchy vegetables	Kumara, yam, parsnip, turnip, swedes (boiled, mashed, baked) Taro (flesh, roots, stalks), green banana, sweet corn kernels, breadfruit, cassava, green banana
16	White breads	Plain white bread, wraps, focaccia, bagels, pita bread, rēwena bread, doughboys, breadcrumbs, including gluten free options, naan (plain)
17	Discretionary breads	Crumpets, scone, savoury muffin, plain croissant, pancakes, waffles, iced bun, savoury pin wheels, garlic bread, fruit bread, roti, naan (garlic)
18	Crackers	All crackers made from grains, cream crackers, Cruskits, rice crackers
19	Whole grain products	Wholegrain breads (High fibre, wholemeal, wholegrain, including gluten free options), grains (Quinoa, buckwheat, bulgur wheat, brown rice, wholemeal pasta, wholegrain gluten free pasta e.g., brown rice)
20	Refined grains	White rice, white pasta, noodles (instant, egg, rice), canned spaghetti, cous cous, including gluten free pasta
21	Refined grain mixed dishes	Macaroni and cheese, carbonara, white rice salad, two-minute noodles
22	Oats	Porridge, rolled oats, oat bran
23	Sweetened cereal	Sultana bran, light and fruity cereal, chocolate-based cereals, Nutri-Grain, honey puffs, milo cereals, fruit loops, oat sachets, all muesli and granola
24	Unsweetened cereals	Weet-Bix, bran cereals, rice bubbles, cornflakes
25	Red meats	Beef, lamb, venison, mince, patties, (including red meat mixed dishes: Stir fry, curry, stew)
26	White meats	Chicken, pork, turkey, (including white meat mixed dishes: Stir fry, curry, stew, casserole)
27	Processed meats	Corned beef (canned), corned silverside, smoked chicken, smoked hock and salami, ham, sausages, frankfurters, bacon, chorizo, luncheon meat
28	Fish and seafood	Canned and fresh (including mixed dishes Oka ika mata, curry, stew) and processed fish products (e.g., fish balls)

29	Eggs	Whole eggs (boiled, poached, fried, plain omelette) including egg mixed dishes (Quiche, frittata, omelette with filling, egg and banana pancake)
30	Legumes and meat alternatives	Baked beans, black beans, dahl, canned or dried legumes, hummus, and legume based vegetarian meals and products (including meat alternatives and soy products edamame beans tofu, tempeh)
31	Peanut butter and peanuts	Peanut butter and peanuts
32	Nuts and seeds	Brazil nuts, walnuts, almond, cashew, pistachio, chia, linseed, pumpkin, sesame
33	Animal fats	Cream, sour cream, reduced cream, butter, lard, dripping, ghee
34	Coconut fats and products	Coconut oil, cream, milk, desiccated, yoghurt, fresh
35	Plant based fats	Avocado (whole fruit), canola, sunflower, olive, vegetable oil, cooking spray, oil-based salad dressings (French/Italian)
36	Margarine	All margarines
37	Creamy dressings and sauces	Creamy readymade meal-based sauces, dips, mayonnaise, aioli, tartare sauce, white sauce, cheese sauce
38	Savoury sauces and condiments	Curry pastes, herb and spices, vinegar, gravy and garlic sauce, oil based condiments such as sundried tomatoes / olives in oil, pesto, tomato, barbeque, mint, soy, gravy, mustard, chutney, miso, pasta sauce, tomato paste, sweet chilli sauce (including savoury spreads Vegemite, marmite)
39	Sweet spreads	Jam, honey, marmalade, syrup (maple, golden), Nutella, chocolate peanut butter, chocolate butter
40	Cake and biscuits	Slices, cakes, loaves, muffins, biscuits, doughnuts, sweet pies, pastries, tarts
41	Puddings and other desserts	Including all milk alternative ice creams, ice cream, custard, milk-based puddings (e.g., rice, instant, semolina, pavlova, sticky date, fruit pies and crumbles, jelly, ice blocks)
42	Sweet snack foods	Fruit and nut mixes, bliss balls, chocolate, lollies, muesli bars
43	Savoury snack foods	Popcorn, potato crisps, corn chips, Twisties, bhuja mix
44	Crumbed and deep fried	Hot chips/fries, hash browns, and packaged home baked chips, wontons, paraoa (fry) bread, schnitzel, nuggets, crumbed fish
45	Fast-food (burgers)	Pies, dumpling, burgers, pizzas, curries, noodle-based dishes, Nandos chicken, egg fu yong
46	Fast food (salads & sushi)	Salads, sandwiches, wraps, sushi, vegetable-based stir fry
47	Fruit and vegetable juice	Fruit and/or vegetable juice, fruit and/or vegetable smoothies
48	Soft drinks and other sugar sweetened beverages	All cordials, flavoured water, sports drinks, soft drinks, fruit drinks, iced tea, energy drinks
49	Diet drinks	All exclusively artificially sweetened beverages
50	Tea	Black, green, herbal, chai, kombucha
51	Coffee	Instant, brewed, espresso, pre-mixed sachet, filter, cold brew
52	Alcoholic beverages	Wine (standard and low alcohol), beer (standard and low alcohol), cider, spirits, RTDs, sherry, port, liqueurs, sake
53	Water	Water (unflavoured, soda, tap)
54	Sugar added to food and drink	All sugar added to food and drink
55	Soups and stock	All soups (instant, canned, packet) and stocks

Supplementary Table 6.2. Component loadings of four habitual dietary patterns

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Explanation of variation of food intake (%)	13.7	6.2	5.7	5.4
Cronbachs's alpha	0.8	0.7	0.7	0.7
Other non-starchy vegetables	0.79			
Green vegetables	0.71		0.33	
Savoury sauces and condiments	0.66			
Low fat cheese	0.65			
Legumes and meat alternatives	0.64			-0.39
Yellow vegetables	0.59		0.44	
High fat cheese	0.58			
Tomatoes	0.57			
Fast food burgers	-0.57	0.38		
Nuts and seeds	0.56		0.41	
Plant based fats	0.56			
Coffee	0.55			
Refined grains mixed dishes	-0.52	0.30		
Alcoholic beverages	0.51			
Unsweetened cereals	-0.49			
Sugar sweetened beverages	-0.47	0.37		
Dairy yoghurt	0.43			
Egg and egg products	0.41			
Crumbed and deep-fried food	-0.41			
Added sugar to food and beverages	-0.34			
Tea	0.33			
Sweet snack foods	0.31			
Water				
Fruit and vegetable juice				
Sweetened milk products		0.55		
Potatoes		0.54		
Discretionary breads	0.32	0.51		
Puddings and desserts		0.50		
Fast food salad and sushi		0.49		
Creamy based sauces and dressings		0.43		
Cakes and biscuits		0.41		
White bread		0.38		
Savoury snack foods		0.32		
Margarine		0.32		
Diet drinks				
Low fat milk				
Refined grains				
Sweet spreads				
Apple, banana, orange			0.57	
Starchy vegetables			0.45	
Other fruit			0.44	
Coconut products			0.44	
Peanut butter and peanuts	0.36		0.43	
Fish and seafood			0.37	0.34
Milk alternatives			0.35	
Wholegrain products		0.31	0.34	
Oats				
Soups and stocks				
Crackers				
Sweetened cereals				
Red meat				0.82
White meat				0.82
Processed meat				0.78
Animal fats				0.43
Full fat milk				

Patterns identified based on component loading >0.3. Kaiser-Meyer-Olkin measure of sampling adequacy=0.795, Bartlett's test of sphericity = 0.000. Total n=287 in each dietary pattern; Pacific: n=126, NZE: n=161. Food groups with loadings <0.3 are not shown.

Supplementary Table 6.3. Habitual food group intake with the odds of being in enterotypes 1 or 3

Food group (g/day)	Enterotype 1 OR (95 % CI)	Enterotype 3 OR (95 % CI)
Full fat milk	1.00 [0.99, 1.00]	1.00 [0.99, 1.00]
Low fat milk	1.00 [0.99, 1.01]	1.01 [1.00, 1.01]
Milk alternatives	1.02 [1.00, 1.03]*	1.01 [1.00, 1.03]
Sweetened milk	1.01 [0.99, 1.02]	1.01 [1.00, 1.02]
Dairy yoghurt	1.00 [0.97, 1.02]	1.00 [0.98, 1.03]
High fat cheese	1.06 [0.97, 1.16]	1.05 [0.96, 1.16]
Low fat cheese	1.07 [0.97, 1.17]	1.10 [1.00, 1.21]
Apple, banana & oranges	0.99 [0.99, 1.00]	0.99 [0.98, 1.00]
Other fruit	0.99 [0.98, 1.01]	1.00 [0.98, 1.01]
Tomatoes	1.02 [0.96, 1.09]	1.02 [0.94, 1.10]
Yellow vegetables	1.04 [1.00, 1.08]	1.01 [0.97, 1.06]
Green vegetables	1.02 [1.00, 1.04]*	1.02 [1.00, 1.04]
Other non-starchy vegetables	1.04 [1.00, 1.08]	1.02 [0.98, 1.07]
Potatoes	0.99 [0.95, 1.02]	0.96 [0.92, 1.00]
Starchy vegetables	1.00 [0.98, 1.03]	0.99 [0.95, 1.02]
White bread	0.98 [0.95, 1.00]	1.01 [0.98, 1.04]
Discretionary breads	1.03 [0.93, 1.13]	1.04 [0.94, 1.16]
Crackers	0.96 [0.87, 1.06]	0.94 [0.83, 1.05]
Wholegrain products	1.01 [0.99, 1.03]	1.00 [0.98, 1.02]
Refined grains	0.97 [0.95, 1.00]*	0.96 [0.93, 0.99]*
Refined grains mixed	1.01 [0.98, 1.03]	1.00 [0.97, 1.04]
Oats	0.99 [0.95, 1.04]	1.00 [0.97, 1.04]
Sweetened cereals	1.06 [1.01, 1.12]*	1.06 [1.01, 1.12]*
Unsweetened cereals	0.97 [0.92, 1.01]	0.94 [0.87, 1.01]
Red meat	1.01 [0.98, 1.04]	0.99 [0.95, 1.03]
White meat	1.00 [0.98, 1.01]	1.00 [0.98, 1.02]
Processed meat	0.99 [0.95, 1.03]	0.99 [0.94, 1.03]
Fish and seafood	0.99 [0.96, 1.03]	1.01 [0.95, 1.06]
Egg and egg products	1.03 [0.99, 1.06]	1.06 [1.02, 1.09]**
Legumes and meat alternatives	1.07 [1.01, 1.13]*	1.06 [1.01, 1.12]*
Peanuts	1.26 [0.97, 1.64]	1.22 [0.93, 1.60]
Nuts and seeds	1.18 [1.01, 1.38]*	1.14 [0.97, 1.33]
Animals fats	0.99 [0.91, 1.09]	1.02 [0.92, 1.13]
Coconut products	1.11 [0.99, 1.26]	1.09 [0.96, 1.23]
Plant based fats	1.08 [0.98, 1.20]	1.05 [0.95, 1.17]
Margarine	1.10 [0.92, 1.31]	1.19 [0.97, 1.46]
Creamy based sauces and dressings	1.13 [0.94, 1.35]	1.21 [0.99, 1.48]
Savoury sauces and condiments	1.02 [0.97, 1.07]	1.01 [0.95, 1.06]
Sweet spreads	1.25 [0.93, 1.68]	1.12 [0.81, 1.54]
Cake and biscuits	1.00 [0.98, 1.03]	0.99 [0.96, 1.02]
Puddings and desserts	0.99 [0.96, 1.01]	0.98 [0.95, 1.01]
Sweet snacks	1.01 [0.98, 1.04]	1.00 [0.97, 1.04]
Savoury snacks	1.00 [0.97, 1.04]	0.82 [0.72, 0.94]**
Crumbed and deep fried	0.99 [0.96, 1.01]	1.00 [0.96, 1.04]
Fast food burgers	1.00 [0.99, 1.00]	1.00 [0.99, 1.01]
Fast food salad and sushi	1.02 [0.98, 1.05]	1.00 [0.96, 1.05]
Fruit and vegetable juice	1.01 [1.00, 1.03]	1.01 [0.99, 1.03]
Sugar sweetened beverages	0.99 [0.99, 1.00]*	0.99 [0.99, 1.00]*
Diet drinks	1.00 [1.00, 1.01]	1.00 [1.00, 1.01]
Tea	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]
Coffee	1.00 [1.00, 1.01]	1.00 [1.00, 1.01]
Alcoholic beverages	1.00 [1.00, 1.00]	1.00 [0.99, 1.00]
Water	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]
Added sugar to food and beverages	1.00 [0.94, 1.07]	0.97 [0.86, 1.09]
Soups and stocks	1.01 [0.99, 1.02]	1.00 [0.99, 1.02]

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. Enterotype 1: n=145, Enterotype 2: n=68, Enterotype 3: n=69. *P value <0.05, **P value ≤0.01, ***P value ≤0.001. Enterotype 2 is the reference category. Model: adjusted for age, ethnicity, energy intake (kJ/d), faecal water content (%), and NZDep2013 index.

Supplementary Table 6.4. Habitual nutrient intake with the odds of being in enterotypes 1 or 3

Nutrient	Enterotype 1 OR (95 % CI)	Enterotype 3 OR (95 % CI)
Energy (kJ/d)	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]
Protein (g/d)	1.01 [0.99, 1.04]	1.00 [0.97, 1.04]
Total fat (g/d)	1.03 [1.01, 1.04]**	1.01 [0.99, 1.03]
SFA (g/d)	1.004[1.00, 1.09]	1.01 [0.96, 1.07]
PUFA (g/d)	1.18 [1.06, 1.32]**	1.11 [0.98, 1.27]
MUFA (g/d)	1.06 [1.01, 1.12]*	1.02 [0.96, 1.08]
Cholesterol (g/d)	1.00 [1.00, 1.01]	1.01 [1.00, 1.01]*
CHO (g/d)	1.00 [0.99, 1.01]	1.00 [0.99, 1.01]
Sugar (g/d)	1.00 [0.99, 1.02]	1.00 [0.98, 1.02]
Starch (g/d)	1.00 [1.00, 1.01]	0.99 [0.98, 1.01]
Dietary Fibre (g/d)	1.09 [1.01, 1.17]*	1.06 [1.98, 1.15]
Thiamin (mg/d)	1.09 [0.51, 2.32]	1.04 [0.43, 2.53]
Riboflavin (mg/d)	1.28 [0.56, 2.91]	2.69 [0.96, 7.58]
Niacin (mg/d)	1.00 [0.94, 1.06]	1.00 [0.93, 1.08]
Niacin eq (mg/d)	1.01 [0.97, 1.06]	1.01 [0.95, 1.07]
Vitamin C (mg/d)	1.01 [1.00, 1.02]	1.01 [1.00, 1.03]
Vitamin E (mg/d)	1.23 [1.07, 1.42]**	1.15 [0.97, 1.36]
Vitamin B6 (mg/d)	1.30 [0.82, 2.04]	1.37 [0.83, 2.26]
Vitamin B12 (µg/d)	1.03 [0.73, 1.46]	1.20 [0.79, 1.84]
Total Folate (µg/d)	1.01 [1.00, 1.01]**	1.01 [1.00, 1.02]***
Total Vitamin A eq (µg/d)	1.00 [1.00, 1.01]**	1.00 [1.00, 1.01]**
Retinol (µg/d)	1.00 [1.00, 1.01]	1.01 [1.00, 1.01]*
Beta carotene (µg/d)	1.00 [1.00, 1.00]**	1.00 [1.00, 1.00]*
Sodium (g/d)	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]
Potassium (mg/d)	1.00 [1.00, 1.00]	1.00 [1.00, 1.01]
Magnesium (mg/d)	1.01 [1.00, 1.01]**	1.01 [1.00, 1.01]
Calcium (mg/d)	1.00 [1.00, 1.01]*	1.00 [1.00, 1.01]**
Phosphorous (mg/d)	1.00 [0.99, 1.01]	1.01 [1.00, 1.02]
Iron (mg/d)	1.13 [0.98, 1.30]	1.07 [0.89, 1.28]
Zinc (mg/d)	1.07 [0.87, 1.31]	0.93 [0.70, 1.24]
Selenium (µg/d)	1.02 [1.00, 1.04]	1.02 [0.99, 1.05]
Iodine (µg/d)	1.02 [1.00, 1.03] **	1.02 [1.01, 1.03]**
Caffeine (mg/d)	1.00 [1.00, 1.01]	1.00 [1.00, 1.01]
Alcohol (g/d)	1.00 [0.97, 1.04]	1.00 [0.96, 1.04]
Water (g/d)	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content [%] data. Enterotype 1: n=145, Enterotype 2: n=68, Enterotype 3: n=69. *P value <0.05, **P value ≤0.01, ***P value ≤0.001. Enterotype 2 is the reference category. Model : adjusted for age, ethnicity, energy intake (kJ/d), faecal water content (%), and NZDep2013 index. SFA = Saturated fat, PUFA = Polyunsaturated fat, MUFA = Monounsaturated fat, CHO = carbohydrate.

Supplementary Table 6.5. Habitual food group intake and the relative abundance of *Faecalibacterium prausnitzii*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.01]	p=0.392
Low fat milk	1.00 (0.995)	[0.99, 1.00]	p=0.012*
Milk alternatives	1.00	[1.00, 1.01]	p=0.242
Sweetened milk	1.00	[1.00, 1.00]	p=0.225
Dairy yoghurt	1.00	[1.00, 1.01]	p=0.999
High fat cheese	1.00	[0.96, 1.04]	p=0.938
Low fat cheese	1.01	[0.97, 1.04]	p=0.749
Apple, banana & oranges	1.00	[0.99, 1.01]	p=0.978
Other fruit	1.00	[0.99, 1.01]	p=0.869
Tomatoes	1.01	[0.96, 1.06]	p=0.706
Yellow vegetables	1.01	[0.98, 1.03]	p=0.652
Green vegetables	1.01	[1.00, 1.02]	p=0.010**
Other non-starchy vegetables	1.02	[1.00, 1.04]	p=0.079
Potatoes	0.99	[0.97, 1.01]	p=0.344
Starchy vegetables	1.00	[0.98, 1.02]	p=0.880
White bread	0.97	[0.95, 0.99]	p=0.002**
Discretionary breads	0.96	[0.91, 1.02]	p=0.184
Crackers	0.99	[0.93, 1.06]	p=0.820
Wholegrain products	1.01	[0.99, 1.02]	p=0.391
Refined grain	1.00	[0.98, 1.01]	p=0.606
Refined grains mixed	1.00	[0.98, 1.01]	p=0.620
Oats	0.99	[0.96, 1.02]	p=0.518
Sweetened cereals	1.01	[0.99, 1.03]	p=0.169
Unsweetened cereals	0.99	[0.95, 1.03]	p=0.580
Red meat	1.01	[0.98, 1.03]	p=0.634
White meat	1.00	[0.99, 1.01]	p=0.852
Processed meat	0.99	[0.96, 1.02]	p=0.420
Fish and seafood	1.00	[0.97, 1.03]	p=0.934
Egg and egg products	1.00	[0.98, 1.02]	p=0.842
Legumes and meat alternatives	1.01	[1.00, 1.02]	p=0.177
Peanuts	1.04	[0.97, 1.12]	p=0.275
Nuts and seeds	1.06	[1.02, 1.10]	p=0.005**
Animal fats	0.99	[0.93, 1.05]	p=0.658
Coconut products	1.03	[0.99, 1.08]	p=0.196
Plant based fats	1.02	[0.99, 1.06]	p=0.234
Margarine	1.01	[0.91, 1.12]	p=0.827
Creamy based sauces and dressings	0.96	[0.89, 1.05]	p=0.380
Savoury sauces and condiments	1.00	[0.97, 1.03]	p=0.819
Sweet spreads	0.98	[0.89, 1.08]	p=0.659
Cake and biscuits	1.00	[0.98, 1.02]	p=0.956
Puddings and desserts	1.00	[0.99, 1.02]	p=0.662
Sweet snacks	1.01	[0.99, 1.04]	p=0.203
Savoury snacks	1.03	[0.99, 1.08]	p=0.153
Crumbed and deep fried	1.00	[0.98, 1.03]	p=0.870
Fast food burgers	1.00	[0.99, 1.00]	p=0.100
Fast food salad and sushi	0.99	[0.97, 1.02]	p=0.618
Fruit and vegetable juice	1.01	[1.00, 1.02]	p=0.199
Other beverages	1.00	[1.00, 1.00]	p=0.576
Diet drinks	1.00	[1.00, 1.00]	p=0.751
Tea	1.00	[1.00, 1.00]	p=0.408
Coffee	1.00	[1.00, 1.00]	p=0.210
Alcoholic beverages	1.00	[1.00, 1.00]	p=0.398
Water	1.00	[1.00, 1.00]	p=0.556
Added sugar to food and beverages	0.97	[0.91, 1.02]	p=0.220
Soups and stocks	1.00	[0.99, 1.01]	p=0.766

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups in g/day. Low relative abundance is the reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. Low relative abundance: n = 140, high relative abundance: n = 142 (cut point is the median)

Supplementary Table 6.6. Habitual food group intake of Pacific and NZ European women stratified by enterotypes

Food group (g/day)	Enterotype		
	1	**2**	**3**
	n=146	*n*=70	*n*=70
Full fat milk	18.7 [0.0, 64.2]^	32.6 [10.5, 60.3]+	10.9 [0.0, 59.5]
Low fat milk	3.8 [0.0, 20.9]	0.0 [0.0, 8.1]+	8.9 [0.0, 91.6]~
Milk alternatives	0.0 [0.0, 30.0]^	0.0 [0.0, 0.0]+	0.0 [0.0, 31.0]
Sweetened milk	4.4 [3.8, 12.2]	5.7 [4.1, 11.7]	4.6 [4.0, 15.7]
Dairy yoghurt	3.9 [2.7, 8.9]^	2.9 [2.2, 4.6]+	4.6 [3.3, 13.2]~
High fat cheese	5.9 [3.6, 10.0]^	3.9 [2.9, 5.3]+	6.1 [4.3, 13.1]
Low fat cheese	4.9 [3.3, 7.8]^	3.2 [0.0, 4.3]+	6.5 [4.6, 12.0]~
Apple, banana & oranges	29.6 [15.6, 53.0]	30.2 [13.8, 51.2]	29.3 [13.7, 62.6]
Other fruit	32.8 [21.4, 50.1]	29.3 [20.7, 47.6]	40.6 [21.9, 56.0]
Tomatoes	29.8 [20.7, 34.5]^	20.0 [18.0, 21.9]+	32.9 [29.9, 36.6]~
Yellow vegetables	16.4 [10.4, 25.5]^	10.1 [7.5, 14.5]+	16.6 [13.0, 24.8]
Green vegetables	30.0 [16.3, 52.6]^	13.6 [8.6, 20.7]+	36.6 [23.9, 50.8]
Other non-starchy vegetables	22.2 [12.9, 34.2]^	12.9 [8.5, 18.6]+	24.8 [18.3, 35.3]
Potatoes	37.8 [32.6, 42.9]	36.8 [33.1, 43.3]	36.9 [32.9, 42.9]
Starchy vegetables	19.3 [15.2, 27.4]	21.3 [17.0, 29.3]+	18.4 [13.9, 23.5]
White bread	8.5 [4.8, 14.1]^	12.0 [5.1, 19.0]	7.9 [5.0, 29.6]
Discretionary breads	5.1 [4.2, 7.5]^	3.8 [3.2, 5.3]+	5.2 [4.3, 8.7]
Crackers	1.5 [1.4, 2.9]	1.4 [1.4, 2.5]	1.4 [1.3, 2.4]~
Wholegrain products	43.0 [30.7, 59.0]^	30.7 [27.0, 45.3]+	39.3 [30.8, 56.0]
Refined grain	15.7 [10.2, 24.8]	18.1 [10.9, 33.7]+	12.4 [8.8, 18.6]
Refined grains mixed	12.5 [0.0, 17.4]^	14.6 [10.3, 19.2]+	9.3 [0.0, 14.0]~
Oats	0.8 [0.7, 1.2]	0.8 [0.7, 0.9]	0.8 [0.7, 0.9]
Sweetened cereals	9.2 [8.7, 12.9]^	9.0 [0.0, 9.5]	9.1 [0.0, 11.7]
Unsweetened cereals	2.6 [0.0, 2.9]^	3.1 [2.8, 7.4]+	0.0 [0.0, 2.6]~
Red meat	20.6 [14.5, 24.7]^	15.6 [11.5, 20.4]+	20.3 [15.4, 25.6]
White meat	36.6 [27.1, 49.1]	38.6 [30.6, 51.4]	36.8 [25.1, 48.1]
Processed meat	11.2 [7.4, 16.7]	9.4 [7.5, 14.5]	10.1 [6.8, 12.7]
Fish and seafood	10.1 [6.9, 14.4]^	11.5 [8.7, 15.6]+	8.1 [6.2, 11.5]~

Food group (g/day)	Enterotype 1	2	3
Egg and egg products	25.4 [18.4, 33.9]^	17.3 [14.7, 23.6]+	31.9 [23.1, 44.4]~
Legumes and meat alternatives	9.6 [5.6, 19.0]^	3.0 [0.0, 6.0]+	11.5 [7.4, 26.5]
Peanuts	0.9 [0.8, 2.7]^	0.8 [0.5, 1.0]+	0.9 [0.9, 2.0]
Nuts and seeds	1.8 [1.3, 4.9]^	1.2 [0.0, 1.4]+	2.1 [1.3, 6.2]
Animals fats	4.6 [2.9, 7.0]	4.1 [3.0, 6.6]	3.8 [2.8, 7.0]
Coconut products	2.8 [2.4, 4.7]	2.8 [2.2, 3.1]	2.4 [2.2, 3.2]~
Plant based fats	2.6 [1.2, 6.8]^	1.1 [0.7, 2.3]+	3.0 [1.3, 7.5]
Margarine	1.7 [0.0, 2.6]	1.8 [1.7, 2.4]	1.7 [0.0, 2.7]
Creamy based sauces and dressings	2.0 [1.5, 3.3]^	1.5 [1.2, 2.1]+	2.4 [1.5, 3.8]
Savoury sauces and condiments	16.4 [10.0, 22.3]^	9.3 [5.7, 15.5]+	16.8 [13.0, 25.2]
Sweet spreads	1.3 [0.9, 2.1]^	1.1 [0.0, 1.7]	1.1 [0.9, 1.6]
Cake and biscuits	24.3 [18.9, 34.6]	23.9 [19.5, 31.6]	22.0 [18.1, 30.5]
Puddings and desserts	8.9 [7.0, 14.8]	9.7 [8.0, 15.1]	11.5 [7.9, 17.8]~
Sweet snacks	12.2 [7.6, 21.4]^	8.9 [5.5, 15.3]+	12.0 [7.6, 21.5]
Savoury snacks	3.1 [1.9, 6.1]	2.7 [1.9, 8.4]+	2.1 [1.6, 3.4]~
Crumbed and deep fried	14.2 [10.7, 21.6]^	20.9 [14.6, 33.9]+	12.1 [10.1, 18.7]
Fast food burgers	66.3 [41.5, 108.1]^	122.5 [80.4, 155.5]+	46.1 [38.6, 70.0]~
Fast food salad and sushi	9.8 [8.6, 14.7]	10.3 [9.1, 15.4]	9.4 [8.8, 12.5]
Fruit and vegetable juice	4.5 [3.0, 9.5]	4.7 [3.7, 10.9]+	3.7 [2.6, 6.6]
Sugar sweetened beverages	226.4 [144.6, 272.8]^	332.6 [252.6, 378.4]+	172.8 [91.1, 227.5]~
Diet drinks	7.4 [0.0, 10.8]	8.2 [0.0, 9.7]+	6.6 [0.0, 8.5]
Tea	41.3 [13.4, 172.2]^	14.7 [11.6, 46.3]+	64.8 [13.1, 259.4]
Coffee	4.1 [1.5, 33.8]^	1.5 [1.3, 3.4]+	5.0 [1.6, 85.9]
Alcoholic beverages	19.0 [8.0, 35.3]^	10.2 [0.0, 23.8]+	19.6 [11.2, 37.8]
Water	766.4 [415.4, 1284.3]	656.8 [344.8, 1143.5]+	867.6 [621.3, 1465.5]
Added sugar to food and beverages	0.5 [0.0, 1.5]^	1.1 [0.3, 3.7]+	0.3 [0.0, 1.0]~
Soups and stocks	28.8 [0.0, 40.8]	22.5 [0.0, 33.5]	31.4 [0.0, 46.3]

All values are reported as medians [25th, 75th centiles]. Total *n* = 286; Pacific *n* = 125 and NZE *n* = 161. Mann Whitney statistical test used to identify a significant difference ($p<0.05$). ^Statistically significant difference between enterotype 1 and 2. ~Statistically significant difference between enterotype 1 and 3. +Statistically significant difference between enterotype 2 and 3.

Supplementary Table 6.7. Habitual food group intake and the presence or absence of *Lactobacillus ruminis*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.01]	p=0.961
Low fat milk	0.99	[0.99, 1.00]	p=0.173
Milk alternatives	1.00	[0.99, 1.01]	p=0.670
Sweetened milk	1.00	[0.99, 1.01]	p=0.743
Dairy yoghurt	0.94	[0.87, 1.02]	p=0.114
High fat cheese	0.86	[0.76, 0.98]	p=0.021*
Low fat cheese	0.91	[0.81, 1.03]	p=0.140
Apple, banana & oranges	1.00	[0.99, 1.01]	p=0.653
Other fruit	1.01	[0.99, 1.02]	p=0.458
Tomatoes	1.02	[0.96, 1.10]	p=0.498
Yellow vegetables	0.99	[0.95, 1.03]	p=0.471
Green vegetables	1.00	[0.98, 1.02]	p=0.882
Other non-starchy vegetables	0.98	[0.94, 1.03]	p=0.444
Potatoes	1.02	[0.98, 1.06]	p=0.379
Starchy vegetables	1.02	[0.99, 1.05]	p=0.183
White bread	0.99	[0.97, 1.02]	p=0.613
Discretionary breads	0.99	[0.88, 1.10]	p=0.806
Crackers	0.79	[0.62, 1.01]	p=0.055
Wholegrain products	1.00	[0.99, 1.02]	p=0.680
Refined grain	1.00	[0.98, 1.01]	p=0.703
Refined grains mixed	0.98	[0.96, 1.01]	p=0.220
Oats	0.97	[0.89, 1.06]	p=0.471
Sweetened cereals	1.00	[0.97, 1.04]	p=0.806
Unsweetened cereals	0.99	[0.94, 1.04]	p=0.716
Red meat	0.98	[0.95, 1.02]	p=0.291
White meat	1.00	[0.98, 1.01]	p=0.643
Processed meat	1.01	[0.97, 1.06]	p=0.679
Fish and seafood	1.03	[0.99, 1.07]	p=0.202
Egg and egg products	0.99	[0.96, 1.02]	p=0.566
Legumes and meat alternatives	1.02	[1.00, 1.05]	p=0.049*
Peanuts	1.05	[0.89, 1.24]	p=0.579
Nuts and seeds	0.78	[0.62, 0.98]	p=0.032*
Animal fats	0.90	[0.79, 1.01]	p=0.080
Coconut products	1.01	[0.92,1.10]	p=0.836
Plant based fats	1.02	[0.94, 1.10]	p=0.686
Margarine	1.20	[1.03, 1.41]	p=0.023*
Creamy based sauces and dressings	0.93	[0.79, 1.11]	p=0.426
Savoury sauces and condiments	0.98	[0.93, 1.03]	p=0.450
Sweet spreads	1.10	[0.93, 1.31]	p=0.252
Cake and biscuits	1.00	[0.97, 1.03]	p=0.800
Puddings and desserts	1.02	[0.99, 1.05]	p=0.226
Sweet snacks	0.98	[0.95, 1.01]	p=0.211
Savoury snacks	0.99	[0.95, 1.03]	p=0.532
Crumbed and deep fried	1.00	[0.98, 1.03]	p=0.842
Fast food burgers	1.00	[0.99, 1.01]	p=0.715
Fast food salad and sushi	1.01	[0.98, 1.05]	p=0.473
Fruit and vegetable juice	1.00	[0.99, 1.01]	p=0.634
Sugar sweetened beverages	1.01	[1.00, 1.01]	p=0.025*
Diet drinks	1.00	[0.99, 1.01]	p=0.791
Tea	1.00	[1.00, 1.00]	p=0.701
Coffee	1.00	[1.00, 1.01]	p=0.692
Alcoholic beverages	1.00	[0.99, 1.01]	p=0.788
Water	1.00	[1.00, 1.00]	p=0.224
Added sugar to food and beverages	0.97	[0.92, 1.04]	p=0.386
Soups and stocks	1.00	[0.98, 1.02]	p=0.849

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups in g/day. Model: Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *Lactobacillus ruminis*: present n=63, absent n =220

Supplementary Table 6.8. Habitual food group intake and the presence or absence of *Bifidobacterium bifidum*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.00]	p=0.767
Low fat milk	1.00	[1.00, 1.01]	p=0.045*
Milk alternatives	0.99	[0.98, 1.00]	p=0.021*
Sweetened milk	1.00	[0.99, 1.01]	p=0.646
Dairy yoghurt	1.00	[0.98, 1.01]	p=0.625
High fat cheese	0.97	[0.92, 1.02]	p=0.211
Low fat cheese	0.97	[0.92, 1.02]	p=0.215
Apple, banana & oranges	1.00	[0.99, 1.00]	p=0.344
Other fruit	0.99	[0.98, 1.00]	p=0.222
Tomatoes	1.00	[0.95, 1.05]	p=0.867
Yellow vegetables	0.99	[0.96, 1.01]	p=0.277
Green vegetables	0.99	[0.98, 1.01]	p=0.298
Other non-starchy vegetables	0.97	[0.95, 1.00]	p=0.040*
Potatoes	1.01	[0.99, 1.04]	p=0.420
Starchy vegetables	1.02	[1.00, 1.05]	p=0.078
White bread	1.00	[0.98, 1.02]	p=0.758
Discretionary breads	1.05	[1.00, 1.11]	p=0.077
Crackers	1.03	[0.95, 1.11]	p=0.490
Wholegrain products	0.99	[0.98, 1.01]	p=0.343
Refined grain	1.01	[0.99, 1.04]	p=0.175
Refined grains mixed	1.00	[0.98, 1.02]	p=0.943
Oats	1.00	[0.99, 1.01]	p=0.785
Sweetened cereals	0.98	[0.95, 1.01]	p=0.204
Unsweetened cereals	1.00	[0.95, 1.04]	p=0.831
Red meat	0.99	[0.97, 1.02]	p=0.441
White meat	1.01	[0.99, 1.02]	p=0.302
Processed meat	1.00	[0.96, 1.03]	p=0.813
Fish and seafood	1.01	[0.98, 1.05]	p=0.515
Egg and egg products	0.99	[0.97, 1.01]	p=0.375
Legumes and meat alternatives	1.00	[0.98, 1.01]	p=0.762
Peanuts	0.83	[0.70, 0.98]	p=0.024*
Nuts and seeds	0.87	[0.80, 0.96]	p=0.004**
Animal fats	0.97	[0.90, 1.04]	p=0.372
Coconut products	0.97	[0.91, 1.03]	p=0.268
Plant based fats	0.93	[0.87, 0.99]	p=0.026*
Margarine	1.05	[0.93, 1.18]	p=0.440
Creamy based sauces and dressings	1.05	[0.96, 1.13]	p=0.289
Savoury sauces and condiments	0.98	[0.95, 1.02]	p=0.310
Sweet spreads	0.88	[0.74, 1.05]	p=0.160
Cake and biscuits	1.01	[0.99, 1.03]	p=0.389
Puddings and desserts	1.01	[0.99, 1.03]	p=0.340
Sweet snacks	0.99	[0.96, 1.01]	p=0.249
Savoury snacks	0.98	[0.94, 1.02]	p=0.254
Crumbed and deep fried	1.00	[0.98, 1.03]	p=0.854
Fast food burgers	1.00	[1.00, 1.01]	p=0.320
Fast food salad and sushi	1.02	[0.99, 1.05]	p=0.191
Fruit and vegetable juice	1.00	[0.99, 1.01]	p=0.357
Sugar sweetened beverages	1.00	[1.00, 1.01]	p=0.081
Diet drinks	1.00	[1.00, 1.00]	p=0.663
Tea	1.00	[1.00, 1.00]	p=0.923
Coffee	1.00	[0.99, 1.00]	p=0.066
Alcoholic beverages	1.00	[0.99, 1.00]	p=0.414
Water	1.00	[1.00, 1.00]	p=0.957
Added sugar to food and beverages	0.98	[0.93, 1.04]	p=0.566
Soups and stocks	0.99	[0.98, 1.00]	p=0.190

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. Food groups all in g/day.
Model: Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content.
Bifidobacterium bifidum: present n=116, absent n =167

Supplementary Table 6.9. Habitual food group intake and the presence or absence of *Bifidobacterium adolescentis*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.01]	p=0.625
Low fat milk	1.00	[1.00, 1.00]	p=0.550
Milk alternatives	1.00	[0.99, 1.01]	p=0.922
Sweetened milk	0.99	[0.99, 1.00]	p=0.027*
Dairy yoghurt	1.00	[0.98, 1.01]	p=0.394
High fat cheese	1.00	[0.96, 1.03]	p=0.795
Low fat cheese	0.97	[0.94, 1.01]	p=0.163
Apple, banana & oranges	1.00	[1.00, 1.01]	p=0.386
Other fruit	1.00	[0.99, 1.02]	p=0.480
Tomatoes	1.01	[0.96, 1.07]	p=0.748
Yellow vegetables	1.03	[1.00, 1.06]	p=0.068
Green vegetables	1.00	[0.99, 1.01]	p=0.770
Other non-starchy vegetables	1.01	[0.98, 1.03]	p=0.605
Potatoes	1.05	[1.02, 1.08]	p=0.001***
Starchy vegetables	1.04	[1.00, 1.08]	p=0.042*
White bread	0.99	[0.97, 1.01]	p=0.350
Discretionary breads	1.06	[0.99, 1.14]	p=0.120
Crackers	1.00	[0.93, 1.09]	p=0.926
Wholegrain products	1.01	[0.99, 1.03]	p=0.227
Refined grain	1.04	[1.01, 1.08]	p=0.012*
Refined grains mixed	0.99	[0.96, 1.01]	p=0.153
Oats	1.00	[0.98, 1.01]	p=0.415
Sweetened cereals	1.00	[0.98, 1.02]	p=0.975
Unsweetened cereals	1.02	[0.96, 1.08]	p=0.494
Red meat	0.97	[0.95, 1.00]	p=0.059
White meat	1.00	[0.99, 1.02]	p=0.667
Processed meat	0.98	[0.95, 1.01]	p=0.250
Fish and seafood	0.99	[0.95, 1.03]	p=0.593
Egg and egg products	0.99	[0.97, 1.01]	p=0.130
Legumes and meat alternatives	1.02	[1.00, 1.04]	p=0.099
Peanuts	0.97	[0.90, 1.04]	p=0.352
Nuts and seeds	0.99	[0.95, 1.03]	p=0.645
Animal fats	0.94	[0.88, 1.00]	p=0.062
Coconut products	0.98	[0.94,1.03]	p=0.493
Plant based fats	0.99	[0.95, 1.02]	p=0.458
Margarine	1.09	[0.95, 1.25]	p=0.198
Creamy based sauces and dressings	1.00	[0.92, 1.10]	p=0.946
Savoury sauces and condiments	1.01	[0.98, 1.05]	p=0.488
Sweet spreads	1.02	[0.91, 1.13]	p=0.788
Cake and biscuits	1.02	[1.00, 1.05]	p=0.098
Puddings and desserts	1.01	[0.98, 1.04]	p=0.457
Sweet snacks	1.01	[0.98, 1.04]	p=0.447
Savoury snacks	0.98	[0.93, 1.02]	p=0.334
Crumbed and deep fried	1.00	[0.97, 1.04]	p=0.848
Fast food burgers	1.00	[0.99, 1.01]	p=0.664
Fast food salad and sushi	1.04	[0.99, 1.08]	p=0.093
Fruit and vegetable juice	1.00	[0.98, 1.02]	p=0.704
Sugar sweetened beverages	1.00	[1.00, 1.01]	p=0.509
Diet drinks	1.00	[1.00, 1.00]	p=0.653
Tea	1.00	[1.00, 1.00]	p=0.718
Coffee	1.00	[1.00, 1.00]	p=0.400
Alcoholic beverages	1.00	[1.00, 1.00]	p=0.295
Water	1.00	[1.00, 1.00]	p=0.461
Added sugar to food and beverages	1.04	[0.94, 1.14]	p=0.492
Soups and stocks	0.99	[0.98, 1.00]	p=0.077

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups are in g/day. Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *Bifidobacterium adolescentis*: present n=220, absent n =63

Supplementary Table 6.10. Habitual food group intake and the relative abundance of *Subdoligranulum species*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.00]	p=0.558
Low fat milk	1.00	[1.10, 1.01]	p=0.293
Milk alternatives	1.00	[0.99, 1.00]	p=0.214
Sweetened milk	1.00	[0.99, 1.00]	p=0.428
Dairy yoghurt	1.01	[0.99, 1.02]	p=0.372
High fat cheese	0.98	[0.95, 1.02]	p=0.266
Low fat cheese	1.01	[0.97, 1.05]	p=0.574
Apple, banana & oranges	1.00	[1.00, 1.01]	p=0.543
Other fruit	1.00	[0.99, 1.01]	p=0.593
Tomatoes	1.03	[0.98, 1.08]	p=0.279
Yellow vegetables	1.00	[0.98, 1.02]	p=0.948
Green vegetables	1.01	[1.00, 1.02]	p=0.087
Other non-starchy vegetables	1.02	[1.00, 1.04]	p=0.135
Potatoes	0.99	[0.97, 1.01]	p=0.380
Starchy vegetables	1.01	[0.99, 1.03]	p=0.450
White bread	1.01	[0.99, 1.03]	p=0.251
Discretionary breads	1.03	[0.97, 1.08]	p=0.342
Crackers	0.98	[0.92, 1.05]	p=0.632
Wholegrain products	1.01	[1.00, 1.03]	p=0.035*
Refined grain	0.98	[0.96, 1.00]	p=0.096
Refined grains mixed	1.00	[0.98, 1.02]	p=0.847
Oats	1.00	[0.99, 1.01]	p=0.463
Sweetened cereals	1.00	[0.98, 1.02]	p=0.983
Unsweetened cereals	1.00	[0.96, 1.04]	p=0.964
Red meat	0.99	[0.96, 1.01]	p=0.234
White meat	1.01	[1.00, 1.02]	p=0.201
Processed meat	1.00	[0.97, 1.03]	p=0.844
Fish and seafood	0.99	[0.96, 1.02]	p=0.482
Egg and egg products	1.02	[1.00, 1.04]	p=0.046*
Legumes and meat alternatives	1.01	[0.99, 1.02]	p=0.275
Peanuts	1.03	[0.96, 1.11]	p=0.467
Nuts and seeds	1.01	[0.97, 1.05]	p=0.559
Animal fats	1.01	[0.95, 1.07]	p=0.795
Coconut products	1.01	[0.96, 1.05]	p=0.805
Plant based fats	1.01	[0.97, 1.05]	p=0.678
Margarine	1.04	[0.94, 1.16]	p=0.446
Creamy based sauces and dressings	1.09	[0.98, 1.20]	p=0.107
Savoury sauces and condiments	1.00	[0.97, 1.03]	p=0.896
Sweet spreads	1.00	[0.90, 1.10]	p=0.930
Cake and biscuits	1.00	[0.98, 1.02]	p=0.889
Puddings and desserts	0.99	[0.98, 1.01]	p=0.529
Sweet snacks	0.98	[0.95, 1.00]	p=0.037*
Savoury snacks	0.98	[0.95, 1.02]	p=0.413
Crumbed and deep fried	1.00	[0.97, 1.02]	p=0.785
Fast food burgers	1.00	[0.99, 1.01]	p=0.710
Fast food salad and sushi	1.00	[0.97, 1.03]	p=0.944
Fruit and vegetable juice	1.01	[1.00, 1.02]	p=0.120
Sugar sweetened beverages	1.00 (0.996)	[0.99, 1.00]	p=0.020*
Diet drinks	1.00	[1.00, 1.00]	p=0.897
Tea	1.00	[1.00, 1.00]	p=0.961
Coffee	1.00	[1.00, 1.00]	p=0.931
Alcoholic beverages	1.00	[1.00, 1.00]	p=0.985
Water	1.00	[1.00, 1.00]	p=0.551
Added sugar to food and beverages	1.01	[0.96, 1.07]	p=0.751
Soups and stocks	1.00	[0.99, 1.01]	p=0.494

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups are in g/day.
Low relative abundance is the reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. Low relative abundance: n = 140, high relative abundance: n = 142 (cut point is the median)

Supplementary Table 6.11. Habitual food group intake and the presence or absence of *Methanobrevibacter smithii*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.00]	p=0.660
Low fat milk	1.00	[1.00, 1.00]	p=0.758
Milk alternatives	1.00	[1.00, 1.01]	p=0.565
Sweetened milk	1.00	[0.99, 1.01]	p=0.693
Dairy yoghurt	1.00	[0.98, 1.01]	p=0.557
High fat cheese	0.99	[0.95, 1.02]	p=0.451
Low fat cheese	0.98	[0.94, 1.02]	p=0.328
Apple, banana & oranges	1.00	[1.00, 1.01]	p=0.440
Other fruit	0.99	[0.98, 1.00]	p=0.030*
Tomatoes	1.00	[0.95, 1.04]	p=0.904
Yellow vegetables	1.01	[0.99, 1.03]	p=0.578
Green vegetables	1.00	[0.99, 1.01]	p=0.552
Other non-starchy vegetables	1.00	[0.98, 1.02]	p=0.969
Potatoes	0.99	[0.97, 1.02]	p=0.461
Starchy vegetables	1.01	[0.99, 1.04]	p=0.206
White bread	0.98	[0.96, 1.00]	p=0.081
Discretionary breads	1.01	[0.97, 1.07]	p=0.572
Crackers	1.00	[0.93, 1.07]	p=0.995
Wholegrain products	1.01	[1.00, 1.02]	p=0.135
Refined grain	0.98	[0.97, 1.00]	p=0.119
Refined grains mixed	0.99	[0.97, 1.01]	p=0.272
Oats	0.99	[0.97, 1.02]	p=0.515
Sweetened cereals	1.00	[0.98, 1.01]	p=0.664
Unsweetened cereals	0.97	[0.93, 1.01]	p=0.173
Red meat	1.00	[0.98, 1.03]	p=0.715
White meat	1.00	[0.99, 1.01]	p=0.916
Processed meat	1.01	[0.98, 1.04]	p=0.482
Fish and seafood	0.99	[0.96, 1.02]	p=0.575
Egg and egg products	1.01	[0.99, 1.02]	p=0.573
Legumes and meat alternatives	1.01	[0.99, 1.02]	p=0.485
Peanuts	1.00	[0.93, 1.08]	p=0.995
Nuts and seeds	0.99	[0.96, 1.03]	p=0.772
Animal fats	1.01	[0.95, 1.07]	p=0.790
Coconut products	0.99	[0.95, 1.04]	p=0.769
Plant based fats	1.00	[0.96, 1.03]	p=0.857
Margarine	1.03	[0.93, 1.14]	p=0.589
Creamy based sauces and dressings	0.97	[0.89, 1.05]	p=0.436
Savoury sauces and condiments	0.98	[0.95, 1.01]	p=0.151
Sweet spreads	1.01	[0.92, 1.11]	p=0.899
Cake and biscuits	1.00	[0.98, 1.02]	p=0.815
Puddings and desserts	1.00	[0.98, 1.02]	p=0.666
Sweet snacks	0.97	[0.94, 0.99]	p=0.011*
Savoury snacks	0.99	[0.95, 1.03]	p=0.549
Crumbed and deep fried	1.01	[0.98, 1.03]	p=0.559
Fast food burgers	1.00	[0.99, 1.01]	p=0.748
Fast food salad and sushi	1.01	[0.99, 1.04]	p=0.288
Fruit and vegetable juice	1.00	[0.99, 1.01]	p=0.646
Sugar sweetened beverages	1.00	[1.00, 1.00]	p=0.408
Diet drinks	1.00	[1.00, 1.00]	p=0.698
Tea	1.00	[1.00, 1.00]	p=0.512
Coffee	1.00	[1.00, 1.00]	p=0.844
Alcoholic beverages	1.00	[1.00, 1.00]	p=0.705
Water	1.00	[1.00, 1.00]	p=0.627
Added sugar to food and beverages	1.02	[0.97, 1.08]	p=0.411
Soups and stocks	0.99	[0.98, 1.00]	p=0.153

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups in g/day. Model: Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *Methanobrevibacter smithii:* present n=114, absent n =169

Supplementary Table 6.12. Habitual food group intake and the presence or absence of *Akkermansia muciniphila*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.01]	p=0.531
Low fat milk	1.00	[1.00, 1.01]	p=0.085
Milk alternatives	1.00	[1.00, 1.01]	p=0.360
Sweetened milk	1.00	[1.00, 1.01]	p=0.453
Dairy yoghurt	1.01	[0.99, 1.03]	p=0.237
High fat cheese	0.99	[0.96, 1.03]	p=0.646
Low fat cheese	1.00	[0.97, 1.04]	p=0.840
Apple, banana & oranges	1.00	[0.99, 1.01]	p=0.975
Other fruit	1.00	[0.99, 1.01]	p=0.473
Tomatoes	1.05	[1.00, 1.10]	p=0.059
Yellow vegetables	1.00	[0.98, 1.03]	p=0.729
Green vegetables	1.00	[0.99, 1.01]	p=0.457
Other non-starchy vegetables	1.03	[1.01, 1.06]	p=0.018*
Potatoes	0.99	[0.96, 1.01]	p=0.290
Starchy vegetables	0.99	[0.96, 1.01]	p=0.221
White bread	1.00	[0.98, 1.02]	p=0.898
Discretionary breads	1.00	[0.95, 1.05]	p=0.917
Crackers	0.96	[0.90, 1.03]	p=0.305
Wholegrain products	1.00	[0.98, 1.01]	p=0.480
Refined grain	0.99	[0.97, 1.01]	p=0.183
Refined grains mixed	0.98	[0.96, 1.01]	p=0.142
Oats	1.00	[0.99, 1.01]	p=0.744
Sweetened cereals	1.02	[0.99, 1.04]	p=0.197
Unsweetened cereals	1.06	[1.01, 1.12]	p=0.028*
Red meat	0.98	[0.96, 1.01]	p=0.219
White meat	0.99	[0.98, 1.01]	p=0.291
Processed meat	0.98	[0.95, 1.01]	p=0.150
Fish and seafood	0.99	[0.96, 1.02]	p=0.420
Egg and egg products	1.01	[0.99, 1.03]	p=0.380
Legumes and meat alternatives	1.00	[0.99, 1.02]	p=0.750
Peanuts	1.05	[0.96, 1.15]	p=0.284
Nuts and seeds	1.03	[0.99, 1.08]	p=0.170
Animal fats	0.98	[0.92, 1.04]	p=0.438
Coconut products	1.01	[0.96, 1.06]	p=0.729
Plant based fats	1.03	[0.99, 1.08]	p=0.194
Margarine	1.09	[0.97, 1.22]	p=0.163
Creamy based sauces and dressings	1.00	[0.92, 1.08]	p=0.902
Savoury sauces and condiments	0.98	[0.95, 1.01]	p=0.272
Sweet spreads	0.94	[0.86, 1.03]	p=0.171
Cake and biscuits	1.02	[1.00, 1.04]	p=0.127
Puddings and desserts	0.98	[0.96, 1.00]	p=0.095
Sweet snacks	1.00	[0.98, 1.02]	p=0.858
Savoury snacks	1.00	[0.96, 1.03]	p=0.842
Crumbed and deep fried	0.99	[0.97, 1.01]	p=0.385
Fast food burgers	1.00	[0.99, 1.00]	p=0.562
Fast food salad and sushi	0.98	[0.96, 1.01]	p=0.177
Fruit and vegetable juice	1.01	[1.00, 1.02]	p=0.329
Sugar sweetened beverages	1.00	[0.99, 1.00]	p=0.160
Diet drinks	1.00	[1.00, 1.01]	p=0.254
Tea	1.00	[1.00, 1.00]	p=0.482
Coffee	1.00	[1.00, 1.00]	p=0.931
Alcoholic beverages	1.00	[1.00, 1.01]	p=0.145
Water	1.00	[1.00, 1.00]	p=0.174
Added sugar to food and beverages	0.99	[0.94, 1.05]	p=0.753
Soups and stocks	1.00	[0.99, 1.01]	p=0.728

Total n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups int g/day. Model: Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *Akkermansia muciniphila* present n=174, absent n =109

Supplementary Table 6.13. Habitual food group intake and the presence or absence of *Ruminococcus bromii*

Model	OR	95 % CI	P value
Full fat milk	1.00	[1.00, 1.01]	p=0.667
Low fat milk	1.00	[1.00, 1.01]	p=0.395
Milk alternatives	1.00	[1.00, 1.01]	p=0.998
Sweetened milk	1.00	[0.99, 1.00]	p=0.517
Dairy yoghurt	1.00	[0.99, 1.02]	p=0.661
High fat cheese	0.98	[0.94, 1.01]	p=0.233
Low fat cheese	1.01	[0.97, 1.05]	p=0.587
Apple, banana & oranges	1.00	[0.99, 1.01]	p=0.858
Other fruit	1.00	[0.99, 1.01]	p=0.884
Tomatoes	1.00	[0.95, 1.04]	p=0.848
Yellow vegetables	1.02	[1.00, 1.05]	p=0.102
Green vegetables	1.02	[1.01, 1.03]	p=0.003**
Other non-starchy vegetables	1.02	[0.99, 1.04]	p=0.162
Potatoes	0.98	[0.96, 1.01]	p=0.188
Starchy vegetables	1.01	[0.99, 1.04]	p=0.444
White bread	1.01	[0.99, 1.03]	p=0.430
Discretionary breads	1.00	[0.95, 1.05]	p=0.877
Crackers	0.96	[0.89, 1.03]	p=0.202
Wholegrain products	1.00	[0.99, 1.02]	p=0.803
Refined grain	1.00	[0.99, 1.02]	p=0.918
Refined grains mixed	1.00	[0.98, 1.02]	p=0.932
Oats	1.00	[0.99, 1.02]	p=0.695
Sweetened cereals	1.00	[0.99, 1.02]	p=0.641
Unsweetened cereals	1.00	[0.96, 1.04]	p=0.972
Red meat	1.01	[0.98, 1.03]	p=0.656
White meat	1.01	[1.00, 1.02]	p=0.136
Processed meat	0.99	[0.96, 1.02]	p=0.506
Fish and seafood	1.00	[0.97, 1.03]	p=0.937
Egg and egg products	1.01	[0.99, 1.03]	p=0.267
Legumes and meat alternatives	1.00	[0.99, 1.01]	p=0.944
Peanuts	1.06	[0.97, 1.16]	p=0.229
Nuts and seeds	1.03	[0.98, 1.08]	p=0.214
Animal fats	1.03	[0.97, 1.11]	p=0.328
Coconut products	1.05	[0.98, 1.11]	p=0.150
Plant based fats	1.03	[0.99, 1.07]	p=0.202
Margarine	1.14	[1.00, 1.30]	p=0.043*
Creamy based sauces and dressings	1.00	[0.92, 1.08]	p=0.928
Savoury sauces and condiments	0.99	[0.96, 1.02]	p=0.476
Sweet spreads	1.01	[0.92, 1.12]	p=0.805
Cake and biscuits	0.99	[0.97, 1.01]	p=0.287
Puddings and desserts	0.98	[0.96, 1.00]	p=0.066
Sweet snacks	0.99	[0.97, 1.01]	p=0.279
Savoury snacks	0.99	[0.96, 1.03]	p=0.681
Crumbed and deep fried	1.01	[0.98, 1.03]	p=0.670
Fast food burgers	1.00	[0.99, 1.00]	p=0.448
Fast food salad and sushi	1.01	[0.98, 1.04]	p=0.507
Fruit and vegetable juice	1.02	[1.00, 1.05]	p=0.086
Sugar sweetened beverages	1.00	[1.00, 1.00]	p=0.325
Diet drinks	1.00	[1.00, 1.00]	p=0.686
Tea	1.00	[1.00, 1.00]	p=0.380
Coffee	1.00	[1.00, 1.00]	p=0.665
Alcoholic beverages	1.00	[1.00, 1.00]	p=0.649
Water	1.00	[1.00, 1.00]	p=0.577
Added sugar to food and beverages	1.05	[0.98, 1.13]	p=0.164
Soups and stocks	1.00	[0.99, 1.01]	p=0.815

Total

n= 282; n=3 with missing deprivation data, n=1 with missing faecal water content (%) data. All food groups in g/day. Model: Absence of species reference category. Adjusted for NZDep2013, age, energy intake (kJ/day), ethnicity and faecal water content. *Ruminococcus bromii* present n=195, absent n = 88

7. Discussion

7.1. Introduction

The main objective of this thesis was to provide greater understanding of the association between habitual diet and the composition of the microbiota, in relation to body fat content and metabolic health markers, in a large, free-living population. This discussion opens with a summary of the study before bringing together the main results from each chapter and concluding remarks. The strengths and limitations will be discussed, and considerations of future directions will bring this thesis to a close.

7.2. Summary of the study

This thesis describes PhD research that utilised a cross-sectional study to explore habitual dietary intake in relation to body fat content, metabolic health markers, and the gut microbiota. The study included 287 Pacific and New Zealand European (NZE) women, who are known to have different metabolic disease risks (Pacific women have a high risk of obesity (70 % obesity) and NZE women have a moderate risk of obesity (30 % obesity) (1)), and body fat profiles (lean and obese). Data were collected between July 2016 and September 2017 from women aged between 18-45 years living in Auckland, NZ. Dietary intake was assessed with a 5-day estimated non-consecutive food record and a validated 220-item semi-quantitative New Zealand Women's FFQ (NZWFFQ) (2). These data were used to estimate habitual dietary intake using the NCI method (3,4). A comprehensive assessment of body composition and a range of metabolic health markers were assessed to explore metabolic disease risk. Metagenomic shotgun sequencing was conducted to explore the composition of the microbiota, including enterotype distributions. Data were used to characterise and explore habitual dietary intake in relation to body fat content, metabolic health markers, and the gut microbiota composition.

Multiple days of dietary assessment can reduce the likelihood of misreporting (5), as does prospectively recording intake rather than relying on memory-based assessment tools (6,7). The NCI modelling process specifically addresses intra-individual variation inherent in dietary data (3), and including the NZWFFQ information improves the precision of the estimation of episodically consumed foods and nutrients (8). The **first objective** of my PhD research was *to explore the relationship between habitual macronutrient intake in relation to body fat content*

and metabolic health markers in Pacific and NZE women. Higher dietary fibre intake was inversely associated with BF% for both Pacific and NZE women, highlighting that the dietary drivers of higher BF% are similar for these women despite different metabolic disease risks.

People choose to eat foods, not nutrients in isolation; hence, dietary pattern assessment enables assessment of the broader diet (and combinations of foods and nutrients consumed together) and how this relates to health outcomes (9–11). The **second objective** was *to characterise posteriori dietary patterns (derived from multiple days of dietary assessment) and to explore the associations with body fat content and metabolic health markers.* A total of four habitual dietary patterns were identified. There was no difference in BF% between Pacific and NZE women; however, higher adherence to dietary patterns characterised by core foods (the "colourful vegetable, plant protein, and dairy" and "fruit, starchy vegetables, and nuts" patterns) were inversely associated with BF%. In contrast, patterns characterised by more 'discretionary' foods ("sweet and fat rich carbohydrate") and less diversity of core foods ("animal meat and fat") were positively associated with BF% for both Pacific and NZE women. For both Pacific and NZE women, the impact of habitual food choice and subsequent risk of obesity is comparable and further suggests that habitual dietary intake could be contributing to NZ's high obesity rates (1,12).

Habitual dietary intake provides substrates for the microbiota (13,14) and, thus, has the potential to influence the functions of the microbiota which, in turn, could subsequently influence human physiology (i.e., indirect disease risk). However, the association between microbiota and habitual dietary intake remains unclear, notably in relation to metabolic health outcomes in healthy free-living populations. The first two objectives within this thesis were focused on thoroughly characterising good quality habitual dietary data (using robust and reliable methodological approaches) to explore the association with body fat content and metabolic health markers. Chapters 3 and 4 highlighted that habitual macronutrient intake and dietary patterns were associated with the same metabolic disease risks for both Pacific and NZE women, and body fat groups. Hence, the **third objective** was to explore *the characteristics of microbiota composition in relation to habitual diet (dietary patterns, foods and nutrients), body fat content and metabolic health markers* within one setting. Metagenomic shotgun sequencing enabled a thorough assessment of the composition of the

microbiota down to the species level. Thus, three enterotypes were characterised based on the differential abundance of particular bacterial species. The “colourful vegetable, plant protein, and dairy” pattern was positively associated with enterotypes 1 and 3 and negatively with enterotype 2. Therefore, habitual diet was associated with the composition of the microbiota. Additionally, stratifying the participants by the composition of their microbiota (i.e., enterotypes) highlighted metabolic disease risks. Furthermore, habitual food group intake was associated with the abundances of bacterial species which characterised the enterotypes.

7.3. Discussion of the main findings

The quality of dietary data is dependent on dietary assessment methodology (i.e., choice of dietary assessment tool, processing of dietary data), and analyses. Most studies exploring habitual dietary intake in relation to the microbiota are limited in that habitual diet is often only assessed with a FFQ (13–20) which, as discussed in the literature review, is prone to overestimation of intake (5–7). Utilising the high-quality dietary data (5DFR) and the validated NCI method to characterise habitual dietary intake (at the nutrient and food group level), strengthens the estimation of habitual dietary intake by reducing the intra-individual variation inherent in self-reported dietary intake (3). In addition, adjusting analyses for covariates that can either influence dietary intake (age, deprivation, energy intake) or the composition of the microbiota (gut transit time) is not routinely done, but improves the validity of the observations within this study.

Habitual diet has previously been associated with the composition of the microbiota (13,14); however, most studies have focused on the nutrient (13,16,20) or food group level (14,18,19) which, fail to consider the whole diet and, thus, delivery of combinations of substrates to the microbiota. A few recent studies have explored *a posteriori* dietary patterns in relation to the microbiota, but the association with adherence to the patterns and microbiota composition was not clear. For example, no associations were observed between dietary patterns and bacterial abundance (21,22) (likely due to small sample size of the study (22)), or the population was elderly (21,23) (which may not be translatable to younger populations due to the physiological changes in the gut that occur with aging). Large population studies in healthy

adults have recently associated aspects of dietary intake with the microbiota (14,18,19); however, due to the large sample size, dietary assessment was limited (e.g., quantitative intakes could not be established (18)) and sequencing was based on 16S rRNA gene amplicon sequencing. Furthermore, very few studies have circled back to associate habitual intake with microbiota and body composition, including multiple metabolic health markers (16).

The current study population was unique, it was a large sample of free-living disease free women, whom had different metabolic disease risk and body fat profiles (lean and obese). Additionally, the metagenomic sequencing employed for this population has enabled the characterisation of the microbiota down to the species level. Although the sample size is not as large as recent population-based studies, this study overcomes these constraints by the quality of dietary assessment and method of microbiota analyses employed, which combined, equate to a substantial sample size for such thorough analyses.

Characterisation of the faecal microbiotas of the study participants revealed three enterotypes. Enterotype 1 did not differ in prevalence between the groups. However, enterotype 2 was predominately found in Pacific women and enterotype 3 predominated in NZE women. There was no difference in BF% between (all) Pacific and NZE women; additionally, when stratifying the microbiota by BF% (at the phylum, family, genus, or species level) there were no distinct differences in microbiota composition between the low- and high- BF% groups. However, when stratifying the participants by enterotypes, there were clear differences in BF%, with women characterised by enterotype 2 having higher adiposity (BF%, visceral adipose tissue (VAT) and android fat % (AF%)). Previous studies have found that Pacific people have higher abdominal fat in comparison to Europeans (24); however, enterotype 1 was characterised by both ethnicities and had significantly lower total and abdominal fat in comparison to enterotype 2. Suggesting these differences in adiposity are not related to ethnic differences, they are associated with habitual dietary intake. Unlike other studies, habitual dietary intake was clearly associated with the composition of the microbiota; because, the "colourful vegetable, plant protein, and dairy" pattern was positively associated with enterotypes 1 and 3 and negatively with enterotype 2.

Chapters 3 and 4 highlighted that, for all participants, habitual diets characterised by higher intake of dietary fibre, and 'core' foods were inversely associated with BF% and metabolic disease risk (irrespective of age, ethnicity, deprivation, or energy intake). Thus, the observations of this study agree with the evidence of the health impact associated with the quality of the food source habitually consumed, including how diet can affect everyone irrespective of age, deprivation, or ethnicity. Based on the evidence, higher intake of core plant-based foods rich in dietary fibre (fruit, vegetables, wholegrains etc.), coupled with lower intake of discretionary foods, is associated with a decreased risk of weight gain, obesity, and developing T2DM (25–28). However, this study contributes to the evidence by clearly highlighting that, habitual dietary intake is associated with microbiota composition, and metabolic disease risks.

Higher adherence to the dietary pattern characterised by nutrient rich core foods (i.e., high in dietary fibre) was positively associated with enterotypes 1 and 3; and, this aligns with the functional capacity of the bacterial species which characterised these enterotypes (i.e., butyrate producers). Additionally, women characterised by these enterotypes had higher alpha diversity (Shannon index) including, lower weight, BMI, BF%, VAT, and circulating concentrations of fasting insulin in comparison to women characterised by enterotype 2. Women characterised by enterotype 2 did not follow this dietary pattern and consumed more discretionary foods (thus lower dietary fibre intake). The microbiota may, therefore, mediate the beneficial effects of higher fibre diets and health outcomes, as plant polysaccharides are largely indigestible to the host. Higher habitual intake of nutrient rich core foods may promote higher microbiota diversity by providing a variety of substrates for the microbiota. This may, in turn, be more beneficial for the host because a diverse microbiota is more resilient to external influences; for instance, has a greater capacity to do different things (i.e., in response to perturbations).

There was no difference in energy intake between tertiles of habitual dietary fibre intake, or between enterotypes; however, there were differences in the foods which contributed to these energy intakes. Dietary fibre intake is likely to be a marker of food choice and thus diet quality (i.e., nutrient density). However, the social context of how these dietary intakes and patterns arise needs to be considered. There were significant differences in deprivation levels

between participants (e.g., tertiles of dietary fibre intake and adherence to dietary patterns, and enterotypes). Despite adjusting all analyses for deprivation and observing clear associations between habitual dietary intake, microbiota composition, and BF%, deprivation is likely significantly contributing to the observed associations. Similar to global trends, living in a deprived area in NZ increases the likelihood of being obese (1), while also decreasing access to plant-based core foods, such as fruit and vegetables (secondary to cost (29,30)). Therefore, habitual diet quality is likely rooted in deprivation. Although we did not assess the qualitative reasons for food choice, the evidence suggests that those with limited financial means will often choose inexpensive energy dense foods to save costs (31,32), and this is an area that requires further investigation.

Overall, the observations of this study have emphasised how enterotypes are a strong classification tool to identify metabolic disease risk and are strongly associated with habitual diet as previously identified (13). Despite women characterised by enterotype 2 being associated with increased multifactorial metabolic disease risk (higher BF%, VAT, AF%, concentrations of fasting insulin etc.), lower microbiota diversity and higher F:B ratio, we cannot confirm if these microbiota characteristics are indicators of higher metabolic disease risk. However, a higher F:B ratio reflects that a higher relative abundance of specialised *Firmicutes* species are likely present to degrade a smaller number of specific plant polysaccharides, in contrast to generalist *Bacteroides* species, which are likely present in lower relative abundance, because *Bacteroides* genomes encode a much larger number of carbohydrate degrading enzymes in comparison to *Firmicutes* (33). Thus, *Bacteroides* species have a greater capacity to utilise complex carbohydrates; which, suggests women characterised by enterotype 1 and 3 habitually consume a diverse range of dietary fibre. It has yet to be determined if the microbiota is a cause or consequence of metabolic health. However, the observations from this study suggest that, higher intake of nutrient rich core foods and dietary fibre, are associated with lower adiposity and metabolic disease risks, in a population of women with different metabolic disease risk, body fat profiles, and deprivation levels.

Due to the design of this study it is difficult to compare results with other studies. For example, few studies have used metagenomic sequencing, explored the differences between disease

free, young, and ethnically diverse populations, analysed habitual dietary data with the NCI method, or explored *a posteriori* dietary patterns in relation to the microbiota. In addition, these enterotypes and dietary patterns are specific to this study population. Regardless, higher intake of core foods (notably plant based foods rich in dietary fibre) and lower intake of discretionary foods are inversely associated with BF%, VAT, AF% and concentrations of fasting insulin levels. If habitually consuming a variety of nutrient rich core foods and dietary fibre is associated with lower adiposity and metabolic health risks globally, including in a population of women with different metabolic disease risk, and body fat profiles, then public health strategies should prioritise addressing the barriers to habitual consumption of these foods. Public health messages (i.e., dietary guidelines) have long advocated the importance of habitually consuming such foods to reduce the risk of chronic diseases; however, there appears to be a disconnect with the public's adherence to these recommendations. This is likely secondary to a myriad of factors, such as, accessibility, affordability, convenience, and personal values; although this study did not explore food choice, the evidence suggests that food choice is associated with these factors (32,34,35). With that in mind, the current food environment does not support habitual consumption of nutrient rich core foods (32,36–38). These foods are typically more expensive than discretionary foods and are not as readily available, or convenient (29,32,36–38). Therefore, to support the healthy choice being the easy choice, the food environment needs to change, which will likely require legislative support to do so (32).

The microbiota may indeed be associated with (if not a marker of) healthier hosts; however, what this study has highlighted is the association of habitual diet and microbiota composition. The way we eat must make sense in the context of our lives, and if plant based, and nutrient rich core foods cost more than discretionary foods, individuals with higher deprivation and limited financial means are more likely to buy inexpensive energy dense foods to save costs. Additionally, if discretionary foods are widely available, relatively inexpensive, and convenient, then irrespective of deprivation levels individuals are more likely to consume more of such products, which may disrupt the microbiota. Exploring the composition of the microbiota may be the first step to highlight increased metabolic disease risk which, with understanding, may enable early intervention and thus prevention of developing chronic diseases. This study has highlighted that habitual dietary intake and the microbiota are linked.

Food choice, likely associated with deprivation, is strongly associated with microbiota composition more so than ethnicity, age, or energy intake. Higher habitual discretionary food intake is associated with increased metabolic health risks, which overtime may disrupt the microbiota. However, questions remain regarding the cause of the observed associations. For instance, is the gut microbiota a product of habitual diet which increases metabolic disease risks or is the gut microbiota driving the metabolic disease risks?

7.4. Concluding remarks

This PhD research aimed to improve the understanding of the association between habitual diet and compositional characteristics of the microbiota in relation to body fat content and metabolic health markers in a large free-living population. Estimating habitual dietary intake with the NCI method improved the validity of the habitual dietary data, by reducing intra-individual variation and improving the estimation of episodically consumed foods. Dietary pattern analysis enabled the whole diet in relation to the microbiota to be explored. Metagenomic shotgun sequencing of the microbiota enabled the characterisation of the microbiota down to the species level, in a relatively large sample of healthy women with different metabolic disease risk and body fat profiles.

Higher intake of dietary fibre, nutrient rich core foods, and lower intake of discretionary foods was inversely associated with BF% and metabolic disease risks for both Pacific and NZE women, despite these women having different metabolic disease risk, body fat profiles, and deprivation levels. Characterisation of the microbiota identified three enterotypes. Higher adherence to the dietary pattern characterised by nutrient rich core foods was positively associated with microbiota composition (enterotypes 1 and 3) and lower metabolic disease risks. Higher discretionary food intake was associated with higher levels of deprivation, microbiota composition (enterotype 2), and higher multifactorial metabolic disease risks. Enterotyping may be a feasible method to identify and characterise the microbiota in future studies to assess metabolic disease risk. In addition, utilising statistical methods to analyse habitual dietary data can increase the accuracy of the dietary data while reducing participant and researcher burden. Further, dietary pattern analyses enables the whole diet to be considered in relation to health outcomes and microbiota composition.

This study highlighted how food choice appears to be strongly associated with levels of deprivation, as higher quality diets (i.e., higher intake of nutrient rich core foods) were associated with older women with lower levels of deprivation. However, this study also emphasised, that certain characteristics of habitual diet are associated with the same metabolic disease risks for these women, irrespective of ethnicity, body fat group, or deprivation level. These results suggest that habitual dietary intake, characterised by higher discretionary food intake, is associated with microbiota composition as well as increased health risks for the host. However, the causes of obesity are complex and multifactorial, and the gut microbiota may be another host risk factor along with suboptimal diet and inactivity. This study highlights the need for future research to consider the social determinants of health, and how the environment in which the individual lives may inadvertently influence the composition of their microbiota. More studies are required to determine if modulation of the microbiota translates to health benefits. However, until this time, the healthy choice needs to be the easy, and affordable choice, and the environment an individual interacts with needs to support this. Habitually consuming more nutrient and dietary fibre rich core foods, could reduce metabolic disease risks; and would provide a variety of substrates to support a diverse microbiota which, may translate to a healthy host and microbiota.

7.5. Strengths of the study

This study had several strengths that must be emphasised

- This study consisted of a large population of women who were selectively recruited based on ethnicity (to assess different metabolic disease risk) and BMI (lean and obese) to explore associations with different body fat profiles.
- Data were collected over a 14-month period reducing the issue of seasonal effect on dietary intake.
- Dietitians reviewed all food records with participants (to address any potential ambiguities) and used standardised protocols to enter all dietary data into analysis software.
- High-quality dietary data were the basis of the habitual dietary data estimated using the NCI method.

- Metagenomic shotgun sequencing provided deep analysis of the composition of the microbiota to be achieved.
- *A posteriori* dietary pattern analysis enabled unique patterns of intake to be identified, as well as enabling the whole diet to be explored in relation to the microbiota.
- Body composition was assessed with the 'gold standard' DXA, which also enabled the whole and regional adiposity of the participants to be assessed.
- A wide range of metabolic markers were assessed to explore metabolic disease risk.
- Faecal water content was objectively measured from the actual faecal sample provided, rather than using a subjective measure such as the Bristol stool chart.
- The data for this PhD research is based on a larger cross-sectional, multi-centre, and multidisciplinary study which, enabled strong methodological and analytical approaches to be conducted, improving the overall quality of the data. For instance, expert microbiologists analysed the microbiota, dietitians led the dietary assessment arm of the study, and epidemiology statisticians were available for discussion about the complex data.
- All regression analyses were adjusted for covariates that could either affect dietary intake or microbiota composition (e.g., age, deprivation level, ethnicity, energy intake, and faecal water content).

7.6. Limitations of the study

This study had several limitations which must be acknowledged.

- The design was cross-sectional; therefore, causality cannot be inferred, and the microbiota and metabolic markers were only assessed at one time point.
- Underreporting and misreporting of dietary intake is an issue with all dietary assessment methods. Participants of the PROMISE study wore triaxial accelerometers; however, at the time of this study, the energy expenditure data were not available to compare to reported energy intake. Therefore, misreporting cut-offs used within this study were based on epidemiological cut-offs.
- The study population was selectively recruited based on BMI and ethnicity. Thus, these results are not generalisable to other populations as it is not a representative sample of NZE or Pacific women living in NZ.

- *A posteriori* dietary patterns are specific to the population and, therefore, cannot be generalised to other populations. Further, there are subjective decisions which can be introduced throughout analyses of dietary patterns, but this study explored the relevant research to guide decisions and aimed to report all decisions made in a transparent and reproducible manner.
- Enterotypes identified cannot be generalised to other populations.
- Conducting research in two distinct ethnicities posed challenges to meet the needs of the individuals, whilst ensuring standardisation of data collection methods.
- Travel to the data collection facility was a barrier for some potential participants.
- Initial intention was to recruit Pacific women with both Pacific parents and a BMI <25 kg/m^2; however, throughout data collection, this proved challenging. Subsequently, for recruitment purposes the criteria was extended to a BMI of 27 kg/m^2 and the requirements were relaxed in that only one parent had to be of Pacific origin, as long as the Pacific women identified as Pacific ethnicity.
- Individuals who participated in this study are more likely to be interested in health and potentially health conscious (i.e., lead healthier lives).

7.7. Recommendations for future research

- To help recruit a wider range of participants, in order to assess a more representative sample of the population; if the target study population predominately lives in other regions, the practicalities of having off site data collection for groups, where travel may be a barrier to participation, should be explored.
- Recording dietary intake online or in a digital format may help increase accuracy and decrease participant and researcher burden. For example, methods such as the Automated Self-Administered 24-Hour (ASA24) dietary assessment tool developed by the NCI, should be explored with regards to how it could be adapted for use in other settings (i.e., based on local/relevant food composition databases).
- To assess the intra- and inter-individual differences and the stability of the microbiota composition overtime, collecting samples at more than one time point (e.g., faecal sample, metabolic markers, and dietary assessment) from the same individual would enable the assessment of adherence to diets (patterns, food choice, etc.) and

directional influences to be inferred. Where feasible, conducting longitudinal studies over several years would facilitate a deeper understanding of the interactions between dietary intake, body fat profiles, metabolic health markers, and microbiota composition, which would be invaluable.

- To assess the stability of enterotypes, an open ended long-term dietary intervention (e.g., getting participants characterised by enterotype 2 to increase their intake of dietary fibre and nutrient rich core foods) to explore the timeframe required, and whether participants could change their enterotype is required. Further, whether an individual is able to maintain a change of enterotype if the cease the dietary intervention, would be interesting to explore.
- To standardised dietary assessment in the diet-microbiota field, the NCI method is a valuable analysis method which should be explored in other settings. It can be administered at a large scale (with multiple 24-hour recalls or food records) and with appropriate considerations could reduce the intra-individual variation inherent within dietary assessment (and potential participant and researcher burden).
- To assess the association of the whole diet in relation to microbiota composition, future studies should consider dietary pattern analysis.
- To determine the efficacy of diet-microbiota interventions during recruitment and/or allocation to treatment, future studies should explore baseline microbiota composition with enterotyping, and habitual diet with the NCI method and dietary pattern analysis.
- To account for external influences which could influence diet and/or microbiota composition, future studies should explore socio-economic deprivation, lifestyle, and environmental influences (e.g., habitual exercise, purchasing behaviours, cooking skills). This could help advise effective public health prevention strategies and advance the development of therapeutic strategies.

7.8. References

1. Ministry of Health. New Zealand Health Survey 2018/19: Body size: Obese [Internet]. 2019 [cited 2019 Nov 14].

2. Beck KL, Houston ZL, McNaughton SA, Kruger R. Development and evaluation of a food frequency questionnaire to assess nutrient intakes of adult women in New Zealand. Nutr Diet. 2018;1–7. doi:10.1111/1747-0080.12472

3. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87. doi:10.1016/j.jada.2006.07.003

4. National Cancer Institute. Usual dietary intakes: SAS macros for the NCI method [Internet]. 2018 [cited 2019 May 6].

5. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018;107(1):80–93. doi:10.1093/ajcn/nqx002

6. Willett W. Nutritional epidemiology. 3rd ed. New York: Oxford University Press; 2012. 1–514 p. doi:10.1093/acprof:oso/9780199754038.001.0001

7. Gibson RS. Principles of nutritional assessment [Internet]. 2nd ed. New York: Oxford University Press; 2005. 1–980 p.

8. Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, et al. Modeling data with excess zeros and measurement error: Application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics. 2009;65(4):1003–10. doi:10.1111/j.1541-0420.2009.01223.x

9. Hu FB. Dietary pattern analysis: A new direction in nutritional epidemiology. Curr Opin Lipidol [Internet]. 2002 Feb;13(1):3–9. doi:10.1097/00041433-200202000-00002

10. Moeller SM, Reedy J, Millen AE, Dixon LB, Newby PK, Tucker KL, et al. Dietary patterns: Challenges and opportunities in dietary patterns research. J Am Diet Assoc. 2007;107(7):1233–9. doi:10.1016/j.jada.2007.03.014

11. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. Nutr Rev. 2004;62(5):177–203. doi:10.1301/nr.2004.may.177

12. Organization for Economic Co-operation and Development (OECD). OECD Obesity update 2017. Paris, France; 2017.

13. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science (80-) [Internet]. 2011 Oct 7;334(6052):105–8. doi:10.1126/science.1208344

14. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. Science (80-) [Internet]. 2016 Apr 29;352(6285):560–4. doi:10.1126/science.aad3503

15. Losasso C, Eckert EM, Mastrorilli E, Villiger J, Barrucci F, Pernthaler J, et al. Assessing the influence of vegan, vegetarian and omnivore oriented westernized dietary styles on human gut microbiota: A cross sectional study. Front Microbiol. 2018;9(March):1–12. doi:10.3389/fmicb.2018.00317

16. Menni C, Jackson MA, Pallister T, Steves CJ, Spector TD, Valdes AM. Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain. Int J Obes. 2017;41(7):1099–105. doi:10.1038/ijo.2017.66

17. Claesson M, Jeffery IB, Conde S, Power S, O'Connor E, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature. 2012;(488):178–84. doi:10.1038/nature11319

18. Partula V, Mondot S, Torres MJ, Kesse-Guyot E, Deschasaux M, Assmann K, et al. Associations between usual diet and gut microbiota composition: Results from the Milieu Intérieur cross-sectional study. Am J Clin Nutr. 2019;109(5):1472–83. doi:10.1093/ajcn/nqz029

19. 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 (80-) [Internet]. 2016 Apr 29;352(6285):565–9. doi:10.1126/science.aad3369

20. Lin D, Peters BA, Friedlander C, Freiman HJ, Goedert JJ, Sinha R, et al. Association of dietary fibre intake and gut microbiota in adults. Br J Nutr [Internet]. 2018;120(9):1014–22. doi:10.1017/s0007114518002465

21. Oluwagbemigun K, Foerster J, Watkins C, Fouhy F, Stanton C, Bergmann MM, et al. Dietary patterns are associated with serum metabolite patterns and their association is influenced by gut bacteria among older german adults. J Nutr [Internet]. 2019 Sep 2;1–10. doi:10.1093/jn/nxz194

22. Kong LC, Holmes BA, Cotillard A, Habi-Rachedi F, Brazeilles R, Gougis S, et al. Dietary patterns differently associate with inflammation and gut microbiota in overweight and obese subjects. PLoS One. 2014;9(10). doi:10.1371/journal.pone.0109434

23. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. Am J Clin Nutr [Internet]. 2019 Oct 1;110(4):1003–14. doi:10.1093/ajcn/nqz174

24. Rush EC, Freitas I, Plank LD. Body size, body composition and fat distribution: Comparative analysis of European, Maori, Pacific Island and Asian Indian adults. Br J Nutr [Internet]. 2009 Aug 10;102(04):632. doi:10.1017/S0007114508207221

25. Seidelmann SB, Claggett B, Cheng S, Henglin M, Shah A, Steffen LM, et al. Dietary carbohydrate intake and mortality: A prospective cohort study and meta-analysis. Lancet Public Heal [Internet]. 2018 Sep;3(9):e419–28. doi:10.1016/S2468-2667(18)30135-X

26. Reynolds A, Mann J, Cummings J, Winter N, Mete E, Te Morenga L. Carbohydrate quality and human health: A series of systematic reviews and meta-analyses. Lancet. 2019; doi:10.1016/s0140-6736(18)31809-9

27. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary

risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72. doi:10.1016/s0140-6736(19)30041-8

28. United States Department of Agriculture. A series of systematic reviews on the relationship between dietary patterns and health outcomes. Department of Health and Human Services. 2014.

29. Amoah I, Cairncross C, Rush E. The battle for better nutrition: the role of the escalating fruit and vegetable prices. N Z Med J. 2017;130(1465):107–8.

30. New Zealand Ministry of Health. New Zealand Health Survey 2017/2018: Nutrition: Vegetable and fruit intake [Internet]. [cited 2019 Apr 11].

31. Darmon N, Briend A, Drewnowski A. Energy-dense diets are associated with lower diet costs: A community study of French adults. Public Health Nutr. 2004;7(1):21–7. doi:10.1079/PHN2003512

32. Swinburn BA, Kraak VI, Allender S, Atkins VJ, Baker PI, Bogard JR, et al. The global syndemic of obesity, undernutrition, and climate change: The lancet commission report. Lancet [Internet]. 2019 Feb;393(10173):791–846. doi:10.1016/S0140-6736(18)32822-8

33. Tannock GW, Liu Y. Guided dietary fibre intake as a means of directing short-chain fatty acid production by the gut microbiota. J R Soc New Zeal [Internet]. 2019 Aug 26;0(0):1–22. doi:10.1080/03036758.2019.1657471

34. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: The development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev Med (Baltim) [Internet]. 1999 Dec;29(6):563–70. doi:10.1006/pmed.1999.0585

35. Nestle M, Wing R, Birch L, DiSogra L, Drewnowski A, Middleton S, et al. Behavioral and social influences on food choice. Nutr Rev [Internet]. 2009 Apr 27;56(5):50–64. doi:10.1111/j.1753-4887.1998.tb01732.x

36. Monteiro CA, Moubarac J, Cannon G, Ng SW, Popkin B. Ultra-processed products are becoming dominant in the global food system. Obes Rev. 2013;14:21–8. doi:10.1111/obr.12107

37. Vandevijvere S, Jaacks LM, Monteiro CA, Moubarac J, Girling-Butcher M, Lee AC, et al. Global trends in ultraprocessed food and drink product sales and their association with adult body mass index trajectories. Obes Rev [Internet]. 2019 Nov 17;20(S2):10–9. doi:10.1111/obr.12860

38. Headey DD, Alderman HH. The relative caloric prices of healthy and unhealthy foods differ systematically across income levels and continents. J Nutr [Internet]. 2019 Nov 1;149(11):2020–33. doi:10.1093/jn/nxz158

8. Appendices

8.1. Appendix 1. Publications, conference presentations and abstracts

2019

Renall N, Merz B, Lawley B, Tannock G.W, Corbin M, Douwes J, Slater J, Kruger R, Breier B.H. (2019). Body composition of New Zealand European and Pacific women is associated with lower dietary fibre intake, and gut microbiota diversity.
Poster presentation at the *13th Federation of European Nutrition Societies (FENS) European Nutrition Conference*. Dublin, Ireland. Oct 2019.

Renall N, Merz B, Corbin M, Douwes J, Tannock G.W, Slater J, Kruger R, Breier B.H. (2019). Habitual dietary fibre intake is associated with discretionary food intake and body composition of New Zealand European and Pacific women.
Poster presentation at the *Food Structures, Digestion and Health International Conference.* Rotorua, New Zealand. Oct 2019.

Renall, N. (2019). Is the key to weight loss understanding the garden in your gut?
Three-minute thesis (3MT) oral presentation at the *Massey University 3MT Finals 2019.* Palmerston North, New Zealand. July 2019.

Kindleysides S, Kruger R, Douwes J, Tannock G, **Renall N**, Slater J, et al. PRedictors linking Obesity and the gut MIcrobiomE (The PROMISE study): Protocol and recruitment strategy of a cross-sectional study on pathways that affect the gut microbiome and its impact on obesity. JMIR Res Protoc. 2019;8:1–16. doi:10.2196/14529

2018

Renall N, Slater J, Kruger R, Richter M, Tannock G.W, Breier B.H. (2018). Dietary intake for New Zealand European and Pacific women from the PROMISE study.
Poster presentation at the *New Zealand Nutrition Society Annual Meeting.* Auckland, New Zealand. Nov 2018.

Renall N, Slater J, Kruger R, Richter M, Tannock G.W, Breier B.H. (2018). Preliminary analysis of the dietary intake for New Zealand European and Pacific women from the PROMISE study.

Poster presentation at the *Riddet Institute Student Colloquium.* Wellington, New Zealand. July 2018.

2017

Renall N, Tannock G.W, Te Morenga L, Richter M, Kruger R, Breier B.H. (2017). Understanding the garden in your gut.

Three-minute thesis (3MT) oral presentation at the *Riddet Institute Student Colloquium.* Auckland, New Zealand. Oct 2017.

2016

Renall N, Tannock G.W, Te Morenga L, Richter M, Kruger R, Breier B.H. (2016). New pathways to obesity prevention and metabolic health: The relationship between diet and the gut microbiome

Oral presentation at the *Riddet Institute Student Colloquium.* Palmerston North, New Zealand. Oct 2016.

8.2. Appendix 2. Contributions of authors

Chapter 4: Habitual macronutrient intake and BF% paper	Nikki Renall	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study. Estimated habitual intake with NCI method, analysed all dietary data, interpreted results, main author of manuscript
	Benedikt Merz	Assisted with NCI modelling of habitual dietary data. Assisted with statistical analyses and interpretation of results. Revised and approved manuscript
	Marine Corbin	Assisted with statistical analyses and interpretation of results
	Jeroen Douwes	Conceptualised and designed PROMISE study wrote the grant application for funding support. Assisted with statistical analyses and interpretation of results
	Joanne Slater	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study
	Gerald W Tannock	Conceptualised and designed PROMISE study wrote the grant application for funding support
	Rozanne Kruger	Conceptualised and designed PROMISE study wrote the grant application for funding support. Revised and approved manuscript
	Bernhard H Breier	Conceptualised and designed PROMISE study wrote the grant application for funding support. Applied for ethics. Revised and approved manuscript

Chapter 5: Dietary pattern and BF% paper	Nikki Renall	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study. Estimated habitual intake with NCI method, analysed all dietary data, interpreted results, main author of manuscript
	Benedikt Merz	Assisted with NCI modelling of habitual dietary data. Assisted with statistical analyses and interpretation of results. Revised and approved manuscript
	Marine Corbin	Assisted with statistical analyses and interpretation of results
	Jeroen Douwes	Conceptualised and designed PROMISE study wrote the grant application for funding support. Assisted with statistical analyses and interpretation of results
	Joanne Slater	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study
	Gerald W Tannock	Conceptualised and designed PROMISE study wrote the grant application for funding support
	Rozanne Kruger	Conceptualised and designed PROMISE study wrote the grant application for funding support. Revised and approved manuscript
	Bernhard H Breier	Conceptualised and designed PROMISE study wrote the grant application for funding support. Applied for ethics. Revised and approved manuscript

Chapter 6: Microbiota and diet paper	Nikki Renall	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study. Estimated habitual intake with NCI method, analysed all dietary data, interpreted results. Analysed microbiota data in QIIME, STAMP and SAS. Main author of manuscript
	Blair Lawley	Processed and analysed microbiota data, assisted with bioinformatics and interpretation of results
	Benedikt Merz	Assisted with NCI modelling of habitual dietary data. Assisted with statistical analyses and interpretation of results. Revised and approved manuscript
	Marine Corbin	Assisted with statistical analyses and interpretation of results
	Jeroen Douwes	Conceptualised and designed PROMISE study wrote the grant application for funding support. Assisted with statistical analyses and interpretation of results
	Joanne Slater	Recruited participants, involved in the organisation and data acquisition of the PROMISE study. Conducted food record reviews and processed dietary data for PROMISE study
	Rozanne Kruger	Conceptualised and designed PROMISE study wrote the grant application for funding support. Revised and approved manuscript
	Bernhard H Breier	Conceptualised and designed PROMISE study wrote the grant application for funding support. Applied for ethics. Revised and approved manuscript
	Gerald W Tannock	Conceptualised and designed PROMISE study wrote the grant application for funding support. Processed and analysed microbiota data and assisted with interpretation of results. Revised and approved manuscript

8.3. Appendix 3. Demographic questionnaire

MASSEY UNIVERSITY
COLLEGE OF HEALTH
TE KURA HAUORA TANGATA

Subject Number:

PROMISE Study

Health and Demographics Questionnaire

What is your first language?
English ☐
Other ☐ Please state:_____

Do you have children? Yes ☐ No ☐
If yes,
How many children do you have? _____
What are their age(s)? _____
When was your youngest child born? __ / __ / ____ (Day/Month/Year)

What is the highest level of education you have received?
Primary School ☐
Secondary School (College e.g. school certificate, Bursary, NCEA level 1-3) ☐
Trade certificate or diploma ☐
University or other tertiary education (e.g. postgraduate diploma and certificate, bachelor's degree, Masters, PhD) ☐
Other (please state) ☐ _____

What would be the total income that the <u>household</u> received from <u>all sources</u> before tax has been taken out in the last 12 months?

Loss	☐	$30,001 - $40,000	☐
Zero	☐	$40,001 - $50,000	☐
$1 - $5,000	☐	$50,001 - $70,000	☐
$5,001 - $10,000	☐	$70,001 – 100,000	☐
$10,001 - $15,000	☐	$100,000 – or more	☐
$15,001 - $20,000	☐	I don't want to answer	☐
$20,001 - $30,000	☐		

During the past 5 years, what was your main occupation or job title? (e.g., teacher, veterinarian, dairy farmer, librarian, social worker, housekeeping, etc.)?
Please be as specific as possible

Fill in 'unemployed' if you were unemployed for most of the past 5 years

	Occupation	**mm/yy**	**mm/yy**	**Still working**
1	_____	from ___/___	to ___/___	☐
2	_____	from ___/___	to ___/___	☐
3	_____	from ___/___	to ___/___	☐
4	_____	from ___/___	to ___/___	☐
5	_____	from ___/___	to ___/___	☐

In your current profession what was the main activity of the Company or Organisation you worked for? (If currently unemployed, refer to your last recent employment)
Please be as specific as possible

Fill in 'not applicable' if you were unemployed for most of the time during the past 5 years or when your main activity was housekeeping

Please describe
1 _____
2_____
3_____4___
_____
5_____

How many hours do you USUALLY work each WEEK? _____ hours

How many hours do you USUALLY work each DAY? _____ hours

What is your USUAL work pattern?
Please tick the box that best applies.
daytime with no shifts ☐
rotating shifts with nights ☐
rotating shifts without nights ☐
permanent nights ☐
irregular or variable ☐
other work pattern ☐ Please specify _____

If you work night shifts, how many do you work in a USUAL week? _____nights out of 7

Diet and Health

Do you follow a specific diet? Yes ☐ No ☐

If yes, please specify the diet that you follow:

_____

If yes, do you follow this diet for health reasons? Yes ☐ No ☐

If no, why do you follow this diet?

_____

Do you follow any diet for cultural or religious reasons? Yes ☐ No ☐

If yes, what type of diet do you follow?

_____

Would you say you eat your meals?

Quickly ☐ At a moderate pace ☐ Slowly ☐

When did your last period start? __ / __ / ____ (Day/Month/Year)

Are you pregnant? Yes ☐ No ☐

Do you have any surgical implants (e.g., pins in bones/teeth, pacemaker, cosmetic)?

Yes ☐ No ☐

Do you smoke cigarettes? Yes ☐ No ☐

If yes, approximately how many cigarettes per day:

_____ _____

Do you drink alcohol? Yes ☐ No ☐

If yes, approximately how many standard drinks per week: _____

[1 standard drink = 1 glass of wine (100ml), 1 bottle/can of beer, 1 nip of spirits (45mL)]

Do you have any allergies? Yes ☐ No ☐

Please specify _____

_____

Are you taking any form of medication, including traditional or homeopathic medicine and contraception?

Yes ☐ No ☐

If yes, please specify the condition, the medication and the dosage in the table below.

Condition	Medication	Dosage	Frequency

Are you taking any probiotics or prebiotics? Yes ☐ No ☐

If yes, please provide the brand name and variety _____
Are you taking any form of supplements, including tablets or drinks?
Yes ☐ No ☐

If yes, please tell us the name, brand and dosage of the supplements you are taking in the table below.

Supplement	Brand	Dosage	Frequency

By what method were <u>you</u> born?
☐ Caesarean
☐ Natural (vaginal)
☐ I don't know
What was your birth weight in pounds? _____

Please tell us how you found out about this study. Did you find out from:

- ☐ A friend?
- ☐ An email list?
  If yes, what is the name of the email list?_____
- ☐ At an event?
  If yes, which event? _____
- ☐ Flyer on noticeboard?
  If yes, where was the noticeboard? _____
- ☐ The Fono medical clinic
- ☐ Other_____

Would you like to receive a brief report summarizing the main findings of the project? Yes ☐ No ☐

Are you willing to be contacted in future research projects within the Massey University School of Food and Nutrition? Yes ☐ No☐

8.4. Appendix 4. Food record

MASSEY UNIVERSITY
COLLEGE OF HEALTH
TE KURA HAUORA TANGATA

PROMISE Study

5 Day Food Record

Thank you very much for taking part in the PROMISE Study. We are extremely grateful for your time, effort and commitment!
If you have any questions, please contact PROMISE staff on:
414 0800 (extn 49013)email: *promise@massey.ac.nz*

All information in this diary will be treated with the strictest confidence. No one outside the PROMISE study will have access to this data.

Please bring this food diary with you to visit 2 at the Nutrition Laboratory

What to do?

- Record all that you eat and drink on the following dates.

_____
_____
_____

- If possible, record food at the time of eating or just after – try to avoid doing it from memory at the end of the day.
- Include all meals, snacks, and drinks, even tap water.
- Include anything you have added to foods such as sauces, gravies, spreads, dressings, etc.
- Write down any information that might indicate **size or weight** of the food to identify the portion size eaten.
- Use a new line for each food and drink. You can use more than one line for a food or drink. See the examples given.
- Use as many pages of the booklet as you need.

Describing Food and Drink

- Provide as much detail as possible about the type of food eaten. For example, **brand names and varieties / types** of food.

General description	**Food record description**
Breakfast example – cereal, milk, sugar	1 cup Sanitarium Natural Muesli 1 cup Pam's whole milk 1 tsp Chelsea white sugar
Coffee	1 tsp Gregg's instant coffee 1 x 200ml cup of water 2 Tbsp Meadow fresh light green milk
Pasta	1 cup San Remo whole grain pasta spirals (boiled)
Pie	Big Ben Classic Mince and Cheese Pie (170g)

- Give details of all the **cooking methods** used. For example, fried, grilled, baked, poached, boiled…

General description	**Food record description**
2 eggs	2 size 7 eggs fried in 2tsp canola oil 2 size 6 eggs (soft boiled)
Fish	100g salmon (no skin) poached in 1 cup of water for 10 minutes

- When using foods that are cooked (e.g., pasta, rice, meat, vegetables, etc), please record the **cooked portion** of food.

General description	Food record description
Rice	1 cup cooked Jasmine rice (cooked on stove top)
Meat	90g lean T-bone steak (fat and bone removed)
Vegetables	½ cup cooked mixed vegetables (Wattie's peas, corn, carrots)

- Please specify the **actual amount of food eaten** (e.g., for leftovers, foods where there is waste)

General description	Food record description
Apple	1 x 120g Granny Smith apple (peeled, core not eaten – core equated to ¼ of the apple)
Fried chicken drumstick	100g chicken drumstick (100g includes skin and bone); fried in 3 Tbsp Fern leaf semi-soft butter

- **Record recipes** of home prepared dishes where possible and the proportion of the dish you ate. There are blank pages for you to add recipes or additional information.

<u>Recording the amounts of food you eat</u>

It is important to also record the quantity of each food and drink consumed. This can be done in several ways.

- By using household measures – for example, cups, teaspoons and tablespoons. e.g., 1 cup frozen peas, 1 heaped teaspoon of sugar.
- By weight marked on the packages – e.g., a 425g tin of baked beans, a 32g cereal bar, 600ml Coke
- For bread – describe the size of the slices of bread (e.g., sandwich, medium, toast) – also include brand and variety.
- Using comparisons – e.g., Meat equal to the size of a pack of cards, a scoop of ice cream equal to the size of a hen's egg.
- Use the food record instructions provided to help describe portion sizes.

General description	Food record description
Cheese	1 heaped tablespoon of grated cheese 1 slice cheese (8.5 x 2.5 x 2mm) 1 cube cheese, match box size Size 10B grated cheese,

- If you go out for meals, describe the food eaten in as much detail as possible.

Please eat as normally as possible - don't adjust what you would normally eat just because you are keeping a diet record and be honest! Your food record will be identified with a number rather than your name.

Example day

Time food was eaten	Complete description of food (food and beverage name, brand, variety, preparation method)	Amount consumed (units, measures, weight)
7:55am	Sanitarium Weet-Bix	2 Weet-Bix
" "	Anchor Blue Top milk	150ml
" "	Chelsea white sugar	2 heaped teaspoons
" "	Orange juice (Citrus Tree with added calcium – nutrition label attached)	1 glass (275 ml)
10.00am	Raw Apple (gala)	Ate all of apple except the core, whole apple was 125g (core was ¼ of whole apple)
12.00pm	Home made pizza (recipe attached)	1 slice (similar size to 1 slice of sandwich bread, 2 Tbsp tomato paste, 4 olives, 2 rashers bacon (fat removed), 1 Tbsp chopped spring onion, 3 Tbsp mozzarella cheese)
1.00pm	Water	500ml plain tap water
3.00pm	Biscuits, chocolate covered Girl Guide biscuits	6 x (standard size)
6.00pm	Lasagne	½ cup cooked mince, 1 cup cooked Budget lasagne shaped pasta , ½ cup Wattie's creamy mushroom and herb pasta sauce, ½ cup mixed vegetables (Pam's carrots, peas and corn), 4 Tbsp grated Edam cheese
6.30pm	Banana cake with chocolate icing (homemade, recipe attached)	1/8 of a cake (22cm diameter, 8 cm high), 2 Tbsp chocolate icing
" "	Tip Top Cookies and Cream ice cream	1 cup (250g)
7.30pm	Coffee	1 tsp Gregg's instant coffee 1 x 300ml cup of water 2 Tbsp Meadow fresh blue top milk

Date_____ DAY 1

Time food was eaten	Complete description of food (food and beverage name, brand, variety, preparation method)	Amount consumed

Date_____ DAY 1 (continued)

Time food was eaten	Complete description of food (food and beverage name, brand, variety, preparation method)	Amount consumed

Recipes (Day 1)

8.5. Appendix 5. Food Frequency Questionnaire

PROMISE Food Frequency Questionnaire

1. Please read carefully before you begin:

Please make sure when filling out this questionnaire that you:

- **Tell us what YOU usually eat (not someone else in your household!).**
- **Fill in the form YOURSELF.**
- **Are correct, but don't spend too much time on each food.**
- **Answer EVERY question; the asterisk symbol (*) at the beginning of each question means that you must answer before moving onto the next question.**

This will help us to get the most accurate information about your usual food intake.

Please answer by ticking the box which best describes HOW OFTEN you ate or drank a particular food or drink in the LAST MONTH and HOW MUCH you would usually have.

For example:

1. EXAMPLE: How often do you usually have sugar? (Please do not fill out)

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Sugar - 1 tsp	○	○	○	○	○	○	○	○	○

If every day you have 2 cups of coffee with 1 tsp sugar, 4 cups of tea with 1 tsp sugar, one bowl of cereal with 1 tsp sugar and sugar on pancakes at dinner, you would choose four or more times per day = '4+ x / day'.

Adjust your portion size and frequency of intake to suit your eating habits.

2. EXAMPLE: How often do you usually eat bread? (Please do not fill out)

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Bread - 1 slice	○	○	○	○	○	○	○	○	○

If every day you have two slices of toast for breakfast, and you have a sandwich for lunch three times per week, you would choose two - three times per day = '2-3x / day'.

Adjust your portion size and frequency of intake to suit your eating habits.

1

PROMISE Food Frequency Questionnaire

2. PROMISE Food Frequency Questionnaire

* 1. Please enter your study ID (if you are unsure or don't know please ask the researcher)

2

PROMISE Food Frequency Questionnaire

3. Eating Pattern

* 1. How would you describe your eating pattern? (Please choose one only)

- ○ Eat a variety of all foods, including animal products
- ○ Eat eggs, dairy products, fish and chicken but avoid other meats
- ○ Eat eggs, dairy products and fish, but avoid chicken and other red meats
- ○ Eat eggs and dairy products, but avoid all meats, chicken and fish
- ○ Eat eggs, but avoid dairy products, all meats and fish
- ○ Eat dairy products, but avoid eggs, all meats and fish
- ○ Eat no animal products
- ○ None of the above

Other (please state)

PROMISE Food Frequency Questionnaire

4. Dairy

* 1. Do you use milk? (e.g. fresh, UHT, powdered)

- ○ Yes
- ○ No

* 2. What type(s) of milk do you have most often? (You can choose up to 3 options, but please only choose the ones you usually have)

- ☐ Not applicable
- ☐ Full cream milk (purple top)
- ☐ Standard milk (blue top)
- ☐ Skim milk (light blue top)
- ☐ Trim milk (green top)
- ☐ Super trim milk (light green top)
- ☐ Calcium enriched milk (yellow top) e.g. Xtra, Calci-Trim
- ☐ Calcium and vitamin enriched milk e.g. Mega, Anlene
- ☐ Calcium and protein enriched milk e.g. Sun Latte
- ☐ Standard soy milk (blue)
- ☐ Light soy milk (light blue)
- ☐ Calcium enriched soy milk (purple) e.g. Calci-Forte, Calci-Plus
- ☐ Calcium, vitamin and omega 3 enriched soy milk e.g. Essential
- ☐ Calcium and high fibre enriched soy milk e.g. Calci-Plus High Fibre
- ☐ Rice milk

Other (please state)

* 3. Choose the one milk you have the most

- [ ] Not applicable
- [ ] Full cream milk (purple top)
- [ ] Standard milk (blue top)
- [ ] Skim milk (light blue)
- [ ] Trim milk (green top)
- [ ] Super trim milk (light green top)
- [ ] Calcium enriched milk (yellow top) e.g. Xtra, Calci-Trim
- [ ] Calcium and vitamin enriched milk e.g. Mega, Anlene
- [ ] Calcium and protein enriched milk e.g. Sun Latte
- [ ] Standard soy milk (blue)
- [ ] Light soy milk (light blue)
- [ ] Calcium enriched soy milk (purple) e.g. Calci-Forte, Calci-Plus
- [ ] Calcium, vitamin and omega 3 enriched soy milk e.g. Essential
- [ ] Calcium and high fibre enriched soy milk e.g. Calci-Plus High Fibre
- [ ] Rice milk

Other (please state)

* 4. On average, how many servings of milk do you have per day? (Please choose one only)
(A 'serving' = 250 mL or 1 cup/glass)
e.g. 5 cups of coffee/tea using 50 mL of milk + ½ cup of milk on cereal = 1 ½ servings per day

- ( ) Not applicable
- ( ) Less than 1 serving
- ( ) 1-2 servings
- ( ) 3-4 servings
- ( ) 5 or more servings

5

* 5. How often do you usually have milk?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Flavoured milk (milkshake, iced coffee, Primo, Nesquik) - 250 mL/ 1 cup									
Milk as a drink - 250 mL / 1 cup									
Milk on breakfast cereals or porridge - 125 mL/ 1/2 cup									
Milk added to water-based hot drinks (coffee, tea) - 50 mL / 1/5 cup									
Milk-based hot drinks (Latte, Milo) - 250 mL / 1 cup									

* 6. How often do you usually eat cheese?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day	4+ x / day
Cheddar (tasty, mild, colby) - 2 heaped Tbsp / matchbox cube										
Edam, Gouda, Swiss - 2 heaped Tbsp / matchbox cube										
Feta, Mozarella, Camembert - 1 heaped Tbsp / 1 med wedge										
Brie, blue and other specialty cheese - 1 heaped Tbsp / 1 med wedge										
Processed cheese slices - 1 slice										
Cream cheese - 2 heaped Tbsp										
Cottage or ricotta cheese - 2 heaped Tbsp										

* 7. How often do you usually eat these dairy based foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Ice cream - 2 scoops									
Custard or dairy food - 1 pottle / ½ cup									
Yoghurt, plain or flavour - 1 pottle / ½ cup									
Milk puddings (semolina, instant) - ½ cup									
Fermented or evaporated milk (buttermilk) - ½ cup									

PROMISE Food Frequency Questionnaire

5. Bread

* 1. Do you eat bread?

- ○ No
- ○ Yes

* 2. What type(s) of bread, rolls or toast do you eat most often? (You can choose up to 3 options, but please only choose the ones you usually have)

- ☐ Not applicable
- ☐ White
- ☐ White – high fibre
- ☐ Wholemeal or wheat meal
- ☐ Wholegrain

Other (please state)

* 3. What type of bread slice do you usually have? (Please choose one only)

- ○ Not applicable
- ○ Sandwich slice
- ○ Toast slice
- ○ Mixture of both sandwich and toast slices

* 4. On average, how many servings of bread do eat per day? (Please choose one only)
(A 'serving' = 1 slice of bread or 1 small roll)

- ○ Not applicable
- ○ Less than 1 serving
- ○ 1–2 servings
- ○ 3–4 servings
- ○ 5–6 servings
- ○ 7 or more servings

* 5. How often do you usually eat these bread based foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Plain white bread - 1 slice	○	○	○	○	○	○	○	○	○
High fibre white bread - 1 slice	○	○	○	○	○	○	○	○	○
Wholemeal or wheat meal - 1 slice	○	○	○	○	○	○	○	○	○
Wholegrain bread - 1 slice	○	○	○	○	○	○	○	○	○
Fruit bread or fruit bun - 1 slice	○	○	○	○	○	○	○	○	○
Wrap - 1 medium	○	○	○	○	○	○	○	○	○
Focaccia, bagel, pita, panini or other speciality breads - 1 medium	○	○	○	○	○	○	○	○	○
Paraoa Parai (fry bread) - 1 slice	○	○	○	○	○	○	○	○	○
Rewena bread - 1 slice	○	○	○	○	○	○	○	○	○
Doughboys or Maori bread - 1 slice	○	○	○	○	○	○	○	○	○

* 6. How often do you usually eat these other bread based foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Crumpet or muffin split - 1 crumpet / 1 whole muffin split	○	○	○	○	○	○	○	○	○
Scone - 1 medium	○	○	○	○	○	○	○	○	○
Bran muffin or savoury muffin - 1 medium	○	○	○	○	○	○	○	○	○
Croissant - 1 medium	○	○	○	○	○	○	○	○	○
Waffle, pancakes or pikelets - 1 medium / 2 small	○	○	○	○	○	○	○	○	○
Iced buns - 1 medium	○	○	○	○	○	○	○	○	○
Crackers (cream crackers, cruskits, corn / rice crackers, vitawheat) - 2 medium	○	○	○	○	○	○	○	○	○

* 7. Do you have butter, margarine or spreads on bread or crackers?

○ No

○ Yes

* 8. What type(s) do you have most often? (You can choose up to 3 options, but please only choose the ones you usually have)

- [ ] Not applicable
- [ ] Butter (all varieties)
- [ ] Monounsaturated fat margarine e.g. Olive, Rice Bran, Canola Oil Spreads
- [ ] Polyunsaturated fat margarine e.g. Sunflower Oil Spreads
- [ ] Light monounsaturated fat margarine e.g. Olivio Spread Light
- [ ] Light polyunsaturated fat margarine e.g. Flora Spread Light
- [ ] Plant sterol enriched margarine e.g. Pro Active, Logical Spreads
- [ ] Light plant sterol enriched margarine e.g. Pro Active Spread Light
- [ ] Butter and margarine blend e.g. Country Soft, Butter Lea

Other (please state)

* 9. Choose the one you have the most

- [ ] Not applicable
- [ ] Butter (all varieties)
- [ ] Monounsaturated fat margarine e.g. Olive, Rice Bran, Canola Oil Spreads
- [ ] Polyunsaturated fat margarine e.g. Sunflower Oil Spreads
- [ ] Light monounsaturated fat margarine e.g. Olivio Spread Light
- [ ] Light polyunsaturated fat margarine e.g. Flora Spread Light
- [ ] Plant sterol enriched margarine e.g. Pro Active, Logical Spreads
- [ ] Light plant sterol enriched margarine e.g. Pro Active Spread Light
- [ ] Butter and margarine blend e.g. Country Soft, Butter Lea
- [ ] Other (please state)

* 10. On average, how many servings of butter, margarine or spreads do you have per day? (Please choose one only)
(A 'serving' = 1 level teaspoon or 5 mL)
e.g. 1 sandwich with butter thinly spread on two pieces of bread = 2 servings

- Not applicable
- Less than 1 serving
- 1–2 servings
- 3–4 servings
- 5–6 servings
- 7 or more servings

10

PROMISE Food Frequency Questionnaire

6. Breakfast Cereals and Porridge

* 1. Do you usually eat breakfast cereal and/or porridge?

- ○ No
- ○ Yes

* 2. What breakfast cereal(s) do you eat most often? (You can choose up to 3 options, but please only choose the ones you usually have)

- ☐ Not applicable
- ☐ Weetbix
- ☐ Refined cereals e.g. Cornflakes or Rice Bubbles
- ☐ Bran based cereals including fruity varieties e.g. Special K, Muesli, All Bran
- ☐ Sweetened e.g. Nutrigrain, Cocoa Pops
- ☐ Porridge

Other (please state)

* 3. On average, how many servings of breakfast cereal or porridge do you have per week? (Please choose one only)
(A 'serving' = ½ cup porridge, muesli, cornflakes or 2 weetbix)
e.g. ½ cup of porridge 3 times per week + 2 weetbix 4 times a week = 7 servings per week

- ○ Not applicable
- ○ Less than 4 servings
- ○ 4–6 servings
- ○ 7–9 servings
- ○ 10–12 servings
- ○ 13–15 servings
- ○ 16 or more servings

* 4. How often do you usually eat porridge or these cereal foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Porridge, rolled oats, oat bran, oat meal - ½ cup	○	○	○	○	○	○	○	○	○
Muesli (all varieties) - ½ cup	○	○	○	○	○	○	○	○	○
Weetbix (all varieties) - 2 weetbix	○	○	○	○	○	○	○	○	○
Cornflakes or rice bubbles - ½ cup	○	○	○	○	○	○	○	○	○
Bran cereals (All Bran, Bran Flakes) - ½ cup	○	○	○	○	○	○	○	○	○
Bran based cereals (Sultana Bran, Sultana Bran Extra) - ½ cup	○	○	○	○	○	○	○	○	○
Light and fruity cereals (Special K, Light and Tasty) - ½ cup	○	○	○	○	○	○	○	○	○
Chocolate based cereals (Milo cereal, Coco Pops) - ½ cup	○	○	○	○	○	○	○	○	○
Sweetened cereals (Nutrigrain, Fruit Loops, Honey Puffs, Frosties) - ½ cup	○	○	○	○	○	○	○	○	○
Breakfast drinks (Up and Go) - Small carton / 250 mL	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

7. Starchy Foods

* 1. Do you eat any type of starchy foods such as rice, pasta, noodles and couscous?

- ○ No
- ○ Yes

* 2. On average, how many servings of starchy foods such as rice, pasta, noodles and couscous do you eat per week? (Please choose one only)
(A 'serving' = 1 cup cooked rice / pasta)
e.g. 1 cup of rice + ½ cup of pasta included in a lasagne pasta dish + 1 cup of spaghetti = 2.5 servings

- ○ Not applicable
- ○ Less than 4 servings
- ○ 4–6 servings
- ○ 7–9 servings
- ○ 10–12 servings
- ○ 13–15 servings
- ○ 16 or more servings

* 3. How often do you usually eat these starchy foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Rice, white - 1 cup	○	○	○	○	○	○	○	○	○
Rice, brown or wild - 1 cup	○	○	○	○	○	○	○	○	○
Pasta, white or wholegrain (spaghetti, vermicelli) - 1 cup	○	○	○	○	○	○	○	○	○
Canned spaghetti (Watties) - 1 cup	○	○	○	○	○	○	○	○	○
Instant noodles (2 minute noodles) - 1 packet	○	○	○	○	○	○	○	○	○
Egg and rice noodles (hokkien noodles, udon) - 1 cup	○	○	○	○	○	○	○	○	○
Other grain (quinoa, couscous, bulgar wheat) - 1 cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

8. Meat

* 1. Do you eat beef, mutton, hogget, lamb, or pork

- ○ No
- ○ Yes

* 2. Do you trim any excess fat (fat you can see) off these meats? (Please choose one only)

- ○ Not applicable
- ○ Always
- ○ Often
- ○ Occasionally
- ○ Never cut the fat off meat

* 3. On average, how many servings of meat e.g. beef, mutton, hogget, lamb or pork do you eat per week? (Please choose one only)
(A 'serving' = palm size or ½ a cup of meat without bone)
e.g. ½ cup of savoury mince + 2 small lamb chops = 2 servings

- ○ Not applicable
- ○ Less than 1 serving
- ○ 1-3 servings
- ○ 4-6 servings
- ○ 7 or more servings

* 4. How often do you usually eat meat?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Beef mince dishes (rissoles, meatloaf, hamburger pattie) - 1 slice / patty / ½ cup	○	○	○	○	○	○	○	○	○
Beef or veal mixed dishes (casserole, stir-fry) - ½ cup	○	○	○	○	○	○	○	○	○
Beef or veal (roast, chop, steak, schnitzel, corned beef) - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Lamb, hogget or mutton mixed dishes (stews, casserole, stir-fry) - ½ cup	○	○	○	○	○	○	○	○	○
Lamb, hogget or mutton (roast, chops, steak) - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Pork (roast, chop, steak) - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Canned corned beef - 1 medium slice	○	○	○	○	○	○	○	○	○

* 5. How often do you usually eat these other meats?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Sausage, frankfurter or saveloy - 1 sausage / frankfurter/ 2 saveloys	○	○	○	○	○	○	○	○	○
Bacon - 2 rashers	○	○	○	○	○	○	○	○	○
Ham - 1 medium slice	○	○	○	○	○	○	○	○	○
Luncheon meats or brawn - 1 slice	○	○	○	○	○	○	○	○	○
Salami or chorizo - 1 slice / cube	○	○	○	○	○	○	○	○	○
Offal (liver, kidneys, pate) - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Venison/game - palm size / ½ cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

9. Poultry

* 1. Do you eat poultry e.g. chicken, turkey or duck?

- ○ No
- ○ Yes

* 2. Do you remove the skin from chicken? (Please choose one only)

- ○ Not applicable
- ○ Always
- ○ Often
- ○ Occasionally
- ○ Never remove the skin from chicken

* 3. On average, how many servings of chicken do you eat per week? (Please choose one only)
(A 'serving' = palm size of chicken or ½ cup)
e.g. 1 chicken breast + 2 chicken drumsticks + 1 chicken thigh = 4 servings per week

- ○ Not applicable
- ○ Less than1 serving
- ○ 1-3 servings
- ○ 4-6 servings
- ○ 7 or more servings

* 4. How often do you usually eat poultry?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Chicken legs or wings - palm size / ½ cup / 1 unit (wing, drumstick)	○	○	○	○	○	○	○	○	○
Chicken breast - palm size / ½ cup / ½ breast	○	○	○	○	○	○	○	○	○
Chicken mixed dishes (casserole, stir-fry) - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Crumbed chicken (nuggets, patties, schnitzel) - 1 medium / 4 nuggets	○	○	○	○	○	○	○	○	○
Turkey or quail - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Mutton bird or duck - palm size / ½ cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

10. Fish and Seafood

* 1. Do you eat any type of fish or seafood?

- ○ No
- ○ Yes

* 2. On average, how many servings of fish and seafood (all types; fresh, frozen, tinned) do you eat per week? (Please choose one only)
(A 'serving' = 80 - 120g or palm size or small tin (85g))
e.g. 1 fish fillet and 1 small tin of tuna = 2 servings per week.

- ○ Not applicable
- ○ Less than 1 serving
- ○ 1-3 servings
- ○ 4-6 servings
- ○ 7 or more servings

* 3. How do you normally cook / eat fish? (You can choose up to 3 options, but please only choose the ones you usually have)

- ☐ Not applicable
- ☐ Raw / I don't cook it
- ☐ Oven baked / Grilled
- ☐ Deep fried
- ☐ Shallow fry
- ☐ Micro waved
- ☐ Steamed
- ☐ Poached
- ☐ Smoked

* 4. How often do you usually eat seafood?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Canned Salmon - 1 small can (85-95g)	○	○	○	○	○	○	○	○	○
Canned Tuna - 1 small can (85-95g)	○	○	○	○	○	○	○	○	○
Canned Mackerel, sardines, anchovies, herring - 1 small can (85-95g)	○	○	○	○	○	○	○	○	○
Frozen crumbed fish (patties, fillets, cakes, fingers, nuggets) - 1 medium / 4 nuggets	○	○	○	○	○	○	○	○	○
Snapper, Tarakihi, Hoki, Cod, Flounder - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Gurnard, Kahawai or Trevally - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Lemon fish or Shark - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Tuna - palm size / ½ cup	○	○	○	○	○	○	○	○	○
Salmon, trout or eel - palm size / ½ cup	○	○	○	○	○	○	○	○	○

* 5. How often do you usually eat seafood?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Shrimp, prawn, lobster or crayfish - ½ cup	○	○	○	○	○	○	○	○	○
Crab or surumi - ½ cup	○	○	○	○	○	○	○	○	○
Scallops, mussels, oysters, paua or clams - ½ cup	○	○	○	○	○	○	○	○	○
Pipi or cockle - ½ cup	○	○	○	○	○	○	○	○	○
Kina - ½ cup	○	○	○	○	○	○	○	○	○
Whitebait - ¼ cup	○	○	○	○	○	○	○	○	○
Roe - ¼ cup	○	○	○	○	○	○	○	○	○
Squid, octopus, calamari, cuttlefish - ½ cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

11. Fats and Oils

* 1. Do you cook meat, chicken, fish, eggs and/or vegetables with fat or oil?

- ○ No
- ○ Yes

* 2. What type(s) do you use most often? (You can choose up to 3 options, but please only choose the ones you usually have)

- ☐ Not applicable
- ☐ Butter (all varieties)
- ☐ Margarines (all varieties)
- ☐ Cooking oils (all varieties)
- ☐ Lard, Dripping, Coconut oil, Ghee (clarified butter)
- ☐ Cooking spray

Other (please state)

* 3. Chose the one you use the most

- ☐ Not applicable
- ☐ Butter (all varieties)
- ☐ Margarines (all varieties)
- ☐ Cooking oils (all varieties)
- ☐ Lard, Dripping, Coconut oil, Ghee (clarified butter)
- ☐ Cooking spray
- ☐ Other (please state)

* 4. When you use fat or oil to cook, how many servings of fat or oil do you use per dish? (Please choose one only)
(A 'serving' = 1 level teaspoon or 5 mL)

- Not applicable
- Less than 1 serving
- 1 serving
- 2 servings
- 3 servings
- 4 servings
- 5 or more servings

* 5. On average, how many servings of fat or oil do you use to cook per week? (Please choose one only)

- Not applicable
- Less than 1 serving
- 1-3 servings
- 4-7 servings
- 8-10 servings
- 11-14 servings
- 15 or more servings

PROMISE Food Frequency Questionnaire

12. Eggs

* 1. Do you eat eggs?

- ○ No
- ○ Yes

* 2. On average, not counting eggs used in baking / cooking, how many eggs do you usually eat per week? (Please choose one only)

- ○ Not applicable
- ○ Less than 1 egg
- ○ 1 egg
- ○ 2 eggs
- ○ 3 eggs
- ○ 4 eggs
- ○ 5 or more eggs

* 3. How often do you usually eat eggs?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Whole eggs (hard-boiled, poached, fried, mashed, omelette, scrambled) - 1 egg	○	○	○	○	○	○	○	○	○
Mixed egg dish (quiche, frittata, other baked egg) - 1 slice	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

13. Legumes

* 1. Do you eat legumes e.g. chickpeas/dried peas, soybeans, dried/canned beans, baked beans, lentils or Dahl?

- No
- Yes

* 2. On average, how many servings of legumes (fresh, frozen, canned, dried) do you eat per week? (Please choose one only)
(A 'serving' = ½ cup or 125g of cooked legumes)

- Not applicable
- Less than 1 serving
- 1 serving
- 2 servings
- 3 servings
- 4-5 servings
- 6-7 servings
- 8 or more servings

* 3. How often do you usually eat these legumes?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Soybeans - ½ cup	○	○	○	○	○	○	○	○	○
Tofu - ½ cup	○	○	○	○	○	○	○	○	○
Dahl - ½ cup	○	○	○	○	○	○	○	○	○
Canned or dried legumes, beans (baked beans, chickpeas, lentils, peas, beans) - ½ cup	○	○	○	○	○	○	○	○	○
Hummus - 2 Tbsp	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

14. Vegetables

* 1. Do you eat vegetables?

- ○ No
- ○ Yes

* 2. On average, how many servings of vegetables (fresh, frozen, canned) do you eat per day? Do NOT include vegetable juices. (Please choose one only)
(A 'serving' = 1 medium potato / kumara or ½ cup cooked vegetables or 1/2 cup of lettuce)
e.g. 2 medium potatoes + ½ cup of peas = 3 servings

- ○ Not applicable
- ○ Less than 1 serving
- ○ 1 serving
- ○ 2 servings
- ○ 3 servings
- ○ 4 or more servings

* 3. How often do you usually eat these vegetables?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Potato (boiled, mashed, baked, roasted) - 1 medium / ½ cup									
Pumpkin (boiled, mashed, baked, roasted) - ½ cup									
Kumara (boiled, mashed, baked, roasted) - 1 medium / ½ cup									
Mixed frozen vegetables - ½ cup									
Green beans - ½ cup									
Silver beet, spinach - ½ cup									
Carrots - 1 medium / ½ cup									
Sweet corn - 1 medium cob / ½ cup									
Mushrooms - ½ cup									
Tomatoes - 1 medium / ½ cup									
Beetroot - 1 medium / ½ cup									
Taro, cassava or breadfruit - 1 medium / ½ cup									

* 4. How often do you usually eat these vegetables?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Green bananas (plantain) - 1 medium / ½ cup									
Sprouts (alfalfa, mung) - ½ cup									
Pacific Island yams - 1 medium / ½ cup									
Turnips, swedes, parsnip or yams - ½ cup									
Onions, celery or leeks - ¼ cup									
Cauliflower, broccoli or broccoflower - ½ cup									
Brussel sprouts, cabbage, red cabbage or kale - ½ cup									
Courgette/zucchini, marrow, eggplant, squash, kamo kamo, asparagus, cucumber - ½ cup									
Capsicum (peppers) - ½ medium / ¼ cup									
Avocado - ¼ avocado									
Lettuce greens (mesculin, cos, iceberg) - ½ cup									
Other green leafy vegetables (whitloof, watercress, taro leaves, puha) - ½ cup									

PROMISE Food Frequency Questionnaire

15. Fruit

* 1. Do you eat fruit?

- ○ No
- ○ Yes

* 2. On average, how many servings of fruit (fresh, frozen, canned or stewed) do you eat per day? Do NOT include fruit juice. (Please choose one only)
(A 'serving' = 1 medium or 2 small pieces of fruit or 1/2 cup of chopped fruit)
e.g. 1 apple + 2 small apricots = 2 servings)

- ○ Not applicable
- ○ Less than one serving
- ○ 1 serving
- ○ 2 servings
- ○ 3 or more servings

* 3. How often do you usually eat these fruits?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Apple - 1 medium / ½ cup	○	○	○	○	○	○	○	○	○
Pear - 1 medium / ½ cup	○	○	○	○	○	○	○	○	○
Banana - 1 medium / ½ cup	○	○	○	○	○	○	○	○	○
Orange, mandarin, tangelo, grapefruit - 1 medium / 2 small	○	○	○	○	○	○	○	○	○
Peach, nectarine, plum or apricot - 1 medium / ½ cup / 2 small	○	○	○	○	○	○	○	○	○
Mango, paw-paw or persimmons / ½ cup	○	○	○	○	○	○	○	○	○
Pineapple - ½ cup	○	○	○	○	○	○	○	○	○
Grapes - ½ cup / 8-10 grapes	○	○	○	○	○	○	○	○	○
Strawberries, other berries, cherries - ½ cup	○	○	○	○	○	○	○	○	○
Melon (watermelon, rockmelon) - ½ cup	○	○	○	○	○	○	○	○	○
Kiwifruit - 1 medium / 2 small	○	○	○	○	○	○	○	○	○
Feijoas - 1 medium / 2 small	○	○	○	○	○	○	○	○	○
Tamarillos - 1 medium / ½ cup	○	○	○	○	○	○	○	○	○
Sultanas, raisins or currants - 1 small box	○	○	○	○	○	○	○	○	○
Other dried fruit (apricots, prunes, dates) - 4 pieces	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

16. Drinks

* 1. On average, how many drinks do you have per day? (Please choose one only)
(A 'serving' = 250 mL or one cup/glass)

- Less than 1 serving
- 1-3 servings
- 4-5 servings
- 6-8 servings
- 9-10 servings
- 11 or more servings

* 2. How often do you usually have these drinks?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Instant soup (Cup of soup) - 250 mL / 1 cup	○	○	○	○	○	○	○	○	○
Fruit juice (Just Juice, Fresh-up, Charlie's, Rio Gold) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Fruit drink (Choice, Rio Splice) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Vegetable juice (tomato juice, V8 juice) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Iced Tea (Lipton ice tea) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Cordial or Powdered drinks (Thriftee, Raro, Vita-fresh) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Low-calorie cordial - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Energy drinks small-medium can (V, Red Bull) - 250-350 mL	○	○	○	○	○	○	○	○	○
Energy drinks large can (Monster, Mother, Demon, large V) - 450-550 mL	○	○	○	○	○	○	○	○	○
Sugar-free Energy drinks (sugar-free V, Monster, Red Bull) - 1 small can	○	○	○	○	○	○	○	○	○
Diet soft/fizzy/carbonated drink (diet sprite) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Soft/fizzy/carbonated drinks (Coke, Sprite) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○
Sport's drinks (Gatorade, Powerade) - 1 bottle	○	○	○	○	○	○	○	○	○
Flavoured water (Mizone, H2Go flavoured) - 1 bottle	○	○	○	○	○	○	○	○	○
Water (unflavoured mineral water, soda water, tap water) - 250 mL / 1 cup/glass	○	○	○	○	○	○	○	○	○

* 3. How often do you usually have these drinks?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Coffee instant or brewed with or without milk (Nescafe, expresso) - 1 cup	○	○	○	○	○	○	○	○	○
Specialty coffees (flat white, cappuccino, lattes) - 1 small cup	○	○	○	○	○	○	○	○	○
Coffee decaffeinated or substitute (Inka) - 1 cup	○	○	○	○	○	○	○	○	○
Hot chocolate drinks (drinking chocolate, hot chocolate, Koko) - 1 cup	○	○	○	○	○	○	○	○	○
Milo - 1 tsp	○	○	○	○	○	○	○	○	○
Tea (English breakfast tea, Earl Grey) - 1 cup	○	○	○	○	○	○	○	○	○
Herbal tea or Green tea - 1 cup	○	○	○	○	○	○	○	○	○
Soy drinks - 1 cup	○	○	○	○	○	○	○	○	○

* 4. How often do you usually have these alcoholic drinks?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Beer – low alcohol - 1 can or bottle	○	○	○	○	○	○	○	○	○
Beer – ordinary - 1 can or bottle	○	○	○	○	○	○	○	○	○
Red wine - 1 small glass	○	○	○	○	○	○	○	○	○
White wine, champagne, sparkling wine - 1 small glass	○	○	○	○	○	○	○	○	○
Wine cooler - 1 small glass / bottle	○	○	○	○	○	○	○	○	○
Sparkling grape juice - 1 glass / cup	○	○	○	○	○	○	○	○	○
Sherry or port - 100 mL	○	○	○	○	○	○	○	○	○
Spirits, liqueurs - 1 shot or 30 mL	○	○	○	○	○	○	○	○	○
RTD (KGB, Vodka Cruiser, Woodstock bourbon) - 1 bottle / can	○	○	○	○	○	○	○	○	○
Cider - 1 glass / cup / bottle	○	○	○	○	○	○	○	○	○
Kava - 1 glass / cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

17. Dressings and Sauces

* 1. How often do you usually have these dressings or sauces?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Butter (all varieties) - 1 tsp	○	○	○	○	○	○	○	○	○
Margarine (all varieties) - 1 tsp	○	○	○	○	○	○	○	○	○
Oil (all varieties) - 1 tsp	○	○	○	○	○	○	○	○	○
Cream or sour cream - 1 Tbsp	○	○	○	○	○	○	○	○	○
Mayonnaise or creamy dressings (aioli, tartae sauce) - 1 Tbsp	○	○	○	○	○	○	○	○	○
Low fat/calorie dressing (reduced fat mayonnaise) - 1 Tbsp	○	○	○	○	○	○	○	○	○
Salad dressing (french, italian) - 1 Tbsp	○	○	○	○	○	○	○	○	○
Sauces (tomato, BBQ, sweet chilli, mint) - 1 Tbsp	○	○	○	○	○	○	○	○	○
Mustard - 1 Tbsp	○	○	○	○	○	○	○	○	○
Soy sauce - 1 Tbsp	○	○	○	○	○	○	○	○	○
Chutney or relish - 1 Tbsp	○	○	○	○	○	○	○	○	○
Gravy homemade - ¼ cup	○	○	○	○	○	○	○	○	○
Instant Gravy (e.g. Maggi) - ¼ cup	○	○	○	○	○	○	○	○	○
White sauce/cheese sauce - ¼ cup	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

18. Miscellaneous - Cakes, Biscuits and Puddings

* 1. How often do you usually eat these baked products?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Cakes, loaves, sweet muffins - 1 slice / 1 muffin	○	○	○	○	○	○	○	○	○
Sweet pies or pastries, tarts, doughnuts - 1 medium	○	○	○	○	○	○	○	○	○
Other puddings or desserts - not including milk-based puddings (sticky date pudding, pavlova) - ½ cup	○	○	○	○	○	○	○	○	○
Plain biscuits, cookies (Round wine, Ginger nut) - 2 biscuits	○	○	○	○	○	○	○	○	○
Fancy biscuits (chocolate, cream) - 2 biscuits	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

19. Miscellaneous

* 1. How often do you usually eat these other foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Jelly - ½ cup	○	○	○	○	○	○	○	○	○
Ice blocks - 1 ice block	○	○	○	○	○	○	○	○	○
Lollies - 2 lollies	○	○	○	○	○	○	○	○	○
Chocolate - including chocolate bars (Moro bars) - 1 small bar	○	○	○	○	○	○	○	○	○
Sugar added to food and drinks - 1 level tsp	○	○	○	○	○	○	○	○	○
Jam, honey, marmalade or syrup - 1 level tsp	○	○	○	○	○	○	○	○	○
Vegemite or marmite - 1 level tsp	○	○	○	○	○	○	○	○	○
Peanut butter or other nut spreads - 1 level Tbsp	○	○	○	○	○	○	○	○	○
Brazil nuts or walnuts - 2	○	○	○	○	○	○	○	○	○
Peanuts - 10	○	○	○	○	○	○	○	○	○
Other nuts (almonds, cashew, pistachio, macadamia) - 10	○	○	○	○	○	○	○	○	○
Seeds (pumpkin, sunflower)	○	○	○	○	○	○	○	○	○
Muesli bars - 1 bar	○	○	○	○	○	○	○	○	○
Coconut cream - ¼ cup	○	○	○	○	○	○	○	○	○
Coconut milk - ¼ cup	○	○	○	○	○	○	○	○	○
Lite coconut milk - ¼ cup	○	○	○	○	○	○	○	○	○
Potato crisps, corn chips, Twisties - ½ cup / handful	○	○	○	○	○	○	○	○	○

* 2. Do you use salt in cooking?

- ○ Never
- ○ Rarely
- ○ Sometimes
- ○ Usually
- ○ Always

* 3. Do you use salt at the table?

- Never
- Rarely
- Sometimes
- Usually
- Always

34

PROMISE Food Frequency Questionnaire

20. Miscellaneous - Takeaways

* 1. On average, how often do you eat takeaways per week? (Please choose one only)

- ○ Never
- ○ Less than 1 times
- ○ 1-2 times
- ○ 3-4 times
- ○ 4-6 times
- ○ More than 7 times

* 2. How often do you usually eat these takeaway foods?

	Never	<1x / month	1-3x / month	1x / week	2-3x / week	4-6x / week	Once / day	2-3x / day	4+ x / day
Meat pie, sausage roll, other savouries - 1 pie / 2 small sausage rolls or savouries	○	○	○	○	○	○	○	○	○
Hot potato chips, kumara chips, french fries, wedges - ½ cup	○	○	○	○	○	○	○	○	○
Chinese - 1 serve	○	○	○	○	○	○	○	○	○
Indian - 1 serve	○	○	○	○	○	○	○	○	○
Thai - 1 serve	○	○	○	○	○	○	○	○	○
Pizza - 1 medium slice	○	○	○	○	○	○	○	○	○
Burgers - 1 medium burger	○	○	○	○	○	○	○	○	○
Battered fish - 1 piece	○	○	○	○	○	○	○	○	○
Fried chicken (KFC, Country fried chicken) - 1 medium piece	○	○	○	○	○	○	○	○	○
Bread based (Kebab, sandwiches, wraps, Pita Pit, Subway) - 1 medium	○	○	○	○	○	○	○	○	○

PROMISE Food Frequency Questionnaire

21. Other

* 1. Are there any other foods or drinks that you can think of that you have on a regular basis that was not covered by this questionnaire?

- No
- Yes

22. Other

1. Please list these foods and drinks including; the serving size, and how many times per week you eat or drink these items (e.g. Pizza, 4 slices, one time per week)

8.6. Appendix 6. Food groups

	Food Group	Food Items Included
1	Full fat dairy milk	Dark blue, purple, silver top milk, lactose free regular fat milk
2	Low fat dairy milk	Lite and trim milk (green or light blue), lactose free reduced fat milk
3	Milk alternatives, including sweetened	Soy, almond, rice, oat, coconut varieties, milk alternative based drinks from cafés
4	Sweetened dairy milk products	Flavoured milk, fermented or evaporated milk, breakfast drinks (e.g., Up & Go), Yakult fermented milk drink, hot chocolates, milk-based smoothies, milk-based drinks from cafés, coffee sachets, coffees made with syrup and cream (e.g., caramel macchiato)
5	Dairy yoghurt	All types of cows milk yoghurt. (Note: Soy yoghurt under soy products, coconut yoghurt under coconut fats)
6	High fat cheese	Cream cheese, goat cheese, haloumi, parmesan, cheddar, processed cheese, blue vein, mascarpone
7	Low fat cheese	Brie, bocconcini, edam, feta, mozzarella, camembert, cottage, ricotta, paneer
8	Apple, banana, orange	Apple, banana orange
9	Other fruit	All other fruit (fresh, canned, dried)
10	Tomatoes	Fresh, canned, cooked tomatoes
11	Dark yellow vegetables	Carrots, pumpkin, butternut squash
12	Green vegetables	Lettuce, spinach, cabbage, broccoli, watercress, green beans, brussel sprouts, courgette
13	Other non-starchy vegetables	Capsicum, onion, mushrooms, frozen mixed vegetables, beetroot, squash
14	Potatoes / potato dishes (excluding chips)	Potato (boiled, mashed, baked, salad, scalloped, roasted)
15	Starchy vegetables	Kumara, yam, parsnip, turnip, swedes (boiled, mashed, baked) Taro (flesh, roots, stalks), green banana, sweet corn kernels, breadfruit, cassava, green banana
16	White breads	Plain white bread, wraps, focaccia, bagels, pita bread, rēwena bread, doughboys, breadcrumbs, including gluten free options, naan (plain)
17	Discretionary breads	Crumpets, scone, savoury muffin, plain croissant, pancakes, waffles, iced bun, savoury pin wheels, garlic bread, fruit bread, roti, naan (garlic)
18	Crackers	All crackers made from grains, cream crackers, Cruskits, rice crackers
19	Whole grain products	Wholegrain breads (High fibre, wholemeal, wholegrain, including gluten free options), grains (Quinoa, buckwheat, bulgur wheat, brown rice, wholemeal pasta, wholegrain gluten free pasta e.g., brown rice)
20	Refined grains	White rice, white pasta, noodles (instant, egg, rice), canned spaghetti, cous cous, including gluten free pasta
21	Refined grain mixed dishes	Macaroni and cheese, carbonara, white rice salad, two-minute noodles
22	Oats	Porridge, rolled oats, oat bran
23	Sweetened cereal	Sultana bran, light and fruity cereal, chocolate-based cereals, Nutri-Grain, honey puffs, milo cereals, fruit loops, oat sachets, all muesli and granola
24	Unsweetened cereals	Weet-Bix, bran cereals, rice bubbles, cornflakes
25	Red meats	Beef, lamb, venison, mince, patties, (including red meat mixed dishes: Stir fry, curry, stew)
26	White meats	Chicken, pork, turkey, (including white meat mixed dishes: Stir fry, curry, stew, casserole)

27	Processed meats	Corned beef (canned), corned silverside, smoked chicken, smoked hock and salami, ham, sausages, frankfurters, bacon, chorizo, luncheon meat
28	Fish and seafood	Canned and fresh (including mixed dishes Oka ika mata, curry, stew) and processed fish products (e.g., fish balls)
29	Eggs	Whole eggs (boiled, poached, fried, plain omelette) including egg mixed dishes (Quiche, frittata, omelette with filling, egg and banana pancake)
30	Legumes and meat alternatives	Baked beans, black beans, dahl, canned or dried legumes, hummus, and legume based vegetarian meals and products (including meat alternatives and soy products edamame beans tofu, tempeh)
31	Peanut butter and peanuts	Peanut butter and peanuts
32	Nuts and seeds	Brazil nuts, walnuts, almond, cashew, pistachio, chia, linseed, pumpkin, sesame
33	Animal fats	Cream, sour cream, reduced cream, butter, lard, dripping, ghee
34	Coconut fats and products	Coconut oil, cream, milk, desiccated, yoghurt, fresh
35	Plant based fats	Avocado (whole fruit), canola, sunflower, olive, vegetable oil, cooking spray, oil-based salad dressings (French/Italian)
36	Margarine	All margarines
37	Creamy dressings and sauces	Creamy readymade meal-based sauces, dips, mayonnaise, aioli, tartare sauce, white sauce, cheese sauce
38	Savoury sauces and condiments	Curry pastes, herb and spices, vinegar, gravy and garlic sauce, oil based condiments such as sundried tomatoes / olives in oil, pesto, tomato, barbeque, mint, soy, gravy, mustard, chutney, miso, pasta sauce, tomato paste, sweet chilli sauce (including savoury spreads Vegemite, marmite)
39	Sweet spreads	Jam, honey, marmalade, syrup (maple, golden), Nutella, chocolate peanut butter, chocolate butter
40	Cake and biscuits	Slices, cakes, loaves, muffins, biscuits, doughnuts, sweet pies, pastries, tarts
41	Puddings and other desserts	Including all milk alternative ice creams, ice cream, custard, milk-based puddings (e.g., rice, instant, semolina, pavlova, sticky date, fruit pies and crumbles, jelly, ice blocks)
42	Sweet snack foods	Fruit and nut mixes, bliss balls, chocolate, lollies, muesli bars
43	Savoury snack foods	Popcorn, potato crisps, corn chips, Twisties, bhuja mix
44	Crumbed and deep fried	Hot chips/fries, hash browns, and packaged home baked chips, wontons, (fry) paraoa bread, schnitzel, nuggets, crumbed fish
45	Fast-food (burgers)	Pies, dumpling, burgers, pizzas, curries, noodle-based dishes, Nandos chicken, egg fu yong
46	Fast food (salads & sushi)	Salads, sandwiches, wraps, sushi, vegetable-based stir fry
47	Fruit and vegetable juice	Fruit and/or vegetable juice, fruit and/or vegetable smoothies
48	Soft drinks and other sugar sweetened beverages	All cordials, flavoured water, sports drinks, soft drinks, fruit drinks, iced tea, energy drinks
49	Diet drinks	All exclusively artificially sweetened beverages
50	Tea	Black, green, herbal, chai, kombucha
51	Coffee	Instant, brewed, espresso, pre-mixed sachet, filter, cold brew
52	Alcoholic beverages	Wine (standard and low alcohol), beer (standard and low alcohol), cider, spirits, RTDs, sherry, port, liqueurs, sake
53	Water	Water (unflavoured, soda, tap)
54	Sugar added to food and drink	All sugar added to food and drink
55	Soups and stock	All soups (instant, canned, packet) and stocks

8.7. Appendix 7. Faecal collection information

MASSEY UNIVERSITY
COLLEGE OF HEALTH
TE KURA HAUORA TANGATA

PROMISE Study

Stool Sample Collection

Your take home pack contains:

1. Two labelled screw-top scoop containers
2. Two labelled larger plastic containers
3. Disposable gloves
4. Kidney dishes
5. One zip-lock plastic bag
6. Two brown paper bags
7. 2 Ice-sheets
8. Chiller carrier bag

To collect the stool samples:

- Each container will be already labelled with your study ID. Please write the **DATE** and **TIME** of collection on all labels (two containers) with a ball-point pen.
- Place something in the toilet to catch the stool, such as the kidney dish provided, potty or an empty plastic food container, or spread clean newspaper over the rim of the toilet. You may line the kidney dish with toilet paper and flush the toilet paper away after you have collected the sample from the kidney dish.
- Make sure the stool doesn't touch the inside of the toilet.
- Use the scoop that is inside the lid of one of the small screw top containers to collect some of the stool, aim to fill around a third of the container (about the size of a walnut).
- Repeat this process with the second small screw top container and fill around a third of the container.
  IT IS IMPORTANT TO PLEASE COLLECT 2 SEPARATE SAMPLES from the same stool by using both of the scoop containers we have provided.
- Screw on the lids of the two small containers containing the stool samples tightly.
- Fill each of the two larger containers with about 2cm of water and place the two smaller scoop containers inside the larger containers so that the small containers are surrounded by water (one small container inside one larger container). Screw on the lid of each of the larger containers tightly.
- Place each large container in a brown paper bag and place both containers into the ziplock bag.
- **PLACE IN THE FREEZER DIRECTLY OR AS SOON AS POSSIBLE.**
- At the same time as you place your samples in the freezer you can prepare your ice-sheets by soaking them in warm water and also placing them in the freezer.
- Deliver the two separate stool samples to us when you return to the Human Nutrition Research Unit on your second visit.
- **PLEASE TRANSPORT YOUR FROZEN STOOL SAMPLES TO US IN THE CHILLER BAG WITH THE ICE-SHEETS AROUND THE SAMPLES TO KEEP THE SAMPLES FROZEN. SAMPLES SHOULD NOT BE LEFT OUT OF THE FREEZER FOR MORE THAN AN HOUR.**