DISSERTATIONES
BIOLOGICAE
UNIVERSITATIS
TARTUENSIS
403

NELE TABA

Diet, blood metabolites, and health

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
403

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
403

NELE TABA

Diet, blood metabolites, and health

UNIVERSITY OF TARTU
Press
1632

Institute of Molecular and Cell Biology, Institute of Genomics, University of Tartu, Estonia

This dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Molecular Biomedicine on July 4, 2022 by the Council of the Institute of Molecular and Cell Biology, University of Tartu.

Supervisors: **Prof. Krista Fischer, PhD.** Institute of Mathematics and Statistics, Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia; member of the Estonian Academy of Sciences

Prof. Tõnu Esko, PhD. Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia

Nicola Pirastu, PhD. Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, Scotland, United Kingdom; Genomics Research Centre, Human Technopole, Milan, Italy

Prof. Andres Metspalu, MD, PhD. Estonian Genome Center, Institute of Genomics, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia; member of the Estonian Academy of Sciences

Reviewer: **Prof. Maido Remm, PhD.** Chair of Bioinformatics, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

Opponent: **Prof. Hannelore Daniel, PhD.** Department of Nutrition and Food Sciences, Technical University of Munich, Munich, Germany; member of the German Academy of Sciences Leopoldina

Commencement: Room No. 105, 23B Riia St., Tartu, on September 9 2022, at 12.15 pm.

The publication of this dissertation is granted by the Institute of Genomics and the Institute of Molecular and Cell Biology at the University of Tartu.

This research was funded by the European Union through the European Regional Development Fund Grant No. 2014-2020.4.2.2 scholarships in smart specialization growth areas, and Project No. 2014-2020.4.01.15-0012 GENTRANSMED; the Estonian Research Council Grants PUT1665, PUT1660, PRG1197, PRG1291, IUT20-60 and SP1GI18045T; the Graduate School in Biomedicine and Biotechnology ASTRA project of the University of Tartu, PER ASPERA; Archimedes Foundation scholarships Dora Plus for short study visits and Dora Plus for PhD student mobility. Data analyses were carried out in part in the High-Performance Computing Center of the University of Tartu.

ISSN 1024-6479
ISBN 978-9949-03-945-6 (print)
ISBN 978-9949-03-946-3 (pdf)

University of Tartu Press
www.tyk.ee

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, referred to in the text by Roman numerals (Ref. I to Ref. III):

I **Taba, N.**; Valge, H.-K.; Metspalu, A.; Esko, T.; Wilson, J.F.; Fischer, K.; Pirastu, N. (2021). **Mendelian Randomization Identifies the Potential Causal Impact of Dietary Patterns on Circulating Blood Metabolites**. Frontiers in Genetics, 12, ARTN 738265. https://doi.org/10.3389/fgene.2021.738265.

II Sarin, H.V.; **Taba, N.**; Fischer, K.; Esko, T.; Kanerva, N.; Moilanen, L.; Saltevo, J.; Joensuu, A.; Borodulin, K.; Männistö, S.; Kristiansson, K.; Perola, M. (2019). **Food neophobia associates with poorer dietary quality, metabolic risk factors, and increased disease outcome risk in population-based cohorts in a metabolomics study.** American Journal of Clinical Nutrition, 110 (1), 233–245. https://doi.org/10.1093/ajcn/nqz100.

III Macdonald-Dunlop, E.; **Taba, N.**; Klarić, L.; Frkatović, A.; Walker, R.; Hayward, C.; Esko, T.; Haley, C.; Fischer, K.; Wilson, J.F.; Joshi, P.K. (2022). **A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk.** Aging, 623–659. https://doi.org/10.18632/aging.203847.

The publications listed above have been reprinted with the permission of the copyright owners.

My contributions to the listed publications were as follows:

Ref. I Co-performed the analysis, prepared all the figures and tables, interpreted the results, drafted the manuscript, revised the manuscript

Ref. II Performed the analysis and prepared the tables and figures in the Estonian Biobank, revised the manuscript

Ref. III Validated and assessed the performance of the omics clocks in the Estonian Biobank, revised the manuscript

LIST OF ABBREVIATIONS

2-sample MR	Two-sample Mendelian Randomization
BA	Biological age
BMI	Body mass index
chronAge	Chronological age
CHD	Coronary heart disease
CRP	C-reactive protein
CUR	Corrected-to-uncorrected ratio
CVD	Cardiovascular diseases
DHA	22:6 docosahexaenoic acid
DILGOM	Finnish Dietary Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome
DNAme	DNA methylation
EstBB	Estonian Biobank
FDR	False discovery rate
FFQ	Food Frequency Questionnaire
FN	Food neophobia
FNS	Food Neophobia Scale
GWAS	Genome-wide association study
HDL	High-density lipoprotein
HDL-C	High-density lipoprotein cholesterol
IDL	Intermediate density lipoprotein
IVW	Inverse-variance weighted
LA	18:2 linoleic acid
LDL	Low-density lipoprotein
LDL-C	Low-density lipoprotein cholesterol
MR	Mendelian Randomization
NMR	Nuclear magnetic resonance spectroscopy
OC	Omics clock
OCA	Omics clock age
OCAA	Omics clock age acceleration
ORCADES	Orkney Complex Disease Study
PC	Principal component
PREDIMED	Prevención con Dieta Mediterránea
RCT	Randomized clinical trial
SBP	Systolic blood pressure
SNP	Single nucleotide polymorhpishm
T2D	Type II diabetes
UKBB	UK Biobank
VLDL	Very-low-density lipoprotein

INTRODUCTION

Dietary intake is known to affect individuals' health by influencing disease incidence and mortality. However, the exact mechanisms of these associations have in many cases remained unclear. This lack of understanding can in turn result in conflicting dietary recommendations and advice. These questions would be best answered by randomized clinical trials (RCT), but RCTs are often expensive, unethical, or difficult to pursue on large samples. To overcome these issues, we assess the effects of dietary items on health-related traits using Mendelian Randomization (MR) analysis, which has the potential to indicate which associations are possibly causal. In its essence MR is an instrumental variable analysis, and thus one of the most important prerequisites for pursuing MR are valid instruments. In the current context the instruments are single nucleotide polymorphisms (SNPs) directly associated with variables that indicate dietary choices. These SNPs can be detected with a genome-wide association study (GWAS) of dietary items.

One possible way to investigate the effect of dietary choices on health is to examine the metabolic profile of blood, which could be mediating the effect. Blood metabolites are known to be associated with several health conditions such as cardiovascular diseases (CVD), type II diabetes (T2D), and mortality, and have previously shown to elevate or lower the risk of these. Thus, determining how dietary choices affect blood metabolic profile will be instrumental in unravelling the possible mechanisms of how diet influences health.

I investigate the interactions between diet and health from three different perspectives. First, by studying the effect of dietary choices on blood metabolic profile with MR analysis. Second, by assessing how a systematic lack of dietary variety expressed as food neophobia affects blood metabolome. Third, by examining whether the blood metabolic profile can be indicative of underlying biological age (BA), and whether the excess of such BA over the chronological age (chronAge) is predictive of some disease groups and subsequently indicative of the health of a specific bodily system.

Thus, the overarching aim of the studies presented in this thesis is to shed more light on how diet affects health. The findings also form a basis for future RCTs, which could ultimately lead to (personalized) dietary recommendations.

1. REVIEW OF THE LITERATURE

1.1. Dietary studies in health context – metabolites as possible explanators

Diet is a widely studied human behavior due to its essential role on individuals' health and general well-being. It has been shown to have an important role in the development and progress of several medical conditions, such as obesity (Popkin, 2001), CVD (Schaefer, 2002; Boeing *et al.*, 2012; Dilis *et al.*, 2012), T2D (Schwingshackl *et al.*, 2017) and cancer (Key *et al.*, 2004; Johnson and Lund, 2007). Due to the high burden these diseases create for individuals, society, healthcare and economy, their prevention is of great interest and importance. However, oftentimes the mechanisms by which diet affects health are still unclear. In this case, blood metabolites might be of help to at least partly provide explanations. Metabolites are small molecules that participate in human metabolism, and include omega-3 and omega-6 fatty acids, high density lipoprotein cholesterol (HDL-C), low density lipoprotein cholesterol (LDL-C), citrate, glycose etc. In the recent years, there has been immense progress in the quantification of such metabolites. This has provided the basis for investigating on the one hand how diet affects blood metabolites and on the other hand, which metabolites are indicative of diseases.

Several studies have researched the relationship between diet and the levels of metabolites in blood or urine. These investigations have identified associations between the metabolic profile and coffee (Guertin *et al.*, 2015), alcohol (Würtz *et al.*, 2016), fruit and vegetable consumption (Menni *et al.*, 2013), and a wide range of dietary patterns, such as Western and Prudent (Chandler *et al.*, 2020), Mediterranean (Li *et al.*, 2020), and healthy diet measured with several Healthy Eating indexes (Kim *et al.*, 2021). Further, other studies have shown metabolites to be important in the onset of several diseases, such as T2D (Suhre *et al.*, 2010; Wang *et al.*, 2011), CVD (Holmes *et al.*, 2014; Würtz *et al.*, 2015), colorectal cancer (Guertin *et al.*, 2015; Shu *et al.*, 2018), dementia (Lee *et al.*, 2018; Tynkkynen *et al.*, 2018) and increase the risk of mortality (Fischer *et al.*, 2014; Deelen *et al.*, 2019). Since blood metabolites are subject to change by diet and are often indicative or risk factors of several diseases, they represent a promising candidate for at least partly providing explanations to the mechanisms of diet-disease interplay.

Successful endeavors have been made in pursuing randomized trials for investigating the effect of dietary patterns on blood metabolites and health (Esko *et al.*, 2017; Michielsen *et al.*, 2019). One of the largest of such studies was the Prevención con Dieta Mediterránea (PREDIMED) trial (Estruch *et al.*, 2018), which primarily aimed to assess the effect of Mediterranean diet on cardiovascular health, but also on mortality, diabetes, various cancers, and neurodegenerative diseases as secondary outcomes (Ros *et al.*, 2014). PREDIMED has received critique of the sample size being not large enough and follow-up not long enough to draw firm conclusions about mortality (Guasch-Ferré *et al.*, 2017),

and with regards issues with the randomization (Agarwal and Ioannidis, 2019). Further, another important aspect that concerns the generalizability of the PREDIMED results is that the sample consisted of elderly people at high risk of CVD. However, to date it is the largest randomized dietary trial systematically assessing the effect of a dietary pattern on cardiovascular health. The results after 5 years of follow-up showed a 30% reduction in the relative risk of CVD in people following a Mediterranean diet compared to a group that was advised to simply reduce dietary fat intake (Estruch *et al.*, 2018). Furthermore, as part of the study, blood metabolites were quantified in a subset of participants. Analysis identified a particular metabolic signature of the Mediterranean diet and subsequently confirmed the association of this signature with CVD (Li *et al.*, 2020). Fernández-Lázaro *et al.* (2021) stated that the findings by Li et al. (2020) open up a new era for personalized nutrition. This further confirms that investigating blood metabolites is a plausible strategy when the effect of diet on health is of interest.

Besides main endpoint of CVD, there have been several other studies published using PREDIMED data, such as investigating dairy intake on plasma metabolites and subsequently the association between these metabolites and risk of incident T2D (Drouin-Chartier *et al.*, 2021), and detecting plasma metabolites associated with consumption of red wine (Hernández-Alonso *et al.*, 2019). Thus, studies such as PREDIMED are of great value in understanding better the mechanisms of how diet acts on health.

Although an extremely valuable study, PREDIMED only investigated the effect of one specific dietary pattern over a limited amount of time with limited number of individuals. Pursuing such study in a longer term (>10 years) on a large sample investigating a wide variety of different dietary items would be rather complex and expensive. Therefore, some alternative ideas could be considered. Namely, one option would be the method of Mendelian Randomization, which is based on the concept that individuals were randomized into groups at conception and this information is in their genomes (Davey Smith and Ebrahim, 2003). A comprehensive description of MR is given in chapter 1.4. Briefly, it enables to utilize the information from GWASs pursued in large biobanks in order to infer possibly causal relationships between variables of interest. GWAS is an observational method, where associations between the trait of interest and genetic variants are tested across the whole genome. It has gained rapid popularity in the last two decades, with GWAS Catalog currently reporting the results of more than 5000 GWASs. In our study we use MR to assess the possible causal effect of dietary intake on blood metabolites.

1.2. Food neophobia: characteristics and quantification

Dietary items are usually consumed together and form dietary patterns or can be described by a broader dietary behavior. An example of the latter is the food neophobia (FN), which can result in a diet with low variety and thereby affect blood metabolic profile and disease outcomes. Food neophobia is a term that

characterizes the behavior in which individual is reluctant to taste or eat new or unfamiliar foods. FN is a highly heritable trait, with heritability estimates up to 69% in adults (Knaapila *et al.*, 2007) and up to 78% in children (Cooke, Haworth and Wardle, 2007; Faith *et al.*, 2013). FN can be reliably measured with the Food Neophobia Scale (FNS) questionnaire (Pliner and Hobden, 1992). This questionnaire comprises of ten statements (Table 1) that have to be rated on a 7-point scale indicating how much the person agrees with the statement (1 – "strongly disagree" up to 7 – "strongly agree"). Half of the statements reflect the lack of FN, therefore the points of these questions have to be reversed (the score subtracted from 8) prior to calculating the sum of the scores. The sum can range between 10–70 with higher scores indicating more severe FN.

Table 1. The statements in the Food Neophobia Scale questionnaire.

Statement	Reverse scoring
I am constantly sampling new and different foods.	+
I don't trust new foods.	
If I don't know what is in a food, I won't try it.	
I like foods from different countries.	+
Ethnic food looks too weird to eat.	
At dinner parties, I will try a new food.	+
I am afraid to eat things I have never had before.	
I am very particular about the foods I will eat.	
I will eat almost anything.	+
I like to try new ethnic restaurant	+

The score for each of the questions ranges from 1 to 7: 1 – "strongly disagree", 2 – "moderately disagree", 3 – "slightly disagree", 4 – "neither agree, nor disagree", 5 – "slightly agree", 6 – "moderately agree", 7 – "strongly agree". When a question is marked with "reverse scoring", the score is subtracted from 8 prior to calculating the final score.

Previous research has shown FN to be associated with several factors, age, education level, living area and socioeconomic status among others (Tuorila *et al.*, 2001; Dovey *et al.*, 2008; Meiselman, King and Gillette, 2010). Thus, these factors have to be taken into account when analyzing the association of FN with other traits. More specifically, FN has been shown to be lower in individuals with higher income, higher education and those living in urban areas (Tuorila *et al.*, 2001; Flight, Leppard and Cox, 2003; Meiselman, King and Gillette, 2010; Rabadán and Bernabéu, 2021). Furthermore, children tend to have higher FN compared to adults and the level of FN varies during childhood and adolescence (Dovey *et al.*, 2008). However, this does not affect the analyses for the current thesis, since the focus is on adults, and FN has been considered to be relatively stable during adulthood (Dovey *et al.*, 2008). The latter holds with an exception of some studies showing that among elderly population the levels of FN are increased (Tuorila *et al.*, 2001; Hazley *et al.*, 2022).

FN has been associated with reduced dietary quality (Knaapila *et al.*, 2015; Rabadán and Bernabéu, 2021), dietary variety (Hazley *et al.*, 2022) and lower adherence to the healthy Mediterranean diet (Predieri *et al.*, 2020). Reduced dietary quality and lower adherence to the Mediterranean diet have in turn been associated with an elevated risk for chronic diseases such as coronary heart disease (CHD) and T2D, and inflammation (Esposito *et al.*, 2004; Sofi *et al.*, 2010; Dilis *et al.*, 2012; Kanerva *et al.*, 2014; Ros *et al.*, 2014; Jannasch, Kröger and Schulze, 2017). Therefore, higher FN may be a risk factor for CHD and T2D, but there are no studies investigating this hypothesis to date. Furthermore, as discussed in the previous chapter, blood metabolic profile is affected by dietary choices and indicative of health outcomes. However, the effect of FN on blood metabolic profile has not been investigated before. Thus, this thesis aims to examine the effect of FN on blood metabolites, CHD and T2D.

1.3. NMR molecular profile

Blood metabolic profile can be measured with different techniques. A popular method for metabolic profiling is the nuclear magnetic resonance (NMR) spectroscopy (Soininen *et al.*, 2009; Würtz *et al.*, 2017). This comprises mostly of lipid measurements, such as levels of cholesterol, phospholipids, free cholesterol and triglycerides in lipoproteins of different densities and sizes, fatty acids, apolipoproteins, amino acids and inflammation biomarkers. Consequently, the inferences in this thesis are limited to the metabolites measured by NMR and could be further broadened by future studies with metabolites measured by some other platform.

NMR metabolites are predictive of a wide range of diseases, such as CVD (Würtz *et al.*, 2015; Joshi *et al.*, 2020; Vojinovic *et al.*, 2021), T2D (Ahola-Olli *et al.*, 2019), severe COVID-19 and severe pneumonia (Julkunen *et al.*, 2021), and increased mortality (Fischer *et al.*, 2014; Deelen *et al.*, 2019). They are also associated with dementia (Lee *et al.*, 2018) and depression (Bot *et al.*, 2020). Several randomized trials have investigated the effect of dietary interventions on metabolic profile measured by NMR. For example, the effect of dietary counseling as intervention (Lehtovirta *et al.*, 2018), effect of monounsaturated fatty acid rich diet, Mediterranean diet and saturated fatty acid rich diet (Michielsen *et al.*, 2019), effect of different dairy products (Hansson *et al.*, 2019), and different fish oils (Rundblad *et al.*, 2017). However, the sample sizes of such studies are rather small compared to large-scale biobank cohorts, while assessing very limited number of different dietary interventions. As stated above, one possible way to utilize large biobanks in order to draw possibly causal associations between long-term dietary habits and blood metabolic profile would be by using the method of MR. Previously, MR has been successfully used for detecting potentially causal effect of body mass index (BMI) on NMR metabolites (Würtz *et al.*, 2014), but no studies have yet utilized MR to assess the effect of a wide range of dietary traits on blood metabolites.

1.4. Mendelian Randomization: possible causal associations

Mendelian Randomization is a term for instrumental variable analysis that aims to estimate causal effect of an exposure variable on an outcome variable, while using SNPs associated with the exposure variable as the instruments (see Figure 1 for a schematic representation). In other words, MR can be seen in a way as a randomized trial, where alleles associated with exposure are randomized by nature. In the context of this thesis, the exposure variables are reflecting the dietary consumption, outcome variables are blood metabolites, and instrumental variables are SNPs associated with dietary traits. The latter can be detected with a GWAS of dietary items.

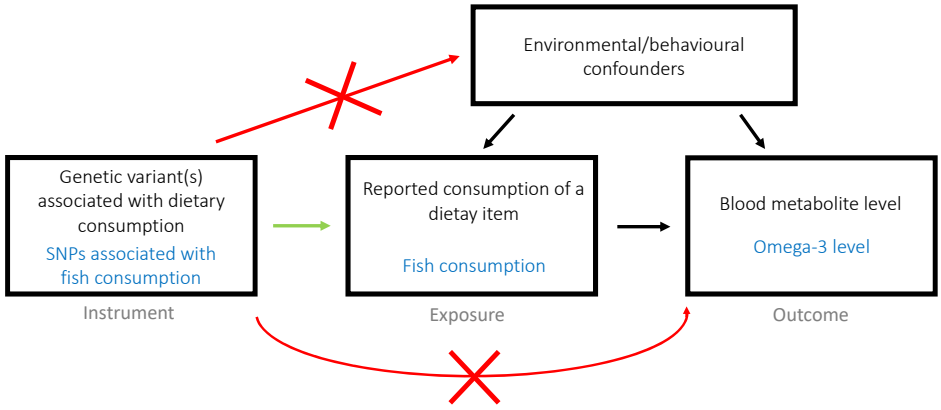

Figure 1. Schematic representation of MR-analysis and accompanying assumptions. Green arrow indicates the assumption of an association that has to be present, and red arrows indicate the assumptions of associations that are not allowed to be present. Under the boxes written in gray are the more general terms. Inside the boxes written in black are general examples in the context of this thesis and in blue are more specific examples.

In an epidemiological setting, unmeasured confounders could bias the effect estimation or create a spurious association, and consequently causality cannot be inferred. To illustrate it in simplified mathematics, let us assume, that there is a confounder C that has simultaneously an effect on exposure E and outcome O, while "error terms" u and v are independent of C:

$$E = aC + u \quad (1)$$

$$O = bC + v \quad (2)$$

It follows that:

$$aC = E - u$$

Therefore:

$$C = \frac{1}{a}E - \frac{1}{a}u$$

Substituting confounder *C* in the outcome-variable equation (2) gives:

$$O = \frac{b}{a}E - \frac{b}{a}u + v$$

From this we can see that if there were an unobserved confounder *C*, it would create an association between outcome *O* and exposure *E* and we would not be able to obtain an unbiased estimate for the direct causal effect of exposure on outcome. However, in the case of instrumental variable estimation, when the assumption of no association between the instrument and confounder is satisfied (marked with red arrow on Figure 1), then the effect of instrument on *E* is not affected by *C*. To illustrate it mathematically, let us assume that there is an instrument *Z* that is independent of *C* (confounder) and *v* (error term for *O*). When we add *Z* to (1) we get:

$$E = aC + dZ + u \tag{3}$$

which we could simplify as:

$$E = dZ + u^* \tag{4}$$

where $u^* = aC + u$ is independent of *Z*. Further, we assume that *O* does not directly depend on *Z* (marked with red arrow on Figure 1). Now, suppose that there is a causal effect of *E* on *O*, marked with g:

$$O = gE + bC + v$$

Then substituting *E* in this formula with *E* from (3) leads to:

$$O = g[aC + dZ + u] + bC + v = gaC + gdZ + gu + bC + v$$

which we could simplify as:

$$O = gdZ + w^* \tag{5}$$

where $w^* = g(aC + u + b) + v$ is independent of *Z*.

Now, when we regress *E* on *Z* (see (4)), we estimate d as $\hat{d}$, and when we regress *O* on *Z* (see (5)), we estimate $\widehat{gd}$. From these two estimates we can estimate g as:

$$\hat{g} = \frac{\widehat{gd}}{\hat{d}} \tag{6}$$

Here $\hat{g}$ is the estimated causal effect of *E* on *O* and is not confounded by *C*. Thus, instrumental variable regression allows us to estimate unbiasedly causal effect of *E* on *O* even when there is an unobserved confounder *C* affecting *E* and *O*.

As RCTs for dietary variables are often too expensive, burdensome or unethical to conduct and conventional observational studies do not allow causal inferences because of unmeasurable confounding, MR is a promising and cost-effective alternative for disentangling causal association structures. Moreover, MR can prove extremely useful as a step between an observational association study and an RCT by helping to find potentially causal relationships while ruling out the ones where there likely is none. This, on the one hand, draws attention to potentially causal associations and on the other hand helps to save resources by narrowing down the number of potential associations to be investigated with RCTs.

One example of the latter is the case of HDL-C and its potential cardio-protective effect. Epidemiological studies have shown that low HDL-C is associated with higher CVD risk (Toth *et al.*, 2013). Since this had a potential to help prevent CVD, clinical trials with HDL-C raising drugs were conducted to investigate the potentially beneficial effect of HDL-C on cardiovascular outcomes. The results of these RCTs were that despite the medications successfully raising HDL-C levels, there was no clinical benefit for cardiovascular health (Barter *et al.*, 2007; The AIM-HIGH Investigators, 2011). Later, MR studies showed that indeed, HDL-C does not seem to have any causal beneficial effect on CVD (Burgess and Harshfield, 2016; Vitali, Khetarpal and Rader, 2017). Thus, by conducting such MR study prior to clinical trial could save time, money, and individuals from possible harmful side-effects. An example where randomized trials can be unethical, but MR could still be used, is investigating the effects of alcohol or smoking on health (Rosoff *et al.*, 2019). Now, that the results of an overwhelmingly large amount of GWASs are available, MR could in several cases be pursued prior to RCT, or for assessing the possible causality of associations detected with observational studies.

In order to pursue MR, the SNPs significantly associated with the exposure variable need to be identified using a large-scale GWAS. For the SNPs to be valid instruments, three assumption need to be satisfied: "1) The variant is associated with the exposure; 2) The variant is not associated with any confounder of the exposure-outcome association; 3) The variant does not affect the outcome, except possibly via its association with the exposure" (Burgess and Thompson, 2015, p. 29).

To bring an example of these assumptions in the context of the current thesis, let us assume that we are interested whether fish consumption has a causal effect on omega-3 fatty acid levels (see Figure 1). In this case our valid instruments would be the SNPs that are: 1) associated with fish consumption; 2) not associated with any confounders of the fish consumption – omega-3 association (an example of confounder: doctor telling the person to simultaneously consume fish and take omega-3 supplement); 3) affecting the omega-3 levels only via fish consumption, not via other pathways or directly. Continuing with this example, if these assumptions hold then the instrument can be seen as a variable that randomizes large groups of individuals to be prone to consume more or less fish. Consequently, if the group of individuals who are "randomized" to consume more fish also have significantly higher omega-3 levels compared to the group that are prone to consume less fish, we can infer that fish consumption causally raises omega-3 levels. In brief, under the assumptions stated, a statistical association between the instrument and the outcome can only be present, if both, an effect of the instrument on the exposure and a causal effect of the exposure on the outcome are present.

1.4.1 Methods of Mendelian Randomization

Several methods can be used for estimating causal effects with MR. Suppose we are interested in using MR to estimate the effect of variable X on variable Y. In this example, our instruments are the SNPs associated with variable X. When there is only one SNP present as an instrument, then the effect of the exposure on the outcome can be estimated with the ratio estimator, also called Wald ratio, as (Burgess and Thompson, 2015):

$$\hat{\beta}_{XY} = \frac{\hat{\beta}_Y}{\hat{\beta}_X} \tag{7}$$

where $\hat{\beta}_{XY}$ represents the estimated causal effect of exposure on outcome, $\hat{\beta}_X$ is the regression coefficient from linear regression of the instrument on exposure and $\hat{\beta}_Y$ is the regression coefficient from linear regression of the instrument on outcome (similarly to (6)).

However, since the assumption of pleiotropy (assumption number 3) cannot be assessed when only one or two SNPs are present, there are more aspects that need to be taken in account. With a single SNP as an instrument, the MR estimator will be biased if the assumption of no pleiotropy is violated, in other words, if the instrument either has a direct causal effect on the outcome itself or on any of the confounders. Possible option to mitigate this problem is to use several instruments simultaneously and combine their effects. This way, one is not dependent on only one instrument. One popular option for using several SNPs as instruments is the Inverse-Variance Weighted (IVW) method that combines the ratio estimates (Burgess, Butterworth and Thompson, 2013). Suppose there are m instruments,

where $\hat{\beta}_{Yk}$ and $\hat{\beta}_{Xk}$ are the estimated regression coefficients for the k-th instrument from outcome and exposure regressions, respectively. To combine the estimates, the fixed-effect meta-analysis idea is used, where estimates $\hat{\theta}_k$ for parameter θ are combined as (Harrer *et al.*, 2021):

$$\hat{\theta} = \frac{\sum_k \hat{\theta}_k * w_k}{\sum_k w_k} \tag{8}$$

where w_k represents the weight of corresponding estimate $\hat{\theta}_k$ and can be written as $1/\sigma_k^2$. The variance of a ratio estimate is (Burgess and Thompson, 2015):

$$\frac{\sigma_{Yk}^2}{\hat{\beta}_{Xk}^2} \tag{9}$$

When substituting in equation (8) the $\hat{\theta}_k$ with the ratio estimate from (7) and with the corresponding variance (9), we get the IVW estimate for the ratio estimates:

$$\hat{\beta}_{IVW} = \frac{\sum_k \frac{\hat{\beta}_{Yk}}{\hat{\beta}_{Xk}} * \frac{\hat{\beta}_{Xk}^2}{\sigma_{Yk}^2}}{\sum_k \frac{\hat{\beta}_{Xk}^2}{\sigma_{Yk}^2}} = \frac{\sum_k \frac{\hat{\beta}_{Yk}\hat{\beta}_{Xk}}{\sigma_{Yk}^2}}{\sum_k \frac{\hat{\beta}_{Xk}^2}{\sigma_{Yk}^2}} = \frac{\sum_k \hat{\beta}_{Yk}\hat{\beta}_{Xk}\,\sigma_{Yk}^{-2}}{\sum_k \hat{\beta}_{Xk}^2\sigma_{Yk}^{-2}}$$

Using the IVW estimate can help mitigate the issue of a biased estimate due to pleiotropy by combining the estimates of several SNPs, but it does not eliminate the problem of pleiotropy itself. One option to overcome the latter issue is to use MR-Egger method, which tackles the bias caused by pleiotropy as similar to so-called small study bias (Bowden, Davey Smith and Burgess, 2015). In this case the intercept of the meta-analysis of several instruments is not constrained to zero. Consequently, when the intercept is estimated to be significantly different from zero, there is an indication of pleiotropy and effect estimate from MR-Egger regression should be used as the causal effect estimate (Bowden, Davey Smith and Burgess, 2015). However, although the MR-Egger estimate is more robust than the IVW estimate, it suffers from power issues, and therefore can be considered more as a sensitivity analysis method rather than the main method for MR analysis.

1.4.2 Two-sample Mendelian Randomization

The methods described above – Wald ratio, IVW, MR-Egger – can be used in the Two-sample Mendelian Randomization (2-sample MR) setting. This means that we do not need all three variables (instrument, exposure, outcome) to be available as individual-level data in one cohort. In contrary, the 2-sample MR can be

performed by utilizing GWAS summary statistics from two different non-overlapping samples, whereby the sizes of these samples do not need to be equal (Pierce and Burgess, 2013). In this case, the summary statistics from the exposure GWAS (in the current context, the dietary item GWAS) and outcome GWAS (here, a GWAS of metabolites) can be utilized to obtain a causal effect estimate of exposure on outcome (dietary item on a metabolite). To continue with the previous example of fish consumption and omega-3, the simplified workflow would be the following: detect the SNPs that are significantly associated with fish consumption; extract the effect estimates for these SNPs from the fish consumption GWAS; extract the effect estimates for the same SNPs from the omega-3 GWAS. If we were to look only at one instrument at a time, the effect estimate from the omega-3 GWAS divided by the effect estimate from the fish GWAS would be the causal estimate (Wald ratio).

As discussed in the previous section, when there is more than one instrument available, one can use the IVW 2-sample MR method, whereas MR-Egger could be used as a sensitivity analysis in case of potential pleiotropy. Furthermore, there are several other 2-sample MR methods developed that can be used as sensitivity analyses. These are for example: MR-Radial (Bowden et al., 2018), which helps to detect outlying SNPs and consequently mitigate the problem of pleiotropy; MR-Median (Bowden et al., 2016), which can estimate the effect when up to half of the information comes from invalid instruments; and MR-RAPS (Zhao et al., 2019), which can handle weak instruments. All these 2-sample MR methods can be implemented using the MR-base via the R-package TwoSampleMR (Hemani *et al.*, 2018, https://www.mrbase.org/).

As stated, an important prerequisite for performing such analyses are valid instruments. In the current context these would be the SNPs significantly associated with various dietary consumption traits in the dietary items GWAS. However, until recently there were no such studies available. To this end, we detected several SNPs associated with various dietary items in UK Biobank (UKBB) in our previous study (Pirastu *et al.*, 2019). There are several aspects that need to be considered when one wants to obtain valid dietary SNPs. Since the ideas of how to avoid possible biases when valid dietary SNPs are of interest is important in the context of current thesis, a brief summary of some methodological aspects based on Pirastu *et al*. (2019) is presented in the next section.

Lastly, another prerequisite for assessing the effect of diet on blood metabolites with 2-sample MR is a GWAS for NMR metabolites. For this we use the summary statistics of the GWAS of 123 NMR-metabolites by Kettunen *et al.* (2016).

1.4.3 Valid instruments for dietary traits

MR has been successfully implemented in nutritional epidemiology for assessing the possibly causal effects of several dietary phenotypes, such as for example the consumption of alcohol (Chen *et al.*, 2008; Andrews, Goate and Anstey, 2020;

Larsson *et al.*, 2020), coffee (Nordestgaard, Thomsen and Nordestgaard, 2015; Lee, 2018) and milk (Bergholdt, Nordestgaard and Ellervik, 2015; Yang *et al.*, 2017). However, up until recently there was a lack of large-scale GWASs of dietary traits. This is an issue also highlighted in a review assessing MR scenery in nutrition research (Larsson, 2021). To fill this void, two large studies reporting GWAS significant SNPs of a wide range of dietary items in the UKBB were recently published (Pirastu *et al.*, 2019; Cole, Florez and Hirschhorn, 2020). Besides reporting a variety of dietary SNPs and assessing subsequently the effect of diet on health, both of these studies detected reverse causality from health to diet, such as CHD and education affecting dietary choices. An important difference in the methodology of these two papers is that Pirastu *et al*. (2019) accounted for the bias originating from reverse causality by correcting the effect estimates for potential health-related confounders.

When identifying SNPs associated with dietary items, there are several pitfalls to account for. Firstly, the biases resulting from self-reported dietary data, such as recall bias and reporting bias, and lack of accuracy. And secondly, mediation of instrument-exposure relationship by health-related traits that dietary advice is given to. Moreover, these mediating traits could potentially affect the reporting of dietary consumption while not affecting actual consumption. This would mean that we would obtain an invalid instrument, since the SNP would actually not be associated with the dietary consumption trait (Pirastu *et al.*, 2019). In order to mitigate these issues, Pirastu *et al*. (2019) proposed the idea of corrected-to-uncorrected ratio (CUR). The main goal of using CUR is to try to identify the SNPs that have only the direct effect on the dietary traits of interest and are not mediated via the traits that cause changes in the reporting of dietary consumption.

The idea of CUR is that if a SNP is directly associated with a dietary item, then its effect estimate should not alter when adjusting for potential confounders. For the latter, the following list was considered: BMI, diastolic and systolic blood pressure, T2D, CHD, ulcerative colitis, Crohn's disease, educational attainment, LDL-C, HDL-C and triglycerides. Further, the ratio between the adjusted (corrected) estimate and raw (uncorrected) estimate is expressed as CUR. When this ratio is close to 1, it means that the corrected and uncorrected effects do not largely differ and therefore the SNP effect of dietary item is direct and such SNP can thus be considered a valid instrument (Pirastu *et al.*, 2019). Furthermore, the issue of recall bias with Food Frequency Questionnaire (FFQ) data has been previously tackled by Bradbury *et al.,* (2018) in UKBB, showing that FFQ was able to reliably rank individuals' intakes of main food groups.

Thus, using the results from Pirastu *et al*. (2019) and Kettunen *et al.* (2016) in a 2-sample MR setting, the possibly causal effect of diet on blood metabolites can be detected. In addition, by observing the effect of FN on blood metabolites, the effect of a systematic lack of dietary variety can be observed. When elaborating further, blood metabolites can be indicative of diseases or general health status. The latter can be examined when assessing whether blood metabolites can be used for describing individual's biological age and whether this BA is indicative of health outcomes.

1.5. Biological age and omics aging clocks

Age as a phenotype that is inevitably affecting every human being, is a risk factor for a wide range of diseases and health outcomes. Several observable changes often accompany higher chronological age, such as lowered physical strength, greying hair, worsening posture and wrinkles. Even though every person encounters such changes, these phenotypes vary notably between individuals with same chronAge. Further, it has been shown that aging is associated as well with molecular markers, such as cellular senescence, telomere shortening, loss of proteostasis, and genomic instability (López-Otín *et al.*, 2013). Much alike as the phenotypes that person can easily observe, these molecular changes vary between individuals with same chronAge. Based on the fact that individuals with the same chronAge vary greatly in terms of age-related diseases and mortality, it has been hypothesized that there is an underlying BA, that is characterized by molecular hallmarks and that affects age-related disease risk (Levine *et al.*, 2018). Consequently, BA might be even more indicative of disease risk than chronAge, and unlike chronAge it might be reversible (Horvath *et al.*, 2020; Noroozi *et al.*, 2021).

Several endeavors using a variety of biomarkers and statistical methods have been pursued in order to construct models that could describe the BA. These models are often called the omics clocks (OCs). From a methodological point of view, OCs that aim to track BA are usually trained on chronAge, while the goal is to identify the set of omics variables with corresponding weights/coefficients that best predict chronAge. This is often done using machine learning methods, more precisely penalized regression methods. One of the most popular approaches for constructing an omics clock is the elastic net regression (Zou and Hastie, 2005) that combines the ridge regression and the lasso regression methods. To put it in a nutshell, assessing BA with an OC built with elastic net regression serves the aim of detecting a set of biological measurements and corresponding parameters that would result in a BA estimation that would correlate strongly with chronAge.

As stated, such models using various types of omics data to estimate BA are called omics clocks. Consequently, the BA estimates can be called omics clocks ages (OCAs). The first OCs describing BA were the epigenetics clocks by Hannum (Hannum *et al.*, 2013) and Horvath (Horvath, 2013). These were based on the DNA methylation (DNAme) data across the genome. When it was shown that the Horvath's clock associates with mortality (Marioni *et al.*, 2015), it was confirmed that such clock is a meaningful measure of BA. After the first epigenetics clocks by Hannum and Horvath were published, several other OCs that aim to track BA have been designed using various types of omics data, such as telomere length (Zhang *et al.*, 2014; Jansen *et al.*, 2021), proteomics (Enroth *et al.*, 2015; Lehallier *et al.*, 2019, 2020; Jansen *et al.*, 2021), glycomics (Krištić *et al.*, 2014) and recently as well metabolomics (van den Akker *et al.*, 2020; Jansen *et al.*, 2021). The study by van den Akker *et al.* (2020) used the data of 56 NMR metabolites of 18 716 individuals from 24 community- and hospital-based

cohorts from Dutch Biobanking and BioMolecular Resources and Research Infrastructure in the Netherlands (BBMRI.nl) to train OC named metaboAge. The correlation between the predicted metaboAge and chronAge was estimated in the test-set to be around 0.65 (van den Akker *et al.*, 2020). The study by Jansen *et al.* (2021) used the data of 231 NMR metabolites of 2910 individuals from the Netherlands Study of Depression and Anxiety cohort to train a metabolomic OC that was estimated to have a correlation coefficient of 0.7 with chronAge. The performance of neither of these clocks have yet been assessed outside of initial study populations, and current author is not aware of any other published OCs based on NMR blood metabolites.

Since BA can be measured using various different omics assays, which can in turn be indicative of different bodily systems, it has been pondered whether there is one single BA for a person measured with varying accuracy depending on the omics data or whether there are actually several different BAs that refer to different bodily systems (Belsky *et al.*, 2017; Jylhävä, Pedersen and Hägg, 2017; Cole *et al.*, 2019; Jansen *et al.*, 2021). One option to assess whether an OC refers to specific aspects of BA is to assess the difference between OCA and chronAge, which we refer to as omics clock age acceleration (OCAA). Thus, if a person's OCA is 45 and chronAge is 40, then the OCAA is 5 – therefore indicating that the person has similar functional capacity and age-related disease risks as the average 5 calendar years older individual. However, if different OCs track different underlying BA, then this comparison would only hold for that specific bodily system.

Further, if there are several underlying BAs, then looking at the correlation between OCAAs of different OCs would tell us whether two OCs describe BAs referring to same or different bodily systems. In theory it is possible for example, that one OC refers more to cardiovascular health, while other refers more to general immune system. For example, the difference between NMR-metabolomics based metaboAge and chronAge was associated with several cardiometabolic risk factors and diseases, such as BMI, total cholesterol, HDL-C, Triglycerides, alcohol usage, high-sensitivity C-reactive protein (CRP), prevalent metabolic syndrome, prevalent T2D, incident cardiovascular events, vascular mortality, and all-cause mortality (van den Akker *et al.*, 2020). Yet, it is important to note that the incident cardiovascular outcomes and mortality were assessed in a cohort of elderly adults (70–82 years) at risk of CVD (van den Akker *et al.*, 2020) and therefore warrant further confirmation in other cohorts prior to generalizing to wider public. Furthermore the metabolomic aging (excess age of metabolomics clock over chronAge) in Jansen *et al.* (2021) was found to be associated with BMI, cardiometabolic diseases and metabolic syndrome. Thus, based on current evidence, NMR metabolites-based clock is potentially indicative of cardiometabolic-related health and aging.

The comparison of other OCs with NMR-based OC is very limited. To date there is only one study comparing NMR-based clock with BAs estimated based on telomere length, DNAme, gene expression and proteomics (Jansen *et al.*, 2021). They observed a small correlation of 0.19 between metabolomic and proteomic

aging, whereas correlations with other OCs were negligible. Furthermore, studies have shown that the correlations between OCAAs of different OCs tend to be rather modest, suggesting that they are indicative of different aspects of aging (Belsky *et al.*, 2017; Jansen *et al.*, 2021). However, it has remained unanswered how NMR-based OC compares with a wider range of OCs, and how the NMR-based OCs perform in other populations. Further, since there is limited evidence assessing the effect of OCAA of NMR-metabolomic clock on health outcomes, studies confirming previous findings and detecting new aspects are instrumental for assessing the utility of such clock as an indicator of the health of specific bodily systems.

1.6. Future directions

The developments of recent years have brought countless options for assessing the interplay between diet, blood metabolites, and health. With biobanks continuously growing larger and gathering more data, subsequently allowing GWASs with larger samples, several dietary GWASs have been recently published (Cole, Florez and Hirschhorn, 2020; Meddens *et al.*, 2021; May-Wilson *et al.*, 2022; Merino *et al.*, 2022). Consequently, this opens new options for following MR studies regarding the effect of diet.

Furthermore, the affordability of metabolite profiling opens up a wide range of research options. NMR is rather cost-effective way (Würtz *et al.*, 2017) (current price 18€ per sample, https://research.nightingalehealth.com/key-benefits-nightingale) of quantifying the metabolic profile and can thus prove useful when tracking changes in metabolome or assessing current health status is of interest. In such way the NMR metabolic profile is utilized for example by Nightingale Health (https://nightingalehealth.com/). Metabolic profile might have the potential in tracking the effect of dietary changes on health. For example, a person with low dietary variety could receive blood metabolic analysis prior to starting treatment (diet change) and after following new diet for 6 months. In this way the potential benefit of dietary changes could be better quantified and such data can be utilized for further research that aim to enhance people's health via dietary changes. On the other hand, seeing the actual measured benefit in blood metabolic profile could motivate people to follow dietary advice. Thus, assessing the effect of diet on NMR metabolites can provide basis for future intervention studies and dietary recommendations.

The studies regarding FN have been with rather limited sample size. Probably this is the reason why to date no GWAS reporting the results of FN is available. Since RCTs assessing the effect of FN on health are not feasible due to the nature of it, the GWAS of FN would have much utility in further investigating the possibly causal effects of FN on health using 2-sample MR. The validated FNS questionnaire is rather short and simple with only 10 statements that have to be rated on a 7-point scale. Moreover, FN has shown to be highly heritable trait (Cooke, Haworth and Wardle, 2007; Knaapila *et al.*, 2007; Faith *et al.*, 2013),

which proves the potential of detecting GWAS significant hits. Adding the FNS questionnaire to the data measured in any of the large biobanks, such as UKBB or Estonian Biobank (EstBB), GWAS and subsequent downstream analyses would become possible. This would help to better assess how severe the effect of FN is on health and could possibly highlight the need to treat FN.

Another exciting field is the one of OCs. Especially of interest are the associations between OCAA and dietary traits. Even though such studies are scarce, there are a few papers reporting that OCAA of DNAme-based clocks is associated with fish intake and blood carotenoid levels (indicator of fruit and vegetable intake) (Quach *et al.*, 2017), and heavy alcohol consumption (Beach *et al.*, 2015; Luo *et al.*, 2020). Moreover, even more enthralling is the question whether it is possible to reduce BA with dietary interventions. This question has been recently tackled with an RCT showing that one-year Mediterranean diet was able to significantly reduce BA measured by DNAme-based Horvath's ageing clock (Gensous *et al.*, 2020). Thus, there is evidence of diet affecting BA. However, current author is not aware of any studies assessing how diet affects NMR-metabolomics based BA and corresponding OCAA. Since NMR-metabolomics based OCs have been associated with cardiometabolic risk factors and diseases (van den Akker *et al.*, 2020; Jansen *et al.*, 2021), there might be considerable scope in assessing the interplay between diet and NMR-based BA. Of greatest interest would be whether it is possible to reduce such BA with dietary interventions.

In summary, the developments in -omics research can potentially contribute to better understanding of the causal effects of dietary choices on health, removing some of the existing controversies and paving paths towards more efficient dietary interventions.

2. AIMS OF THE STUDY

The aim of this thesis is to examine the interplay between diet, blood metabolites, and health in order to shed more light to the possible mechanisms of how dietary choices affect health.

To fulfill that aim, the specific objectives of the thesis are:

- To establish possibly causal relationships between dietary traits and blood metabolites, using the framework of Mendelian Randomization
- To investigate how systematic lack of dietary variety expressed as food neophobia affects blood metabolic profile, type II diabetes and coronary heart disease.
- To assess to what extent the biological age constructed based on blood metabolic profile correlates with chronological age, and whether the excess of such biological age over the chronological age is predictive of specific disease groups and subsequently indicative of the health of a specific bodily system. Furthermore, how the biological age based on blood metabolites compares to the biological ages that are based on proteomics and DNA methylation data

3. RESULTS AND DISCUSSION

3.1. Mendelian randomization identifies the potential causal impact of dietary patterns on circulating blood metabolites (Ref. I)

The aim of this paper was to detect the potentially causal effects of dietary consumption traits on blood metabolic profile measured by NMR. The main method of this paper was the 2-sample MR, which utilized the data of two different large GWASs. The results of this study provide grounds for future RCTs that aim to establish the effect of various dietary items on blood metabolome.

3.1.1 Description of cohorts and methods

The MR method is more thoroughly described in chapter 1.4. Briefly, the instruments for using in MR were obtained from a GWAS of 25 individual and 15 principal-component (PC) dietary traits performed in UKBB (Pirastu *et al.*, 2019). The estimates for the same SNPs on the outcomes (the blood metabolites) were obtained from a large GWAS of 123 NMR blood metabolites performed in Estonia and Finland (Kettunen *et al.*, 2016). The workflow of performing the analysis for this paper is described in Figure 2 (adapted from Ref. I Figure 1). This figure presents the relevant steps performed and reported in previous studies (Kettunen *et al.*, 2016; Pirastu *et al.*, 2019), and the main 2-sample MR methods along with sensitivity analyses methods performed in Ref. I.

The valid instruments were available for 25 individual dietary traits: consumption of beef, pork, processed meat, poultry, lamb, beer, champagne/ white wine, red wine, spirits, tea, ground coffee, decaffeinated coffee, instant coffee, water adjusted for coffee, bread, cheese, cooked vegetables, salad, dried fruit, fresh fruit, non-oily fish, oily fish, salt, and vegetarianism and drink temperature. Since dietary items are often consumed together, we also looked at 15 PC-traits that were obtained in Pirastu *et al.* (2019). To obtain PCs, items were first clustered together using iCLUST hierarchical clustering algorithm (Revelle, 1979; Revelle and Zinbarg, 2009) on genetic correlation matrix of the dietary traits. Thereafter, the clusters were formed based on splitting the resulting tree dendrogram into several layers. The PC-traits obtained and used for later analysis were: Fish PC1, Fruit PC1, Vegetables PC1, Alcohol PC1-2, Coffee PC1, Meat PC1, Healthy PC1-3, Psychoactive PC1-2, All PC1-3. Some examples of the composition of the PC-traits: oily fish and non-oily fish were clustered together as Fish PC1 describing the overall fish consumption, while different categories of coffee and alcohol were clustered together as psychoactive drinks. The composition of each PC-trait can be viewed in Ref. I Figure 2. Due to splitting the tree dendrogram to several layers, one item can participate in PCs from different layers. For example, beer was included in Alcohol PCs, Psychoactive PCs and All PCs (the latter includes all dietary items of interest).

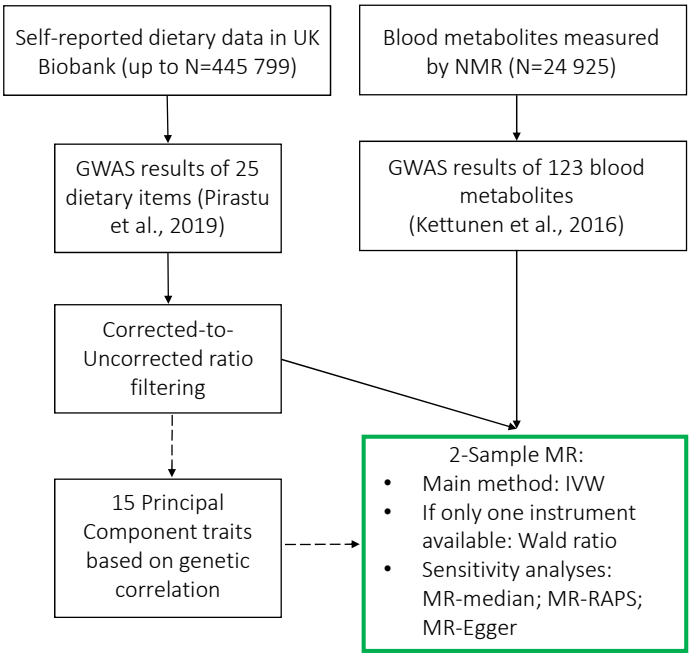

Figure 2. Workflow and selection of instrumental variables. Black boxes describe the relevant steps performed and reported in previous studies (Kettunen *et al.*, 2016; Pirastu *et al.*, 2019) and green box describes the 2-sample MR main methods and sensitivity analyses pursued in Ref. I.

After defining groups based on clustering, the next step was to perform PC-analysis for each group in order to obtain the SNP effect estimates for the PC-traits. For that, the PC-analysis was applied on the genetic correlation matrix of the traits that comprised the corresponding group. Subsequently the SNP effects for the PC-traits were estimated based on the SNP effects on each of the components and corresponding loadings from PC-analysis.

The loadings of the PC-traits usually help to interpret the meaning of the corresponding PC. For the main PC traits of interest, the loadings can be observed from Ref. I Figure 3. For many of the PC-traits the loadings are with the same direction and self-explanatory. For example, the Alcohol PC1 has positive loadings from all alcohol-related traits. This means that if a SNP has a positive effect on any of the alcohol traits, it has also a positive effect on the Alcohol PC1. However, some PCs were more difficult to interpret. For example, All PC1 has positive loadings from fish, fruits, salad, vegetables and negative loadings from different meats, coffees and alcohols. Therefore, one could interpret All PC1 as partly describing healthy-unhealthy axis of dietary consumption, with higher values referring to more healthy diet in general.

The GWAS results of the traits described above provided instruments for pursuing 2-sample MR. More precisely, for using as instruments we chose the independent ($r2<0.001$) SNPs that were significantly (p-value $< 5 \times 10^{-8}$)

associated with the traits in the dietary (the exposure) GWAS (Pirastu *et al.*, 2019), whereas additionally the CUR=1±0.05 filtering was applied (see section 1.4.3 for description of CUR). The effect estimates for the same SNPs for the metabolites (the outcomes) were obtained from the GWAS results of 123 metabolites in 24 925 individuals (Kettunen *et al.*, 2016).

We used IVW method as the main method for the 2-sample MR analysis (see section 1.4.1 for description of MR methods). To address the issue of horizontal pleiotropy, we used the method of MR-Radial (Bowden *et al.*, 2018). As sensitivity analyses we used MR-Median (Bowden *et al.*, 2016), MR-RAPS (Zhao *et al.*, 2020), and MR-Egger (Bowden, Davey Smith and Burgess, 2015). In a situation where only one instrument was available, we used the method of Wald ratio.

We corrected the results for multiple testing by using the false discovery rate (FDR<0.05) via Storey's q-values (Storey, 2003). After this correction, 413 potentially causal relationships remained statistically significant. Most of the significant effects of dietary items on blood metabolites were related to atherogenic lipoproteins: very-low density lipoproteins (VLDL), intermediate-density lipoproteins (IDL), and low-density lipoproteins (LDL). These all contain Apolipoprotein B, a particle whose higher amount is considered to be indicative of elevated CVD risk (Leslie, 2017; Ference *et al.*, 2019). The heatmaps describing the results of the metabolites that showed significant associations with the food items or groups are depicted in Ref. I Figure 4 and Ref. I Figure 5. In these figures, darker red corresponds to stronger positive association whereas darker blue corresponds to stronger negative association. The FDR-significant results are marked with "*" in the middle of the corresponding square. The most interesting findings are described in the following subsections.

3.1.2 Alcohol and coffee

Very clear patterns emerged regarding alcohol and coffee (Ref. I Figure 4). It can be noted that for most of the atherogenic lipoproteins measurements alcohol and coffee posed similar elevating effects, especially with regards the components of IDL and LDL particles and serum total cholesterol. However, the association patterns were clearly distinct for VLDL-related measurements: the components of medium VLDL and small VLDL particles were significantly affected by coffee consumption, whereas no clear effect by alcoholic beverages were observed. Moreover, even the effect directions of the two groups of drinks were not coherent, indicating that this difference is not due to power issues. In this case Coffee PC1 shows elevating effect, whereby the betas of all the coffee subgroups agree with the direction of the effect. One reason for such effect of coffee on VLDL-related measurements could be that cafestol – a common ingredient in coffee – has been shown to increase VLDL particle assembly rate in the liver (de Roos *et al.*, 2001). Therefore, this mechanism could be specific to coffee, while not applying to alcohol. Furthermore, our results replicate the previously consistently

shown LDL-C elevating effect of coffee (Poole *et al.*, 2017). In contrary, the effect of coffee and alcohol on IDL particles – where we saw consistent strong elevating effects – has previously not been shown, and therefore our results provide basis for future research aiming to understand the effect of coffee and alcohol on atherogenic lipoproteins.

An interesting outcast appeared when examining the results of alcoholic beverages and levels of high-density lipoprotein (HDL) related measurements and 22:6 docosahexaenoic acid (DHA) (Ref I Figure 5). Namely, beer was clearly the odd one out compared to other alcoholic beverages by having significant positive effects on several HDL-related measurements and DHA. This pattern was especially clear in terms of DHA, which is a subgroup of omega-3 fatty acids. DHA has been shown to possess anti-inflammatory properties (Calder, 2010) and be protective of CVD (Bernasconi *et al.*, 2021). While DHA is considered beneficial for cardiovascular health, the evidence regarding the effect of HDL-related measurements is conflicting. Several epidemiological and observational studies have proposed that moderate alcohol intake could have a cardioprotective effect (Brien *et al.*, 2011; Ronksley *et al.*, 2011) and this positive effect has mostly been associated with elevated HDL-C levels (Brien *et al.*, 2011). Even though our results seem to indicate a positive impact of alcohol on HDL-related measurements, with the effect of beer being statistically significant, it is currently not possible to confirm that this would have a beneficial outcome on cardiovascular health since the results regarding the actual benefit of higher HDL-C on cardiovascular health are conflicting (Briel *et al.*, 2009). While recent epidemiological studies demonstrate that HDL-C has a U-shaped association with CVD, the actual role of HDL in the context of CVD is unclear (Casula *et al.*, 2021; Yi *et al.*, 2022).

It is important to note that our results regarding alcohol and the lipid profile agree with another recent MR-study (Rosoff *et al.*, 2019), while elaborating the findings further in terms of alcohol subgroups. In addition, our findings conflict with some observational studies (Würtz *et al.*, 2016; Du *et al.*, 2020) with regards to the effect on IDL-related measurements.

3.1.3 Vegetarianism

Our results showed that vegetarianism raises the levels of omega-6 fatty acids and 18:2 linoleic acid (LA), which is a subgroup of omega-6 fatty acids. Of note, these results were based only on one instrument and therefore one should remain cautious with a definite interpretation. However, the effect of vegetarianism on omega-6 fatty acids confirms a previous observational finding by Kornsteiner, Singer and Elmadfa (2008), and expands the knowledge further by proposing causality. Moreover, the elevating effect of vegetarianism on 18:2 linoleic acid has been previously shown with a randomized clinical trial on a subset of individuals with T2D (Kahleova *et al.*, 2013), whereas our findings propose that this association holds as well for the general population.

3.1.4 Discussion

The aim of Ref. I was to detect potentially causal relationships between diet and blood metabolites. A total of 413 detected relationships remained significant after FDR-correction. These findings provide a good starting point for future RCTs that aim to establish causal effects of diet on health. In addition, our results demonstrate the utility of MR in the context of dietary studies by replicating previously known causal relationships, such as positive effect of oily fish consumption on omega-3 fatty acids and DHA (Horrocks and Yeo, 1999) and the lowering effect of vegetarianism on LA (Kahleova *et al.*, 2013), while conflicting with previous observational findings of alcohol on IDL-related measurements (Würtz *et al.*, 2016; Du *et al.*, 2020). To the best of our knowledge, our results did not conflict with any previous RCTs. Consequently, our results not agreeing with some observational studies while confirming previously known causal relationships validates the utility of MR approach for dietary studies and demonstrates its strength as an intermediate step between observational studies and clinical trials.

In addition, we detected several associations that are of interest in health context. The results regarding alcohol and coffee on the levels and constituents of VLDL, IDL and LDL can be important in the context of cardiovascular health. Higher levels of atherogenic lipoprotein particles and their components is associated with elevated risk for several cardiovascular diseases (Carmena, Duriez and Fruchart, 2004; Ference *et al.*, 2017) and is therefore part of a less favorable lipoprotein profile. However, previously published results regarding the effect of coffee on cardiovascular health are controversial with studies showing that coffee either has beneficial or neutral effects (Chrysant, 2015) or that moderate consumption is not likely to cause adverse effects (Rebello and van Dam, 2013). In our results there was no indication of a beneficial effect of coffee. Thus, the total effect of coffee on cardiovascular health remains unclear, but our results suggest that a potentially harmful effect exists and it possibly is mediated by VLDL, IDL and LDL.

Coffee consists of a wide variety of components and it is plausible that some components pose harmful effects, while others are beneficial. Our results encompass mostly the lipid profile and therefore we cannot exclude the possibility of coffee having beneficial effects on metabolites measured with other assays. In any case, our results add to the ongoing discussion of the effects of coffee on health and provide further basis for future RCTs. A possible study design that is feasible due to the cost-efficiency of NMR, would be to randomize study groups as follows: participants who continue their regular coffee consumption and participants who will reduce or quit drinking coffee for a period of time. Measuring NMR metabolites before and after this period could answer the question whether the possibly causal effects we detected are indeed causal. Our results need to be confirmed by such RCTs before it is possible to give dietary advice, but our findings propose that coffee consumption should be limited in people who are at higher risk of cardiovascular diseases.

The effect of beer on HDL-related measurements and DHA is also interesting in the context of cardiovascular health. Since other alcoholic beverages did not show such positive effect on HDL-related measurements and DHA, it possibly originates from other components in beer rather than alcohol. Therefore, these results pose an interesting hypothesis for future studies to investigate whether alcohol-free beer has the same potentially beneficial effects as beer, while not having the harmful effects of alcohol.

In order to provide dietary suggestions, our results need to be confirmed with RCTs. However, our findings provide strong basis for such RCTs and have the potential to contribute to future dietary guidelines. Furthermore, we demonstrated the utility of using MR in dietary studies, which is important in the context of several new dietary GWASs published recently (Cole, Florez and Hirschhorn, 2020; Meddens *et al.*, 2021; May-Wilson *et al.*, 2022; Merino *et al.*, 2022). The study by Cole, Florez and Hirschhorn (2020) reports the GWAS results of 85 single dietary traits and 85 PC-traits corresponding to dietary patterns. A possible next step from our study would be to confirm our results by using their findings when defining the instruments for 2-sample MR. In addition, a recent study presented results of a food-liking GWAS, where they detected 1401 significant associations (May-Wilson *et al.*, 2022). Subsequently, another future direction would be to pursue the same analysis as we performed, but use food-liking SNPs as instruments instead of the ones associated with food consumption.

3.2. Food neophobia associates with poorer dietary quality, metabolic risk factors, and increased disease outcome risk in population-based cohorts in a metabolomics study (Ref. II)

The primary aim of this paper was to detect whether food neophobia associates with blood metabolic profile measured by NMR. The secondary aim was to analyze the effect of FN on disease outcomes, such as T2D and CHD.

3.2.1 Description of cohorts and methods

The analyses of this study were conducted in the Finnish Dietary Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome (DILGOM) cohort and replication was performed in the EstBB cohort. The study design and samples are described in Figure 3 (adapted from Ref. II Figure 1). More precisely, the analyses were performed on the subsamples of Finnish DILGOM (N=2982, age range 18–83) and EstBB (N=1109; age range 25–74) for which the data regarding FN and metabolic profile were present. FN was assessed with the validated FNS questionnaire (see more detailed description of questionnaire in chapter 1.2). A higher score indicates a greater level of FN. Since there are no official cut-off

scores, we divided the individuals artificially into three groups for the descriptive analysis: food neophilics (score from 10 to 24), median group (25 to 39), and food neophobics (40 to 70). The descriptive statistics for both DILGOM and EstBB cohorts by FNS categories and total can be viewed in Ref. II Table 1. Consistent with previous research (Tuorila *et al.*, 2001; Meiselman, King and Gillette, 2010; Rabadán and Bernabéu, 2021), in both cohorts the food neophobics group, compared to other groups, was with older age, lower education level and higher proportion of individuals living in rural areas.

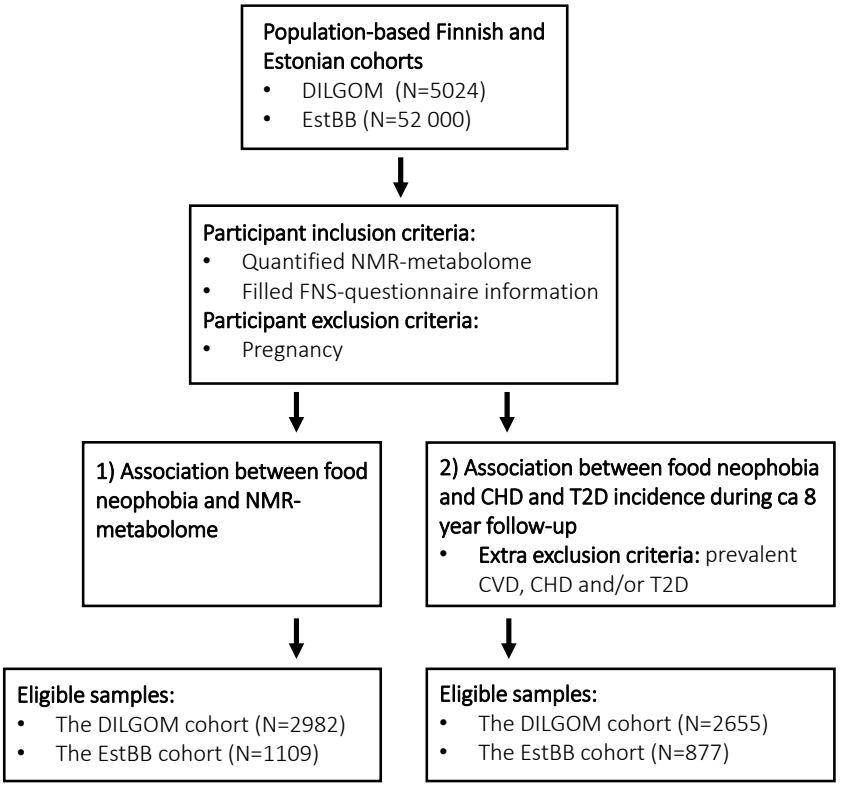

Figure 3. Flowchart of the study design and participant selection.

In this study, the same set of NMR blood metabolites as described in chapter 1.3 and analyzed in Ref. I was of interest. However, for this paper the set of variables was larger, additionally containing a wide variety of ratios describing lipoprotein compositions. The exclusion criteria for the metabolomic analysis was pregnancy. Prior to analysis, the outliers (standard deviation >4 threshold) were removed and the scaled logarithmic transformation was applied for all metabolites. Linear regression was used for assessing the association between FNS score and blood metabolite levels. All analyses assessing metabolite levels in EstBB were adjusted for following confounders: age, BMI, living region (urban/rural), education level, smoking status, prevalence of diabetes, and sex.

The disease outcomes of CHD and T2D were collected from national health registries: the Finnish Hospital Discharge Registry and the Estonian Health Insurance Fund. The end-point of interest were the incident diseases diagnosed during follow-up (average follow-up was approximately 8 years in both cohorts; see the exact follow-up times in Ref. II Table 4). The disease outcomes were analyzed using the Cox proportional hazards model, whereas the prevalent cases of T2D, CVD and CHD were excluded prior to analysis. All survival analyses in EstBB were adjusted for sex, BMI, smoking status and statin treatment, whereas age was used as the time-scale.

Bonferroni correction was used to account for the family-wise error rate. However, many of the assessed metabolites correlate highly with each other, with 24 principal components explaining >95% of the variance of the data. Therefore, we set $p<0.0021$ (0.05/24) as the threshold for the corrected p-value in the analyses of the metabolome.

3.2.2 The association of food neophobia with NMR-metabolites

We found that higher FN associates with adverse metabolic profiles. There were several significant associations between FN and metabolic profile in Finnish DILGOM cohort and in EstBB cohort after correcting for multiple testing. In DILGOM cohort, higher FN was positively associated with: the ratio of MUFA to total fatty acids, α1-acid glycoprotein and citrate; and negatively associated with: the ratio of omega-3 fatty acids to total fatty acids, cholesterol esters in very large HDL, total cholesterol in very large HDL, and the concentration of very large HDL. In EstBB cohort, higher FN was negatively associated with: the ratio of omega-3 fatty acids to total fatty acids, the level of omega-3 fatty acids, the ratio of DHA to total fatty acids, and the level of DHA. Thus, the association between the severity of FN and the ratio of omega-3 fatty acids to total fatty acids replicated in the EstBB cohort. Since DHA belongs to the group of omega-3 fatty acids, the findings that were significant in both Finnish and Estonian cohorts, are related to omega-3 fatty acid associated metabolites.

In general, more severe FN was associated with adverse cardiometabolic outcomes on lipid metabolites (e.g. lower levels of omega-3 related fatty acids in both cohorts, lower concentration of very large HDL particles in DILGOM) and inflammation-related biomarker (higher levels of α1-acid glycoprotein in DILGOM). Consequently, lower levels of FN were associated with more favorable levels of these biomarkers. Since this study is the first to broadly investigate the association between FN and NMR-metabolites, there are no other studies to directly compare the results to. However, the results can be interpreted in terms of health context.

3.2.3 The association of food neophobia with disease outcomes

We assessed the effect of FN on the incidence of T2D and CHD in both, the DILGOM and EstBB cohorts. The results of these analyses are presented in Ref. II Table 4. Higher levels of FN predicted higher incidence of T2D in the DILGOM cohort, but not in the EstBB cohort. Furthermore, higher levels of food neophobia associated with higher incidence of CHD in the EstBB cohort, but not in the DILGOM cohort. Since the effect of FN on CHD and T2D had not been researched before and our results show conflicting evidence between the two cohorts, further studies are needed in order to establish the effect of FN on T2D and CHD risk.

3.2.4 Discussion

In general, our analysis demonstrated that FN has an adverse effect on metabolic profile, and potential risk-elevating effect on T2D and CHD incidence. However, since the results regarding disease associations were conflicting between the two samples analyzed, further studies are needed before firm conclusions can be made.

We observed significant associations consistently between FN and several omega-3 fatty acids related measurements. One explanation for this could be that FN has been associated with lower adherence to the Mediterranean diet (Predieri *et al.*, 2020) and less frequent intake of fish (Helland *et al.*, 2017). Both of these have been shown to associate with altered levels of omega-3 fatty acid and its subgroup DHA (Horrocks and Yeo, 1999; Mantzioris, Muhlhausler and Villani, 2022). The positive effect of higher fish consumption on omega-3 fatty acids and DHA can also be observed in the results of Ref. I.

In addition, omega-3 fatty acids have anti-inflammatory properties (Calder, 2010), and supplementation with omega-3 has been found to be effective in the prevention of CHD (Bernasconi *et al.*, 2021). In contrary, the effect of omega-3 on T2D is not clear, with epidemiologic studies associating higher levels of omega-3 with lower T2D risk (Qian *et al.*, 2021), while meta-analysis of RCTs concluded that there is little or no effect (Brown *et al.*, 2019). Since there is a consistent negative association between FN and omega-3 related measurements, one could expect to see elevated CHD risk with higher FN. We witnessed this effect in EstBB, but not in DILGOM cohort. Whether there is a pathway of FN leading to lower omega-3 related measurements subsequently leading to higher risk of CHD is an interesting and important hypothesis to test in future studies.

Furthermore, negative associations with HDL-related measurements were observed in the DILGOM cohort. As discussed, HDL-C has been consistently associated with CVD in epidemiological studies, but the causal effect has not been proven and the actual role of HDL in the context of CVD is not clear (Casula *et al.*, 2021). Despite the missing causal link, higher levels of HDL-related measurements are indicative of better cardiovascular health. Moreover, higher levels of omega-3 fatty acids have been associated with lower risk of mortality (Harris

et al., 2021), indicating that lower levels of omega-3 associate with higher risk. Therefore, the negative association we observed between FN and omega-3 related and HDL-related measurements are indicative of adverse health effects of higher FN.

Despite the fact that our study was not designed in a way that would allow to provide dietary advice, it suggests that FN should be addressed and further studied in order to prevent unwanted outcomes on health. As randomized trials with FN are not feasible due to the nature of it, future perspectives regarding possibly causal effects of FN on health will become possible when FN is recorded in larger biobank samples, for example entire UKBB or entire EstBB. Thereafter the FN GWAS would be possible and would open the option to pursue MR studies assessing the effect of FN on blood metabolites, CHD and T2D. This would grant us a better indication of the possibly causal effect of FN on health and disease.

3.3. A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk (Ref. III)

The general aim of this study was to investigate whether there is one single biological age for individuals, which could be tracked with different omics clocks with varying accuracy, or whether different OCs are tracking different underlying BAs. This paper considers a wide range of OCs: fifteen clocks, out of which eleven were newly developed and four were previously published. However, in the current thesis the ones that were calculated also on EstBB cohort are discussed, and specifically of interest is the one related to NMR metabolites. The latter is based on the same platform of blood metabolites as Ref. I and Ref. II. Consequently, the aims of this paper in the context of current thesis are: to assess whether the BA measured by NMR-metabolites based OC correlates with chronAge; to investigate whether NMR-based OC tracks the same BA as other OCs; to study whether the difference between the NMR-based BA and chronAge is predictive of diseases and risk factors; and to determine whether the associated health outcomes are related to some specific bodily system. These findings could further elaborate the knowledge of the impact of diet on health via the changes in the blood metabolome.

3.3.1 Description of cohorts and methods

The simplified workflow is depiceted on Figure 4. Very broadly, eleven out of the fifteen omics clocks presented in Ref. III were developed in Orkney Complex Disease Study (ORCADES) cohort using elastic net regression, with the mixing parameter of $\alpha = 0.5$. The clocks were trained on 75% of the sample (training sample) and evaluated on 25% of the sample (testing sample). Tenfold cross validation was used in the training sample to select the shrinkage parameters λ

for the penalized regression. Four out of the fifteen OCs presented in Ref. III were calculated based on previously developed OCs to enable comparison with already established OCs. These four are: Hannum 2013 (Hannum *et al.*, 2013), Horvath 2013 (Horvath, 2013), MetaboAge (van den Akker *et al.*, 2020) and GlycanAge (Krištić *et al.*, 2014).

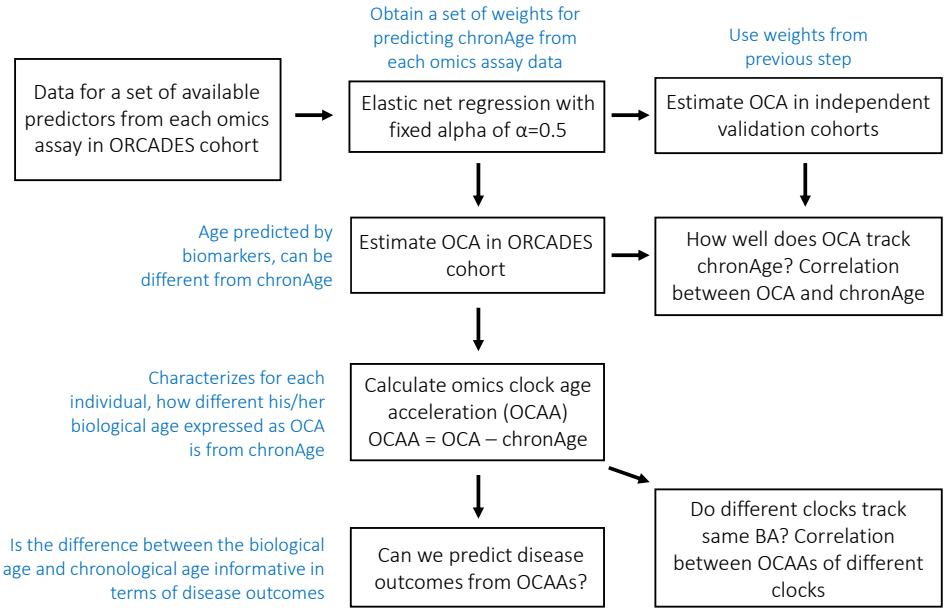

Figure 4. Description of the general workflow. General steps of the analysis workflow are described in black boxes connected with arrows. In blue are additional explanations about corresponding step.

The performance of some OCs developed on ORCADES cohort was subsequently validated in independent validation cohorts, among others also in EstBB. More specifically, EstBB was utilized to validate the following clocks: "PEA Proteomics" on a subset of proteins, measured with Olink CVDII (cardiovascular 2), CVDIII (cardiovascular 3), INF1 (inflammation 1) and ONCII (oncology 2) panels; "DNAme Horvath CpGs"; "DNAme Hannum CpGs"; and "NMR Metabolomics". Of note, DNAme Horvath CpGs and Horvath 2013 refer to different OCs, whereby Horvath 2013 is calculated using the already established published clock (Horvath, 2013), while DNAme Horvath CpGs is a DNAme-based clock developed on ORCADES cohort while using the CpG sites of Horvath's epigenetic clock as potential set of predictors. The same holds for Hannum DNAme CpGs and Hannum 2013. The validation procedure in EstBB consisted of estimating OCA for every individual in EstBB who had corresponding data available. This was carried out using the weights obtained from the elastic net regression that was pursued in ORCADES cohort. Thereafter the performance of each of the four clocks was assessed by calculating the Pearson

correlation coefficient between the OCA and chronAge of the sample, and inspected visually with scatterplots of OCA and chronAge. The same assessment of performance was conducted in ORCADES cohort testing sample.

3.3.2 The performance of the omics clocks in ORCADES and EstBB

For each of the four clocks validated in EstBB the sample sizes and the correlations between chronAge and OCA in ORCADES and in EstBB are presented in Table 2 (adapted from Ref. III Table 1, where characteristics of all other OCs in ORCADES can also be observed. For visual inspection of the performance of the OCs in ORCADES see Ref III. Figure 1). It can be seen that the proteomics- and DNAme-based clocks, which correlated extremely strongly with chronAge in ORCADES (r from 0.93 to 0.96), had comparable performance in EstBB (0.91 to 0.98). This indicates that proteomics and DNAme data can objectively estimate chronAge with very high accuracy. However, the NMR metabolomics clock, which had rather strong correlation (r=0.74) with chronAge in ORCADES, did not perform well in EstBB (r=0.26), possibly meaning that NMR metabolomics clock is more cohort-specific compared to for example DNAme- and proteomics-based clocks.

Table 2. **Description of omics clocks that were validated in EstBB cohort.**

	ORCADES		**EstBB**	
	N	**r**	**N**	**r**
NMR metabolomics	1643	0.74	6704	0.26
PEA proteomics	805	0.93	247	0.91
DNAme Hannum CpGs	1033	0.96	282	0.98
DNAme Horvath CpGs	957	0.93	229	0.97

N corresponds to number of individuals that the omics clock was assessed with. r corresponds to the correlation between omics clock predicted age and chronological age

Assessing the correlation between OCA and chronAge can provide a measure of how well the clock performs, but achieving extremely strong correlation might not always be the ultimate goal. One of our aims was to assess whether the OCAA, that describes the difference between OCA and chronAge, is predictive of diseases and risk factors. An OC that is extremely accurate in predicting chronAge has consequently OCAA with very small variability. This means that individuals with same chronAge would not differ much in their OCA. Consequently, such OCAA could have too little variation for predicting health outcomes, whereas OCAA with a slightly larger variability might convey more information for that purpose.

Consecutively, the next aim was to assess whether there is one underlying BA that is tracked with varying accuracy, or whether different OCs track BAs

referring to different bodily systems. Observing the correlation of the OCAAs of different OCs could provide a measure to what extent they describe similar BA.

3.3.3 Correlations between the omics clocks

The correlation structure between the OCAAs of all fifteen clocks assessed in Ref. III are presented in Ref. III Figure 3. It can be noted that the age acceleration of NMR-based OC correlates weakly with those of proteomics- and DNAme-based OCs (r up to 0.28), indicating that they describe different underlying BAs. A plausible reason behind this can be the earlier described low variability of the OCAAs of proteomics- and DNAme-based clocks. In contrary, the OCAA of the NMR-based OC correlates moderately to strongly with the OCAAs of all other investigated clocks (r ranging from 0.5 to 0.7). This indicates that despite NMR-based OC describing different underlying BA compared to proteomics- and DNAme-based OCs, it describes partly the same underlying BA with several other clocks. The strongest correlation of NMR-based OCAA is with MS Fatty Acid Lipidomics OCAA. Since NMR-metabolomics platform comprises mostly of lipid measurements and MS Fatty Acid Lipidomics describes plasma lipidome, the correlation of the OCAAs of these two OCs is plausible.

The correlations between OCAAs of different OCs can provide a measure to examine whether they describe similar of distinct underlying BA. However, one of the utilities of such ageing clocks could be to provide information about a person's health when we know whether his/her OCA is considerably younger/older compared to chronAge. Therefore, it is of interest to detect whether the OCAAs are predictive of diseases and risk factors.

3.3.4 OCAA in disease prediction, comparison with chronAge

The effect of OCAAs was assessed on those diseases and risk factors that were *a priori* thought to associate with age. For the disease outcomes this comprises of the ICD-10 blocks C (Neoplasms), E (Endocrine, nutritional and metabolic diseases), I (Diseases of the circulatory system) and J (Diseases of the respiratory system). For the disease risk outcomes this comprises of BMI, systolic blood pressure (SBP), CRP, total cholesterol, cortisol, creatinine, and expiratory volume in 1 second. The disease outcomes are measured as hospitalization with corresponding disease in the ORCADES cohort.

In order to observe how OCAA compares to chronAge, the first step was to assess the effect of chronAge on the chosen outcomes. Thereafter, the diseases and risk factors that did not associate significantly with chronAge were filtered out (FDR>10%). Additional exclusion criteria of having less than 5 incident cases was applied for the disease outcomes. After such exclusions, 32/44 disease groups (see list of groups analyzed in Ref. III Supplementary Table 4) and all risk factors

were taken forward to the next step, which was analyzing the effect of OCAA on these outcomes. All the OCAA models were adjusted for chronAge and sex.

The results of the relative effects of OCAAs on disease outcomes in ORCADES cohort are described in Ref. III Figure 5. It should be noted that overall there were not many associations that would have passed FDR-corrections, whereas the NMR metabolomics OCAA did not have any associations below the FDR-correction threshold. However, there were several interesting associations with p-value lower than the nominal significance level ($p<0.05$). Namely, the risk elevating effect that NMR-based OCAA had on the disease groups of "E50-E64 Other nutritional deficiencies", "E10-E14 Diabetes mellitus", and "C15-C26 Malignant neoplasms of digestive organs". This indicates that the NMR-based OC tracks partly the BA that is related to dietary behavior.

The associations of OCAAs with disease risk factors are depicted on Ref. III Figure 6. Since the NMR metabolomics platform contains the measurements of total cholesterol and creatinine, the association of NMR metabolomics OCAA with these two risk factors was not assessed. Among other risk factors, NMR Metabolomics OCAA was significantly associated after FDR-correction with CRP, BMI and SBP. The effect of OCAA on CRP was notably above 1, meaning that one year of OCAA had larger effect than one year of chronAge. All these three risk factors are known to be associated with dietary choices (Appel *et al.*, 1997; Neale, Batterham and Tapsell, 2016), further supporting the idea that NMR Metabolomics clock tracks the BA that is partly related to dietary intake.

3.3.5 Discussion

In Ref. III, fifteen OCs were presented, whereas for four of these the validation was sought in EstBB. Of particular interest in the context of this thesis is the NMR metabolomics OC that uses the same platform of metabolites as Ref. I and Ref. II. We found that the NMR metabolomics OC correlated strongly with chronAge in ORCADES cohort ($r=0.74$), but not in EstBB cohort ($r=0.26$). Contrary to that, the PEA proteomics, DNAme Hannum CpGs and DNAme Horvath CpGs clocks correlated extremely strongly with chronAge in both ORCADES and EstBB (all correlations >0.9). This raises the question whether the NMR metabolomics OC is more population specific. The same phenomenon can be observed when we look at the performance of the previously developed metaboAge (van den Akker *et al.*, 2020) in ORCADES cohort (Ref. III Table 1). The correlation between metaboAge OCA and chronAge is 0.21 in the ORCADES cohort, whereas in the original metaboAge paper the corresponding correlation in the study sample was estimated around 0.65 (van den Akker *et al.*, 2020). These results indicate that one must be cautious when interpreting the NMR metabolomics OC results, since it might reflect partly the peculiarities of specific cohorts. This can be especially the case with ORCADES cohort, since it represents a population isolate.

The specific reason behind NMR metabolites-based ageing clocks not showing good replication yet warrants further investigation. The studies that try to validate previous NMR OCs are scarce and it could also be that the low replications seen in this paper are due to the specifics of ORCADES cohort. Future studies comparing NMR OCs derived on different populations might shed more light into this. One possible study-design for this would be to develop NMR OC on several different populations and thereafter assess the performance of each population-specific clock on other samples. As measuring the NMR metabolome has been gaining popularity in the recent years, NMR-based studies with larger populations and new samples might become feasible in the near future.

Furthermore, we saw that NMR-based OCAA correlates rather weakly with proteomics- and DNAme-based OCAAs, which is in line with Jansen *et al.* (2021), where a small correlation of 0.19 between metabolomics and proteomics aging was observed. One interpretation for such results can be that NMR-based OC tracks different BA compared to proteomics and DNAme assays. However, one must note here that the proteomics- and DNAme-based OCs were extremely accurate in predicting chronAge and therefore corresponding OCAAs are with very low variability.

The NMR-based OCAA has previously been associated with several cardiometabolic risk factors and diseases, such as BMI, HDL-C, alcohol usage, CRP, prevalent metabolic syndrome, prevalent T2D, incident cardiovascular events, vascular mortality, and all-cause mortality (van den Akker *et al.*, 2020; Jansen *et al.*, 2021). Our NMR metabolomics OC results are partly in line with those studies by showing significant associations with CRP and BMI. However, NMR metabolomics OCAA associated significantly with SBP in our results, whereas in van den Akker *et al.* (2020) such association was not seen. Further, we saw nominally significant p-values for the disease groups of "E10-E14 Diabetes mellitus" and "E50-E64 Other nutritional deficiencies", confirming that NMR OC is reflecting metabolic health. However, in the ICD-10 block describing diseases of the circulatory system, only the group of hypertensive diseases had nominally significant association with NMR OCAA, which was unexpected given the previous results regarding cardiovascular health and given that there were significant associations with cardiometabolic risk factors. Of note, the associations between the incident cardiovascular outcomes and metaboAge OCAA were assessed in a cohort of elderly adults at risk of CVD (van den Akker *et al.*, 2020), which might have affected the results. Therefore, further studies are needed to clearly establish the associations between NMR OCAA and disease outcomes.

Despite the need for further confirmation of the associations with disease outcomes by future studies, the present evidence suggests that NMR OC tracks partly the BA that is related to cardiometabolic health and could consequently reflect as well dietary behavior. Previously, the association between BA and dietary behavior has been investigated for the BA characterized by DNAme data. More specifically, the OCAA of DNAme-based OC has been found to be associated with fish intake and blood carotenoid levels (indicator of fruit and vegetable intake) (Quach *et al.*, 2017), and heavy alcohol consumption (Beach *et al.*, 2015;

Luo *et al.*, 2020). Moreover, results of a randomized trial detected that one-year Mediterranean diet was able to significantly reduce BA (Gensous *et al.*, 2020). Current author is not aware of any similar studies regarding NMR-based OCs, but there seems to be considerable potential for using NMR-based OC for assessing the usefulness of dietary interventions with regards to aging related to cardiometabolic health. However, before that, the associations with disease outcomes need further confirmation. This should be feasible in the near future, when larger NMR-samples become available. Finally, it would be very interesting to assess whether the intervention with Mediterranean diet is able to reduce NMR-based BA.

CONCLUSIONS

Understanding the mechanisms of how dietary choices affect health has the potential to provide grounds for personalized health-improving dietary advice. NMR blood metabolic profile measurements are a valuable resource for further elucidating the diet-disease relationships. In addition, recent large-scale GWASs provide great asset for pursuing MR analyses. New knowledge regarding the diet, metabolites and health interplay forms a basis for future dietary RCTs that can ultimately lead to well-informed and personalized dietary recommendations.

The main conclusions drawn from this thesis are as follows:

- MR approach can be used for detecting possibly causal effect of dietary traits on blood metabolites. This was confirmed by our results agreeing with previous RCTs and conflicting with some observational studies. In total we detected 413 potentially causal associations. For example, we found that coffee and alcohol traits pose similar elevating effects on LDL- and IDL-related measurements, whereas the VLDL-related measurements seem to be elevated by coffee consumption and not by alcohol consumption. Furthermore, beer consumption appeared to have an elevating effect on DHA, while we observed a potential elevating effect of vegetarianism on omega-6 fatty acids.
- Lack of dietary variety expressed as FN has adverse effects on blood metabolic profile. More precisely, higher level of FN is negatively associated with several omega-3 related measurements. The effect of FN on T2D and CHD warrants further investigation with larger samples, since our results were conflicting. Namely, FN elevated the risk of CHD in EstBB cohort, but not in DILGOM cohort, whereas FN elevated the risk of T2D in DILGOM cohort, but not in EstBB cohort. These results highlight that FN is harmful for health and there is need for further studies with larger samples.
- The NMR OC developed in ORCADES training sample correlated strongly (r=0.74) with chronological age in ORCADES testing sample and weakly (r=0.26) in EstBB sample. This indicates that the NMR-based omics clock may be population-specific. Further, NMR-based OC seems to describe different underlying biological age compared to proteomics- and DNAme-based OCs. The excess of NMR-metabolites based biological age over the chronological age is associated with cardiovascular risk factors. In addition, the associations with nutritional and metabolic diseases were nominally significant, but not after correcting for multiple testing. Larger samples could confirm these findings. If the associations with nutritional and metabolic diseases are confirmed, NMR metabolomics clock has the potential to be associated with and affected by dietary behavior.

SUMMARY IN ESTONIAN

Toitumine, vere metaboliidid ja tervis

Toitumisvalikute seoseid tervisega on laialdaselt erinevates teadustöödes käsitletud ning on näidatud, et need avaldavad mõju haigusriskidele, erinevatele riskifaktoritele ja ka suremusele. Kuigi teemat on mitmekülgselt uuritud, on siiski kohati selgusetu, milliste täpsete mehhanismide läbi toitumine tervist mõjutab. Parema teadmise puudumine võib aga omakorda kaasa tuua vastuolulised toitumissoovitused.

Kõige parem oleks toitumise mõju tervisele uurida randomiseeritud kliiniliste uuringutega (RCT), kuid neid on tihtipeale väga keeruline suurtel valimitel läbi viia. Lisaks on mõningate tunnuste, näiteks alkoholitarbimise, puhul sellised uuringud ebaeetilised. Viimaste kümnendite tormilised arengud DNA järjestuste määramises ja analüüsimises on toonud kaasa uued võimalused potentsiaalsete põhjuslike seoste uurimiseks, seda ka toitumist kirjeldavate tunnuste ning tervisenäitajate vahel. Nimelt on võimalik kasutada Mendeli Randomiseerimise (MR) meetodit, mis oma olemuselt on instrument-tunnuse analüüs. MR eripära on seejuures, et instrumentidena kasutatakse geeniandmetest saadavat informatsiooni. Siinkohal ongi oluline nimetatud kiire areng DNA uurimises, mille tulemusel on viimastel aastatel väga mitmed teadustööd raporteerinud ülegenoomsete assotsiatsioonanalüüside (GWAS) tulemusi. GWAS on hüpoteesivaba testimine üle terve genoomi, mille abil on võimalik leida, millised ühenukleotiidsed polümorfismid (SNP) on seotud huvipakkuvate tunnustega. Nimetatud SNPd ongi instrumentideks MR analüüsis.

Teisalt on viimasel aastakümnel hoo sisse saanud verest metaboliitide määramine suurematel valimitel ning erinevad biopangad hõlmavad endas muuhulgas ka vere metaboliite kirjeldavaid andmestikke. Korraga määratakse suur hulk erinevaid metaboliite – mõningad näited nendest on HDL-kolesterool, LDL-kolesterool, omega-3 ja omega-6 rasvhapped. Sääraste metaboliitide määramiseks verest on mitmeid meetodeid. Antud töös käsitletavad metaboliidid on määratud tuumamagnetresonants (NMR) tehnoloogial, mis on viimastel aastatel üha levinumaks meetodiks saanud biopankade andmestike loomisel. Kuna vere metaboliidid on mõjutatud toitumisest ning on ise omakorda erinevate haiguste ennustajateks, võib nende uurimise kaudu olla võimalik osaliselt selgitada, kuidas toitumine tervisele mõju avaldab.

Uurides toitumise mõju vere metaboliitidele, saame antud tulemusi tõlgendada varasemate tervisenäitajaid ja haiguseid kirjeldavate teadusuuringute taustal. Samas laiema pildi huvides on oluline teada, kas NMR metaboliitide profiil suudab kirjeldada tervise hetkeseisukorda, ning veel täpsemalt mõne kindla elundkonna tervist. Selle osas on viimastel aastatel palju põnevust tekitanud bioloogilise vanuse uurimine kasutades selleks oomika kellasid. Need on inimese bioloogilise vanuse arvutamiseks loodud algoritmid, mis tuginevad masinõppe meetodite rakendamisel erinevatel -oomika andmetel (näiteks metaboloomika, proteoomika, DNA metülatsioon). Sealjuures on inimese bioloogilise vanuse ja

kronoloogilise vanuse vaheline erinevus (tähistatud antud töös OCAA) potentsiaalselt informatiivne tema tervise hetkeseisu kohta.

Käesolev doktoritöö uurib toitumise mõju tervisele kolmest vaatenurgast. Esiteks uurib MR analüüsi abil, milline on toitumise potentsiaalselt põhjuslik mõju vere metaboliitidele. Seejärel analüüsib, millist mõju omab vere metaboliitidele dieedi süsteemne väike variatiivsus kirjeldatult toidu neofoobia abil (käitumine, kus inimene keeldub söömast või maitsmast uusi või võõraid toite). Ning viimaks uurib, kas NMR metaboliitide profiil kirjeldab inimese bioloogilist vanust ning kas erinevus säärase bioloogilise vanuse ja kronoloogilise vanuse vahel ennustab tervisenäitajaid. Nimetatud uuringud on käsitletud doktoritöö eksperimentaalosas. Kokkuvõtvalt on antud doktoritöö üldisemaks eesmärgiks anda lisateadmisi toitumise, vere metaboliitide ja tervise omavaheliste seoste kohta, et aidata astuda samm lähemale tulevikule, kus on teadlikumad ja personaliseeritud tervist edendavad toitumissoovitused.

Töö esimene osa annab teaduskirjandusele toetudes ülevaate, milliste uuringutega on siiani vaadeldud toitumise, vere metaboliitide ja tervise omavahelisi seoseid. Lisaks kirjeldab toidu neofoobia olemust ning selle seoseid muude tunnustega. Seejärel selgitab, kuidas töötab MR analüüs, millised on selle levinumad meetodid, ning millised on MR rakendamise eripärad toitumistunnuste mõju uurimise korral. Samuti toob välja, mida mõeldakse oomika kellade all, millistele andmetele tuginedes neid on konstrueeritud ning milliste tervisenäitajatega on seostatud NMR metaboliitide põhjal arvutatud bioloogilist vanust. Kirjanduse ülevaate viimane osa sisaldab mõtisklusi võimalike tulevikuväljavaadete osas olukorras, kus viimaste aastatega on toitumise, metaboliitide ja tervise omavaheliste seoste uurimiseks avanenud rohkelt uusi võimalusi.

Doktoritöö eksperimentaalosa esimene artikkel otsis erinevate toitumist kirjeldavate tunnuste põhjuslikku mõju vere metaboliitidele. Kuna antud töö on teadaolevalt esimene, mis kasutas MR-analüüsi, et laialdaselt uurida erinevate toitumistunnuste ja vere metaboliitide vahelisi seoseid, on oluline välja tuua, et artikli tulemused ühtisid varasemate teadaolevate põhjuslike seostega ning varasemate RCTdega. Autoritele teadaolevalt ei olnud artikli tulemused vastuolus mitte ühegi RCTga. Küll aga leidus vastuolusid mõningate vaatlusuuringutega, mis annab põhjust arvata, et vaatlusuuringute tulemused võisid olla segajate poolt mõjutatud. Seega kinnitas antud artikkel, et MR-analüüs on sobiv ja kasulik meetod toitumise mõju metaboliitidele uurimiseks. Antud töö tulemusel leiti 413 potentsiaalselt põhjuslikku seost. Muuhulgas selgus, et kohv ja alkohoolsed joogid avaldavad võrdlemisi sarnast mõju mitmetele vaadeldud metaboliitidele. Eriti silmatorkav selline muster madala ja keskmise tihedusega lipoproteiinidega seotud mõõtmiste osas, kus suurem kohvi ja alkohoolsete jookide tarbimine seostub kõrgemate vaadeldud metaboliitide tasemetega. Täiesti erinev oli nende jookide mõju aga väga madala tihedusega lipoproteiinidega seotud mõõtmistele: nimelt seostus suurem kohvi tarbimine oluliselt kõrgemate metaboliitide tasemetega samas kui alkoholil märgatavat mõju ei olnud. Huvipakkuv on ka õlle positiivne mõju DHA'le, mis on üks omega-3 rasvhapetest ning mille kõrgem tase on varasema kirjanduse põhjal tervisele kasulik. Samas oli alkohoolsetest jookidest see

mõju vaid õllel, näidates, et mõju polnud tõenäoliselt põhjustatud alkoholist endast, vaid mõnest muust komponendist. Lisaks leidsime, et taimetoitlus tõstab omega-6 rasvhapete tasemeid veres. Kuna see tulemus oli leitud vaid ühe instrument-tunnuse baasil, vajab see enne kindlamaid järeldusi järgnevate uuringute poolt kinnitamist. Kokkuvõtvalt, antud artikkel tõestas, et MR on sobiv ja kasulik vahend, millega uurida toitumistunnuste mõju vere metaboliitidele ning leidis mitmeid seoseid, mis millisel moel toitumine tervist mõjutab.

Järgnevalt näitab eksperimentaalosa teine artikkel, et toidu neofoobial on vere metaboolsele profiilile ebasoodne mõju. Täpsemalt, kõrgem tulemus FN skaalal assotsieerub madalamate omega-3'ga seotud mõõtmistega. Samas FN mõju südameveresoonkonna haigustele (CHD) ja tüüp II diabeedile (T2D) jäi osalt selgusetuks. Seda seetõttu, et FN tõstis CHD riski eestlaste kohordis, aga mitte soomlaste kohordis, samas kui FN tõstis T2D riski soomlaste kohordis, aga mitte eestlaste kohordis. Seega on küll põhjust arvata, et FN on tervisele kahjulik, aga täpsema selguse saamiseks CHD ja T2D osas on vaja tulevasi uurimusi suurema valimiga.

Eksperimentaalosa kolmas peatükk uuris oomika kellasid, mis väljendavad inimese bioloogilist vanust. Täpsemalt keskenduti NMR metaboliitide profiili põhjal konstrueeritud oomika kellale. Tulemuste põhjal on põhjust arvata, et NMR andmetel baseeruv oomika kell on osalt populatsiooni-spetsiifiline, erinevalt näiteks nendest, mis tuginevad proteoomika ja DNA metülatsiooni andmetel. Lisaks näitasid tulemused, et NMR andmetel baseeruv bioloogiline vanus on seotud kardiovaskulaarsete riskifaktoritega ning toitumise ja ainevahetusega seotud haigustega. Koos varasematest uuringutest saadud teadmistega, kus on näidatud, et NMR andmetel baseeruv bioloogiline vanus on seotud erinevate kardiometaboolsete haigustega ja riskifaktoritega, on põhjust arvata, et NMR oomika kell võib eelkõige olla kasulik kardiometaboolse tervise kirjeldamiseks. Nähtud seosed erinevate haigustega ei olnud küll pärast mitmesele testimisele korrigeerimist statistiliselt olulised ning vajavad enne kindlamaid järeldusi uute ja suuremate uuringute poolt kinnitamist. Kui need tulemused leiavad kinnituse, on võimalik, et NMR andmetel baseeruv bioloogiline vanus on potentsiaalselt mõjutatav toitumiskäitumisest.

Käesolev doktoritöö vaatles toitumise, vere metaboliitide ja tervise omavahelisi seoseid kolmest erinevast küljest ning leidis mitmeid assotsiatsioone, mis aitavad kirjeldada toitumise ja tervise vaheliste seoste mehhanisme. Leitud seosed on aluseks tulevastele randomiseeritud uuringutele, mille järel on võimalik saada tugevam alus tervist edendavatele teadlikutele ja personaliseeritud toitumis-soovitustele.

REFERENCES

Agarwal, A. and Ioannidis, J.P.A. (2019) 'PREDIMED trial of Mediterranean diet: retracted, republished, still trusted?', *BMJ*, p. l341. doi:10.1136/bmj.l341.

Ahola-Olli, A.V., Mustelin, L., Kalimeri, M., Kettunen, J., Jokelainen, J., Auvinen, J., Puukka, K., Havulinna, A.S., Lehtimäki, T., Kähönen, M., Juonala, M., Keinänen-Kiukaanniemi, S., Salomaa, V., Perola, M., Järvelin, M.-R., Ala-Korpela, M., Raitakari, O. and Würtz, P. (2019) 'Circulating metabolites and the risk of type 2 diabetes: a prospective study of 11,896 young adults from four Finnish cohorts', *Diabetologia*, 62(12), pp. 2298–2309. doi:10.1007/s00125-019-05001-w.

van den Akker, E.B., Trompet, S., Barkey Wolf, J.J.H., Beekman, M., Suchiman, H.E.D., Deelen, J., Asselbergs, F.W., Boersma, E., Cats, D., Elders, P.M., Geleijnse, J.M., Ikram, M.A., Kloppenburg, M., Mei, H., Meulenbelt, I., Mooijaart, S.P., Nelissen, R.G.H.H., Netea, M.G., Penninx, B.W.J.H., Slofstra, M., Stehouwer, C.D.A., Swertz, M.A., Teunissen, C.E., Terwindt, G.M., 't Hart, L.M., van den Maagdenberg, A.M.J.M., van der Harst, P., van der Horst, I.C.C., van der Kallen, C.J.H., van Greevenbroek, M.M.J., van Spil, W.E., Wijmenga, C., Zhernakova, A., Zwinderman, A.H., Sattar, N., Jukema, J.W., van Duijn, C.M., Boomsma, D.I., Reinders, M.J.T. and Slagboom, P.E. (2020) 'Metabolic Age Based on the BBMRI-NL ^{1} H-NMR Metabolomics Repository as Biomarker of Age-related Disease', *Circulation: Genomic and Precision Medicine*, 13(5), pp. 541–547. doi:10.1161/CIRCGEN.119.002610.

Andrews, S.J., Goate, A. and Anstey, K.J. (2020) 'Association between alcohol consumption and Alzheimer's disease: A Mendelian randomization study', *Alzheimer's & Dementia*, 16(2), pp. 345–353. doi:10.1016/j.jalz.2019.09.086.

Appel, L.J., Moore, T.J., Obarzanek, E., Vollmer, W.M., Svetkey, L.P., Sacks, F.M., Bray, G.A., Vogt, T.M., Cutler, J.A., Windhauser, M.M., Lin, P.-H., Karanja, N., Simons-Morton, D., McCullough, M., Swain, J., Steele, P., Evans, M.A., Miller, E.R. and Harsha, D.W. (1997) 'A Clinical Trial of the Effects of Dietary Patterns on Blood Pressure', *New England Journal of Medicine*, 336(16), pp. 1117–1124. doi:10.1056/NEJM199704173361601.

Barter, P.J., Caulfield, M., Eriksson, M., Grundy, S.M., Kastelein, J.J.P., Komajda, M., Lopez-Sendon, J., Mosca, L., Tardif, J.-C., Waters, D.D., Shear, C.L., Revkin, J.H., Buhr, K.A., Fisher, M.R., Tall, A.R. and Brewer, B. (2007) 'Effects of Torcetrapib in Patients at High Risk for Coronary Events', *New England Journal of Medicine*, 357(21), pp. 2109–2122. doi:10.1056/NEJMoa0706628.

Beach, S.R.H., Dogan, M.V., Lei, M.-K., Cutrona, C.E., Gerrard, M., Gibbons, F.X., Simons, R.L., Brody, G.H. and Philibert, R.A. (2015) 'Methylomic Aging as a Window onto the Influence of Lifestyle: Tobacco and Alcohol Use Alter the Rate of Biological Aging', *Journal of the American Geriatrics Society*, 63(12), pp. 2519–2525. doi:10.1111/jgs.13830.

Belsky, D.W., Moffitt, T.E., Cohen, A.A., Corcoran, D.L., Levine, M.E., Prinz, J.A., Schaefer, J., Sugden, K., Williams, B., Poulton, R. and Caspi, A. (2017) 'Eleven Telomere, Epigenetic Clock, and Biomarker-Composite Quantifications of Biological Aging: Do They Measure the Same Thing?', *American Journal of Epidemiology* [Preprint]. doi:10.1093/aje/kwx346.

Bergholdt, H.K., Nordestgaard, B.G. and Ellervik, C. (2015) 'Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals', *The American Journal of Clinical Nutrition*, 102(2), pp. 487–496. doi:10.3945/ajcn.114.105049.

Bernasconi, A.A., Wiest, M.M., Lavie, C.J., Milani, R.V. and Laukkanen, J.A. (2021) 'Effect of Omega-3 Dosage on Cardiovascular Outcomes', *Mayo Clinic Proceedings*, 96(2), pp. 304–313. doi:10.1016/j.mayocp.2020.08.034.

Boeing, H., Bechthold, A., Bub, A., Ellinger, S., Haller, D., Kroke, A., Leschik-Bonnet, E., Müller, M.J., Oberritter, H., Schulze, M., Stehle, P. and Watzl, B. (2012) 'Critical review: vegetables and fruit in the prevention of chronic diseases', *European Journal of Nutrition*, 51(6), pp. 637–663. doi:10.1007/s00394-012-0380-y.

Bot, M., Milaneschi, Y., Al-Shehri, T., Amin, Najaf, Garmaeva, S., Onderwater, Gerrit L.J., Pool, Rene, Thesing, C.S., Vijfhuizen, L.S., Vogelzangs, N., Arts, I.C.W., Demirkan, Ayse, van Duijn, C., van Greevenbroek, M., van der Kallen, Carla J.H., Köhler, S., Ligthart, L., van den Maagdenberg, Arn M.J.M., Mook-Kanamori, D.O., de Mutsert, R., Tiemeier, H., Schram, M.T., Stehouwer, Coen D.A., Terwindt, Gisela M., Willems van Dijk, K., Fu, Jingyuan, Zhernakova, A., Beekman, Marian, Slagboom, P.E., Boomsma, D.I., Penninx, Brenda W.J.H., Beekman, M., Suchiman, H.E.D., Deelen, J., Amin, N., Beulens, J.W., van der Bom, J.A., Bomer, N., Demirkan, A., van Hilten, J.A., Meessen, J.M.T.A., Pool, R., Moed, M.H., Fu, J., Onderwater, G.L.J., Rutters, F., So-Osman, C., van der Flier, W.M., van der Heijden, A.A.W.A., van der Spek, A., Asselbergs, F.W., Boersma, E., Elders, P.M., Geleijnse, J.M., Ikram, M.A., Kloppenburg, M., Meulenbelt, I., Mooijaart, S.P., Nelissen, R.G.H.H., Netea, M.G., Penninx, B.W.J.H., Stehouwer, C.D.A., Teunissen, C.E., Terwindt, G.M., 't Hart, L.M., van den Maagdenberg, A.M.J.M., van der Harst, P., van der Horst, I.C.C., van der Kallen, C.J.H., van Greevenbroek, M.M.J., van Spil, W.E., Wijmenga, C., Zwinderman, A.H., Zhernikova, A., Jukema, J.W. and Sattar, N. (2020) 'Metabolomics Profile in Depression: A Pooled Analysis of 230 Metabolic Markers in 5283 Cases With Depression and 10,145 Controls', *Biological Psychiatry*, 87(5), pp. 409–418. doi:10.1016/j.biopsych.2019.08.016.

Bowden, J., Davey Smith, G. and Burgess, S. (2015) 'Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression', *International Journal of Epidemiology*, 44(2), pp. 512–525. doi:10.1093/ije/dyv080.

Bowden, J., Davey Smith, G., Haycock, P.C. and Burgess, S. (2016) 'Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator', *Genetic Epidemiology*, 40(4), pp. 304–314. doi:10.1002/gepi.21965.

Bowden, J., Spiller, W., Del Greco M, F., Sheehan, N., Thompson, J., Minelli, C. and Davey Smith, G. (2018) 'Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression', *International Journal of Epidemiology*, 47(4), pp. 1264–1278. doi:10.1093/ije/dyy101.

Bradbury, K.E., Young, H.J., Guo, W. and Key, T.J. (2018) 'Dietary assessment in UK Biobank: an evaluation of the performance of the touchscreen dietary questionnaire', *Journal of Nutritional Science*, 7, p. e6. doi:10.1017/jns.2017.66.

Briel, M., Ferreira-Gonzalez, I., You, J.J., Karanicolas, P.J., Akl, E.A., Wu, P., Blechacz, B., Bassler, D., Wei, X., Sharman, A., Whitt, I., Alves da Silva, S., Khalid, Z., Nordmann, A.J., Zhou, Q., Walter, S.D., Vale, N., Bhatnagar, N., O'Regan, C., Mills, E.J., Bucher, H.C., Montori, V.M. and Guyatt, G.H. (2009) 'Association between

change in high density lipoprotein cholesterol and cardiovascular disease morbidity and mortality: systematic review and meta-regression analysis', *BMJ*, 338(feb16 1), pp. b92–b92. doi:10.1136/bmj.b92.

Brien, S.E., Ronksley, P.E., Turner, B.J., Mukamal, K.J. and Ghali, W.A. (2011) 'Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies', *BMJ*, 342(feb22 1), pp. d636–d636. doi:10.1136/bmj.d636.

Brown, T.J., Brainard, J., Song, F., Wang, X., Abdelhamid, A. and Hooper, L. (2019) 'Omega-3, omega-6, and total dietary polyunsaturated fat for prevention and treatment of type 2 diabetes mellitus: systematic review and meta-analysis of randomised controlled trials', *BMJ*, p. l4697. doi:10.1136/bmj.l4697.

Burgess, S., Butterworth, A. and Thompson, S.G. (2013) 'Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data', *Genetic Epidemiology*, 37(7), pp. 658–665. doi:10.1002/gepi.21758.

Burgess, S. and Harshfield, E. (2016) 'Mendelian randomization to assess causal effects of blood lipids on coronary heart disease: lessons from the past and applications to the future', *Current Opinion in Endocrinology, Diabetes & Obesity*, 23(2), pp. 124–130. doi:10.1097/MED.0000000000000230.

Burgess, S. and Thompson, S.G. (2015) *Mendelian randomization: methods for using genetic variants in causal estimation*. CRC Press.

Calder, P.C. (2010) 'Omega-3 Fatty Acids and Inflammatory Processes', *Nutrients*, 2(3), pp. 355–374. doi:10.3390/nu2030355.

Carmena, R., Duriez, P. and Fruchart, J.-C. (2004) 'Atherogenic Lipoprotein Particles in Atherosclerosis', *Circulation*, 109(23_suppl_1). doi:10.1161/01.CIR.0000131511.50734.44.

Casula, M., Colpani, O., Xie, S., Catapano, A.L. and Baragetti, A. (2021) 'HDL in Atherosclerotic Cardiovascular Disease: In Search of a Role', *Cells*, 10(8), p. 1869. doi:10.3390/cells10081869.

Chandler, P.D., Balasubramanian, R., Paynter, N., Giulianini, F., Fung, T., Tinker, L.F., Snetselaar, L., Liu, S., Eaton, C., Tobias, D.K., Tabung, F.K., Manson, J.E., Giovannucci, E.L., Clish, C. and Rexrode, K.M. (2020) 'Metabolic signatures associated with Western and Prudent dietary patterns in women', *The American Journal of Clinical Nutrition*, 112(2), pp. 268–283. doi:10.1093/ajcn/nqaa131.

Chen, L., Davey Smith, G., Harbord, R.M. and Lewis, S.J. (2008) 'Alcohol Intake and Blood Pressure: A Systematic Review Implementing a Mendelian Randomization Approach', *PLoS Medicine*. Edited by C. Minelli, 5(3), p. e52. doi:10.1371/journal.pmed.0050052.

Chrysant, S.G. (2015) 'Coffee Consumption and Cardiovascular Health', *The American Journal of Cardiology*, 116(5), pp. 818–821. doi:10.1016/j.amjcard.2015.05.057.

Cole, J.B., Florez, J.C. and Hirschhorn, J.N. (2020) 'Comprehensive genomic analysis of dietary habits in UK Biobank identifies hundreds of genetic associations', *Nature Communications*, 11(1), p. 1467. doi:10.1038/s41467-020-15193-0.

Cole, J.H., Marioni, R.E., Harris, S.E. and Deary, I.J. (2019) 'Brain age and other bodily "ages": implications for neuropsychiatry', *Molecular Psychiatry*, 24(2), pp. 266–281. doi:10.1038/s41380-018-0098-1.

Cooke, L.J., Haworth, C.M. and Wardle, J. (2007) 'Genetic and environmental influences on children's food neophobia', *The American Journal of Clinical Nutrition*, 86(2), pp. 428–433. doi:10.1093/ajcn/86.2.428.

Davey Smith, G. and Ebrahim, S. (2003) '"Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease?*', *International Journal of Epidemiology*, 32(1), pp. 1–22. doi:10.1093/ije/dyg070.

Deelen, J., Kettunen, J., Fischer, K., van der Spek, A., Trompet, S., Kastenmüller, G., Boyd, A., Zierer, J., van den Akker, E.B., Ala-Korpela, M., Amin, N., Demirkan, A., Ghanbari, M., van Heemst, D., Ikram, M.A., van Klinken, J.B., Mooijaart, S.P., Peters, A., Salomaa, V., Sattar, N., Spector, T.D., Tiemeier, H., Verhoeven, A., Waldenberger, M., Würtz, P., Davey Smith, G., Metspalu, A., Perola, M., Menni, C., Geleijnse, J.M., Drenos, F., Beekman, M., Jukema, J.W., van Duijn, C.M. and Slagboom, P.E. (2019) 'A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals', *Nature Communications*, 10(1), p. 3346. doi:10.1038/s41467-019-11311-9.

Dilis, V., Katsoulis, M., Lagiou, P., Trichopoulos, D., Naska, A. and Trichopoulou, A. (2012) 'Mediterranean diet and CHD: the Greek European Prospective Investigation into Cancer and Nutrition cohort', *British Journal of Nutrition*, 108(4), pp. 699–709. doi:10.1017/S0007114512001821.

Dovey, T.M., Staples, P.A., Gibson, E.L. and Halford, J.C.G. (2008) 'Food neophobia and "picky/fussy" eating in children: A review', *Appetite*, 50(2–3), pp. 181–193. doi:10.1016/j.appet.2007.09.009.

Drouin-Chartier, J.-P., Hernández-Alonso, P., Guasch-Ferré, M., Ruiz-Canela, M., Li, J., Wittenbecher, C., Razquin, C., Toledo, E., Dennis, C., Corella, D., Estruch, R., Fitó, M., Eliassen, A.H., Tobias, D.K., Ascherio, A., Mucci, L.A., Rexrode, K.M., Karlson, E.W., Costenbader, K.H., Fuchs, C.S., Liang, L., Clish, C.B., Martínez-González, M.A., Salas-Salvadó, J. and Hu, F.B. (2021) 'Dairy consumption, plasma metabolites, and risk of type 2 diabetes', *The American Journal of Clinical Nutrition*, 114(1), pp. 163–174. doi:10.1093/ajcn/nqab047.

Du, D., Bruno, R., Blizzard, L., Venn, A., Dwyer, T., Smith, K.J., Magnussen, C.G. and Gall, S. (2020) 'The metabolomic signatures of alcohol consumption in young adults', *European Journal of Preventive Cardiology*, 27(8), pp. 840–849. doi:10.1177/2047487319834767.

Enroth, S., Enroth, S.B., Johansson, Å. and Gyllensten, U. (2015) 'Protein profiling reveals consequences of lifestyle choices on predicted biological aging', *Scientific Reports*, 5(1), p. 17282. doi:10.1038/srep17282.

Esko, T., Hirschhorn, J.N., Feldman, H.A., Hsu, Y.-H.H., Deik, A.A., Clish, C.B., Ebbeling, C.B. and Ludwig, D.S. (2017) 'Metabolomic profiles as reliable biomarkers of dietary composition', *The American Journal of Clinical Nutrition*, 105(3), pp. 547–554. doi:10.3945/ajcn.116.144428.

Esposito, K., Marfella, R., Ciotola, M., Di Palo, C., Giugliano, F., Giugliano, G., D'Armiento, M., D'Andrea, F. and Giugliano, D. (2004) 'Effect of a Mediterranean-Style Diet on Endothelial Dysfunction and Markers of Vascular Inflammation in the Metabolic Syndrome: A Randomized Trial', *JAMA*, 292(12), p. 1440. doi:10.1001/jama.292.12.1440.

Estruch, R., Ros, E., Salas-Salvadó, J., Covas, M.-I., Corella, D., Arós, F., Gómez-Gracia, E., Ruiz-Gutiérrez, V., Fiol, M., Lapetra, J., Lamuela-Raventos, R.M., Serra-Majem, L., Pintó, X., Basora, J., Muñoz, M.A., Sorlí, J.V., Martínez, J.A., Fitó, M., Gea, A., Hernán, M.A. and Martínez-González, M.A. (2018) 'Primary Prevention of Cardiovascular Disease with a Mediterranean Diet Supplemented with Extra-Virgin Olive Oil or Nuts', *New England Journal of Medicine*, 378(25), p. e34. doi:10.1056/NEJMoa1800389.

Faith, M.S., Heo, M., Keller, K.L. and Pietrobelli, A. (2013) 'Child food neophobia is heritable, associated with less compliant eating, and moderates familial resemblance for BMI: Child Food Neophobia is Heritable', *Obesity*, 21(8), pp. 1650–1655. doi:10.1002/oby.20369.

Ference, B.A., Ginsberg, H.N., Graham, I., Ray, K.K., Packard, C.J., Bruckert, E., Hegele, R.A., Krauss, R.M., Raal, F.J., Schunkert, H., Watts, G.F., Borén, J., Fazio, S., Horton, J.D., Masana, L., Nicholls, S.J., Nordestgaard, B.G., van de Sluis, B., Taskinen, M.-R., Tokgözoğlu, L., Landmesser, U., Laufs, U., Wiklund, O., Stock, J.K., Chapman, M.J. and Catapano, A.L. (2017) 'Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel', *European Heart Journal*, 38(32), pp. 2459–2472. doi:10.1093/eurheartj/ehx144.

Ference, B.A., Kastelein, J.J.P., Ray, K.K., Ginsberg, H.N., Chapman, M.J., Packard, C.J., Laufs, U., Oliver-Williams, C., Wood, A.M., Butterworth, A.S., Di Angelantonio, E., Danesh, J., Nicholls, S.J., Bhatt, D.L., Sabatine, M.S. and Catapano, A.L. (2019) 'Association of Triglyceride-Lowering *LPL* Variants and LDL-C–Lowering *LDLR* Variants With Risk of Coronary Heart Disease', *JAMA*, 321(4), p. 364. doi:10.1001/jama.2018.20045.

Fernández-Lázaro, C.I., Ruiz-Canela, M. and Martínez-González, M.Á. (2021) 'Deep dive to the secrets of the PREDIMED trial', *Current Opinion in Lipidology*, 32(1), pp. 62–69. doi:10.1097/MOL.0000000000000731.

Fischer, K., Kettunen, J., Würtz, P., Haller, T., Havulinna, A.S., Kangas, A.J., Soininen, P., Esko, T., Tammesoo, M.-L., Mägi, R., Smit, S., Palotie, A., Ripatti, S., Salomaa, V., Ala-Korpela, M., Perola, M. and Metspalu, A. (2014) 'Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of All-Cause Mortality: An Observational Study of 17,345 Persons', *PLoS Medicine*. Edited by C. Minelli, 11(2), p. e1001606. doi:10.1371/journal.pmed.1001606.

Flight, I., Leppard, P. and Cox, D.N. (2003) 'Food neophobia and associations with cultural diversity and socio-economic status amongst rural and urban Australian adolescents', *Appetite*, 41(1), pp. 51–59. doi:10.1016/S0195-6663(03)00039-4.

Gensous, N., Garagnani, P., Santoro, A., Giuliani, C., Ostan, R., Fabbri, C., Milazzo, M., Gentilini, D., di Blasio, A.M., Pietruszka, B., Madej, D., Bialecka-Debek, A., Brzozowska, A., Franceschi, C. and Bacalini, M.G. (2020) 'One-year Mediterranean diet promotes epigenetic rejuvenation with country- and sex-specific effects: a pilot study from the NU-AGE project', *GeroScience*, 42(2), pp. 687–701. doi:10.1007/s11357-019-00149-0.

Guasch-Ferré, M., Salas-Salvadó, J., Ros, E., Estruch, R., Corella, D., Fitó, M., Martínez-González, M.A., Arós, F., Gómez-Gracia, E., Fiol, M., Lapetra, J., Lamuela-Raventos, R.M., Tur, J., Martinez, J.A., Serra-Majem, L. and Pintó, X. (2017) 'The PREDIMED trial, Mediterranean diet and health outcomes: How strong is the evidence?', *Nutrition, Metabolism and Cardiovascular Diseases*, 27(7), pp. 624–632. doi:10.1016/j.numecd.2017.05.004.

Guertin, K.A., Loftfield, E., Boca, S.M., Sampson, J.N., Moore, S.C., Xiao, Q., Huang, W.-Y., Xiong, X., Freedman, N.D., Cross, A.J. and Sinha, R. (2015) 'Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer', *American Journal of Clinical Nutrition*, 101(5), pp. 1000–1011. doi:10.3945/ajcn.114.096099.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.-B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T. and Zhang, K. (2013) 'Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates', *Molecular Cell*, 49(2), pp. 359–367. doi:10.1016/j.molcel.2012.10.016.

Hansson, P., Holven, K.B., Øyri, L.K.L., Brekke, H.K., Gjevestad, G.O., Thoresen, M. and Ulven, S.M. (2019) 'Sex differences in postprandial responses to different dairy products on lipoprotein subclasses: a randomised controlled cross-over trial', *British Journal of Nutrition*, 122(07), pp. 780–789. doi:10.1017/S0007114519001429.

Harrer, M., Cuijpers, P., Furukawa, T.A. and Ebert, D.D. (2021) *Doing Meta-Analysis with R: A Hands-On Guide*. Chapman and Hall/CRC.

Harris, W.S., Tintle, N.L., Imamura, F., Qian, F., Korat, A.V.A., Marklund, M., Djoussé, L., Bassett, J.K., Carmichael, P.-H., Chen, Y.-Y., Hirakawa, Y., Küpers, L.K., Laguzzi, F., Lankinen, M., Murphy, R.A., Samieri, C., Senn, M.K., Shi, P., Virtanen, J.K., Brouwer, I.A., Chien, K.-L., Eiriksdottir, G., Forouhi, N.G., Geleijnse, J.M., Giles, G.G., Gudnason, V., Helmer, C., Hodge, A., Jackson, R., Khaw, K.-T., Laakso, M., Lai, H., Laurin, D., Leander, K., Lindsay, J., Micha, R., Mursu, J., Ninomiya, T., Post, W., Psaty, B.M., Risérus, U., Robinson, J.G., Shadyab, A.H., Snetselaar, L., Sala-Vila, A., Sun, Y., Steffen, L.M., Tsai, M.Y., Wareham, N.J., Wood, A.C., Wu, J.H.Y., Hu, F., Sun, Q., Siscovick, D.S., Lemaitre, R.N., Mozaffarian, D. and The Fatty Acids and Outcomes Research Consortium (FORCE) (2021) 'Blood n-3 fatty acid levels and total and cause-specific mortality from 17 prospective studies', *Nature Communications*, 12(1), p. 2329. doi:10.1038/s41467-021-22370-2.

Hazley, D., McCarthy, S.N., Stack, M., Walton, J., McNulty, B.A., Flynn, A. and Kearney, J.M. (2022) 'Food neophobia and its relationship with dietary variety and quality in Irish adults: Findings from a national cross-sectional study', *Appetite*, 169, p. 105859. doi:10.1016/j.appet.2021.105859.

Helland, S.H., Bere, E., Bjørnarå, H.B. and Øverby, N.C. (2017) 'Food neophobia and its association with intake of fish and other selected foods in a Norwegian sample of toddlers: A cross-sectional study', *Appetite*, 114, pp. 110–117. doi:10.1016/j.appet.2017.03.025.

Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V.Y., Yarmolinsky, J., Shihab, H.A., Timpson, N.J., Evans, D.M., Relton, C., Martin, R.M., Davey Smith, G., Gaunt, T.R. and Haycock, P.C. (2018) 'The MR-Base platform supports systematic causal inference across the human phenome', *eLife*, 7, p. e34408. doi:10.7554/eLife.34408.

Hernández-Alonso, P., Papandreou, C., Bulló, M., Ruiz-Canela, M., Dennis, C., Deik, A., Wang, D.D., Guasch-Ferré, M., Yu, E., Toledo, E., Razquin, C., Corella, D., Estruch, R., Ros, E., Fitó, M., Arós, F., Fiol, M., Serra-Majem, L., Liang, L., Clish, C.B., Martínez-González, M.A., Hu, F.B. and Salas-Salvadó, J. (2019) 'Plasma Metabolites Associated with Frequent Red Wine Consumption: A Metabolomics Approach within the PREDIMED Study', *Molecular Nutrition & Food Research*, 63(17), p. 1900140. doi:10.1002/mnfr.201900140.

Holmes, M.V., Dale, C.E., Zuccolo, L., Silverwood, R.J., Guo, Y., Ye, Z., Prieto-Merino, D., Dehghan, A., Trompet, S., Wong, A., Cavadino, A., Drogan, D., Padmanabhan, S., Li, S., Yesupriya, A., Leusink, M., Sundstrom, J., Hubacek, J.A., Pikhart, H., Swerdlow, D.I., Panayiotou, A.G., Borinskaya, S.A., Finan, C., Shah, S., Kuchenbaecker, K.B., Shah, T., Engmann, J., Folkersen, L., Eriksson, P., Ricceri, F., Melander, O., Sacerdote, C., Gamble, D.M., Rayaprolu, S., Ross, O.A., McLachlan, S.,

Vikhireva, O., Sluijs, I., Scott, R.A., Adamkova, V., Flicker, L., Bockxmeer, F.M. v., Power, C., Marques-Vidal, P., Meade, T., Marmot, M.G., Ferro, J.M., Paulos-Pinheiro, S., Humphries, S.E., Talmud, P.J., Leach, I.M., Verweij, N., Linneberg, A., Skaaby, T., Doevendans, P.A., Cramer, M.J., Harst, P. v. d., Klungel, O.H., Dowling, N.F., Dominiczak, A.F., Kumari, M., Nicolaides, A.N., Weikert, C., Boeing, H., Ebrahim, S., Gaunt, T.R., Price, J.F., Lannfelt, L., Peasey, A., Kubinova, R., Pajak, A., Malyutina, S., Voevoda, M.I., Tamosiunas, A., Maitland-van der Zee, A.H., Norman, P.E., Hankey, G.J., Bergmann, M.M., Hofman, A., Franco, O.H., Cooper, J., Palmen, J., Spiering, W., Jong, P.A. d., Kuh, D., Hardy, R., Uitterlinden, A.G., Ikram, M.A., Ford, I., Hypponen, E., Almeida, O.P., Wareham, N.J., Khaw, K.-T., Hamsten, A., Husemoen, L.L.N., Tjonneland, A., Tolstrup, J.S., Rimm, E., Beulens, J.W.J., Verschuren, W.M.M., Onland-Moret, N.C., Hofker, M.H., Wannamethee, S.G., Whincup, P.H., Morris, R., Vicente, A.M., Watkins, H., Farrall, M., Jukema, J.W., Meschia, J., Cupples, L.A., Sharp, S.J., Fornage, M., Kooperberg, C., LaCroix, A.Z., Dai, J.Y., Lanktree, M.B., Siscovick, D.S., Jorgenson, E., Spring, B., Coresh, J., Li, Y.R., Buxbaum, S.G., Schreiner, P.J., Ellison, R.C., Tsai, M.Y., Patel, S.R., Redline, S., Johnson, A.D., Hoogeveen, R.C., Hakonarson, H., Rotter, J.I., Boerwinkle, E., Bakker, P.I.W. d., Kivimaki, M., Asselbergs, F.W., Sattar, N., Lawlor, D.A., Whittaker, J., Davey Smith, G., Mukamal, K., Psaty, B.M., Wilson, J.G., Lange, L.A., Hamidovic, A., Hingorani, A.D., Nordestgaard, B.G., Bobak, M., Leon, D.A., Langenberg, C., Palmer, T.M., Reiner, A.P., Keating, B.J., Dudbridge, F., Casas, J.P., and on behalf of The InterAct Consortium (2014) 'Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data', *BMJ*, 349(jul10 6), pp. g4164–g4164. doi:10.1136/bmj.g4164.

Horrocks, L.A. and Yeo, Y.K. (1999) 'Health Benefits of Docosahexaeonic Acid (DHA)', *Pharmacological Research*, 40(3), pp. 211–225. doi:10.1006/phrs.1999.0495.

Horvath, S. (2013) 'DNA methylation age of human tissues and cell types', *Genome Biology*, 14(10), p. R115. doi:10.1186/gb-2013-14-10-r115.

Horvath, S., Singh, K., Raj, K., Khairnar, S., Sanghavi, A., Shrivastava, A., Zoller, J.A., Li, C.Z., Herenu, C.B., Canatelli-Mallat, M., Lehmann, M., Solberg Woods, L.C., Martinez, A.G., Wang, T., Chiavellini, P., Levine, A.J., Chen, H., Goya, R.G. and Katcher, H.L. (2020) *Reversing age: dual species measurement of epigenetic age with a single clock*. preprint. Developmental Biology. doi:10.1101/2020.05.07.082917.

Jannasch, F., Kröger, J. and Schulze, M.B. (2017) 'Dietary Patterns and Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Prospective Studies', *The Journal of Nutrition*, 147(6), pp. 1174–1182. doi:10.3945/jn.116.242552.

Jansen, R., Han, L.K., Verhoeven, J.E., Aberg, K.A., van den Oord, E.C., Milaneschi, Y. and Penninx, B.W. (2021) 'An integrative study of five biological clocks in somatic and mental health', *eLife*, 10, p. e59479. doi:10.7554/eLife.59479.

Johnson, I.T. and Lund, E.K. (2007) 'Review article: nutrition, obesity and colorectal cancer', *Alimentary Pharmacology & Therapeutics*, 26(2), pp. 161–181. doi:10.1111/j.1365-2036.2007.03371.x.

Joshi, R., Wannamethee, S.G., Engmann, J., Gaunt, T., Lawlor, D.A., Price, J., Papacosta, O., Shah, T., Tillin, T., Chaturvedi, N., Kivimaki, M., Kuh, D., Kumari, M., Hughes, A.D., Casas, J.P., Humphries, S., Hingorani, A.D. and Schmidt, A.F. (2020) 'Triglyceride-containing lipoprotein sub-fractions and risk of coronary heart disease and stroke: A prospective analysis in 11,560 adults', *European Journal of Preventive Cardiology*, 27(15), pp. 1617–1626. doi:10.1177/2047487319899621.

Julkunen, H., Cichońska, A., Slagboom, P.E., Würtz, P., and Nightingale Health UK Biobank Initiative (2021) 'Metabolic biomarker profiling for identification of susceptibility to severe pneumonia and COVID-19 in the general population', *eLife*, 10, p. e63033. doi:10.7554/eLife.63033.

Jylhävä, J., Pedersen, N.L. and Hägg, S. (2017) 'Biological Age Predictors', *EBioMedicine*, 21, pp. 29–36. doi:10.1016/j.ebiom.2017.03.046.

Kahleova, H., Matoulek, M., Bratova, M., Malinska, H., Kazdova, L., Hill, M. and Pelikanova, T. (2013) 'Vegetarian diet-induced increase in linoleic acid in serum phospholipids is associated with improved insulin sensitivity in subjects with type 2 diabetes', *Nutrition & Diabetes*, 3(6), pp. e75–e75. doi:10.1038/nutd.2013.12.

Kanerva, N., Kaartinen, N.E., Rissanen, H., Knekt, P., Eriksson, J.G., Sääksjärvi, K., Sundvall, J. and Männistö, S. (2014) 'Associations of the Baltic Sea diet with cardiometabolic risk factors – a meta-analysis of three Finnish studies', *British Journal of Nutrition*, 112(4), pp. 616–626. doi:10.1017/S0007114514001159.

Kettunen, J., Demirkan, A., Würtz, P., Draisma, H.H.M., Haller, T., Rawal, R., Vaarhorst, A., Kangas, A.J., Lyytikäinen, L.-P., Pirinen, M., Pool, R., Sarin, A.-P., Soininen, P., Tukiainen, T., Wang, Q., Tiainen, M., Tynkkynen, T., Amin, N., Zeller, T., Beekman, M., Deelen, J., van Dijk, K.W., Esko, T., Hottenga, J.-J., van Leeuwen, E.M., Lehtimäki, T., Mihailov, E., Rose, R.J., de Craen, A.J.M., Gieger, C., Kähönen, M., Perola, M., Blankenberg, S., Savolainen, M.J., Verhoeven, A., Viikari, J., Willemsen, G., Boomsma, D.I., van Duijn, C.M., Eriksson, J., Jula, A., Järvelin, M.-R., Kaprio, J., Metspalu, A., Raitakari, O., Salomaa, V., Slagboom, P.E., Waldenberger, M., Ripatti, S. and Ala-Korpela, M. (2016) 'Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA', *Nature Communications*, 7(1), p. 11122. doi:10.1038/ncomms11122.

Key, T.J., Schatzkin, A., Willett, W.C., Allen, N.E., Spencer, E.A. and Travis, R.C. (2004) 'Diet, nutrition and the prevention of cancer', *Public Health Nutrition*, 7(1a), pp. 187–200. doi:10.1079/PHN2003588.

Kim, H., Hu, E.A., E Wong, K., Yu, B., Steffen, L.M., Seidelmann, S.B., Boerwinkle, E., Coresh, J. and Rebholz, C.M. (2021) 'Serum Metabolites Associated with Healthy Diets in African Americans and European Americans', *The Journal of Nutrition*, 151(1), pp. 40–49. doi:10.1093/jn/nxaa338.

Knaapila, A., Tuorila, H., Silventoinen, K., Keskitalo, K., Kallela, M., Wessman, M., Peltonen, L., Cherkas, L.F., Spector, T.D. and Perola, M. (2007) 'Food neophobia shows heritable variation in humans', *Physiology & Behavior*, 91(5), pp. 573–578. doi:10.1016/j.physbeh.2007.03.019.

Knaapila, A.J., Sandell, M.A., Vaarno, J., Hoppu, U., Puolimatka, T., Kaljonen, A. and Lagström, H. (2015) 'Food neophobia associates with lower dietary quality and higher BMI in Finnish adults', *Public Health Nutrition*, 18(12), pp. 2161–2171. doi:10.1017/S1368980014003024.

Kornsteiner, M., Singer, I. and Elmadfa, I. (2008) 'Very Low n–3 Long-Chain Polyunsaturated Fatty Acid Status in Austrian Vegetarians and Vegans', *Annals of Nutrition and Metabolism*, 52(1), pp. 37–47. doi:10.1159/000118629.

Krištić, J., Vučković, F., Menni, C., Klarić, L., Keser, T., Beceheli, I., Pučić-Baković, M., Novokmet, M., Mangino, M., Thaqi, K., Rudan, P., Novokmet, N., Šarac, J., Missoni, S., Kolčić, I., Polašek, O., Rudan, I., Campbell, H., Hayward, C., Aulchenko, Y., Valdes, A., Wilson, J.F., Gornik, O., Primorac, D., Zoldoš, V., Spector, T. and Lauc, G. (2014) 'Glycans Are a Novel Biomarker of Chronological and Biological Ages', *The Journals of Gerontology: Series A*, 69(7), pp. 779–789. doi:10.1093/gerona/glt190.

Larsson, S.C. (2021) 'Mendelian randomization as a tool for causal inference in human nutrition and metabolism', *Current Opinion in Lipidology*, 32(1), pp. 1–8. doi:10.1097/MOL.0000000000000721.

Larsson, S.C., Burgess, S., Mason, A.M. and Michaëlsson, K. (2020) 'Alcohol Consumption and Cardiovascular Disease: A Mendelian Randomization Study', *Circulation: Genomic and Precision Medicine*, 13(3), p. e002814. doi:10.1161/CIRCGEN.119.002814.

Lee, S.J., Teunissen, C.E., Pool, R., Shipley, M.J., Teumer, A., Chouraki, V., Melo van Lent, D., Tynkkynen, J., Fischer, K., Hernesniemi, J., Haller, T., Singh-Manoux, A., Verhoeven, A., Willemsen, G., Leeuw, F.A., Wagner, H., Dongen, J., Hertel, J., Budde, K., Willems van Dijk, K., Weinhold, L., Ikram, M.A., Pietzner, M., Perola, M., Wagner, M., Friedrich, N., Slagboom, P.E., Scheltens, P., Yang, Q., Gertzen, R.E., Egert, S., Li, S., Hankemeier, T., Beijsterveldt, C.E.M., Vasan, R.S., Maier, W., Peeters, C.F.W., Jörgen Grabe, H., Ramirez, A., Seshadri, S., Metspalu, A., Kivimäki, M., Salomaa, V., Demirkan, A., Boomsma, D.I., Flier, W.M., Amin, N. and Duijn, C.M. (2018) 'Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies', *Alzheimer's & Dementia*, 14(6), pp. 707–722. doi:10.1016/j.jalz.2017.11.012.

Lee, Y.H. (2018) 'Investigating the possible causal association of coffee consumption with osteoarthritis risk using a Mendelian randomization analysis', *Clinical Rheumatology*, 37(11), pp. 3133–3139. doi:10.1007/s10067-018-4252-6.

Lehallier, B., Gate, D., Schaum, N., Nanasi, T., Lee, S.E., Yousef, H., Moran Losada, P., Berdnik, D., Keller, A., Verghese, J., Sathyan, S., Franceschi, C., Milman, S., Barzilai, N. and Wyss-Coray, T. (2019) 'Undulating changes in human plasma proteome profiles across the lifespan', *Nature Medicine*, 25(12), pp. 1843–1850. doi:10.1038/s41591-019-0673-2.

Lehallier, B., Shokhirev, M.N., Wyss-Coray, T. and Johnson, A.A. (2020) 'Data mining of human plasma proteins generates a multitude of highly predictive aging clocks that reflect different aspects of aging', *Aging Cell*, 19(11). doi:10.1111/acel.13256.

Lehtovirta, M., Pahkala, K., Niinikoski, H., Kangas, A.J., Soininen, P., Lagström, H., Viikari, J.S.A., Rönnemaa, T., Jula, A., Ala-Korpela, M., Würtz, P. and Raitakari, O.T. (2018) 'Effect of Dietary Counseling on a Comprehensive Metabolic Profile from Childhood to Adulthood', *The Journal of Pediatrics*, 195, pp. 190-198.e3. doi:10.1016/j.jpeds.2017.11.057.

Leslie, M. (2017) 'To help save the heart, is it time to retire cholesterol tests?', *Science*, 358(6368), pp. 1237–1238. doi:10.1126/science.358.6368.1237.

Levine, M.E., Lu, A.T., Quach, A., Chen, B.H., Assimes, T.L., Bandinelli, S., Hou, L., Baccarelli, A.A., Stewart, J.D., Li, Y., Whitsel, E.A., Wilson, J.G., Reiner, A.P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L. and Horvath, S. (2018) 'An epigenetic biomarker of aging for lifespan and healthspan', *Aging*, 10(4), pp. 573–591. doi:10.18632/aging.101414.

Li, J., Guasch-Ferré, M., Chung, W., Ruiz-Canela, M., Toledo, E., Corella, D., Bhupathiraju, S.N., Tobias, D.K., Tabung, F.K., Hu, J., Zhao, T., Turman, C., Feng, Y.-C.A., Clish, C.B., Mucci, L., Eliassen, A.H., Costenbader, K.H., Karlson, E.W., Wolpin, B.M., Ascherio, A., Rimm, E.B., Manson, J.E., Qi, L., Martínez-González, M.Á., Salas-Salvadó, J., Hu, F.B. and Liang, L. (2020) 'The Mediterranean diet, plasma metabolome, and cardiovascular disease risk', *European Heart Journal*, 41(28), pp. 2645–2656. doi:10.1093/eurheartj/ehaa209.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M. and Kroemer, G. (2013) 'The Hallmarks of Aging', *Cell*, 153(6), pp. 1194–1217. doi:10.1016/j.cell.2013.05.039.

Luo, A., Jung, J., Longley, M., Rosoff, D.B., Charlet, K., Muench, C., Lee, J., Hodgkinson, C.A., Goldman, D., Horvath, S., Kaminsky, Z.A. and Lohoff, F.W. (2020) 'Epigenetic aging is accelerated in alcohol use disorder and regulated by genetic variation in APOL2', *Neuropsychopharmacology*, 45(2), pp. 327–336. doi:10.1038/s41386-019-0500-y.

Mantzioris, E., Muhlhausler, B.S. and Villani, A. (2022) 'Impact of the Mediterranean Dietary pattern on n-3 fatty acid tissue levels–A systematic review.', *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 176, p. 102387. doi:10.1016/j.plefa.2021.102387.

Marioni, R.E., Shah, S., McRae, A.F., Chen, B.H., Colicino, E., Harris, S.E., Gibson, J., Henders, A.K., Redmond, P., Cox, S.R., Pattie, A., Corley, J., Murphy, L., Martin, N.G., Montgomery, G.W., Feinberg, A.P., Fallin, M.D., Multhaup, M.L., Jaffe, A.E., Joehanes, R., Schwartz, J., Just, A.C., Lunetta, K.L., Murabito, J.M., Starr, J.M., Horvath, S., Baccarelli, A.A., Levy, D., Visscher, P.M., Wray, N.R. and Deary, I.J. (2015) 'DNA methylation age of blood predicts all-cause mortality in later life', *Genome Biology*, 16(1), p. 25. doi:10.1186/s13059-015-0584-6.

May-Wilson, S., Matoba, N., Wade, K.H., Hottenga, J.-J., Concas, M.P., Mangino, M., Grzeszkowiak, E.J., Menni, C., Gasparini, P., Timpson, N.J., Veldhuizen, M.G., de Geus, E., Wilson, J.F. and Pirastu, N. (2022) 'Large-scale GWAS of food liking reveals genetic determinants and genetic correlations with distinct neurophysiological traits', *Nature Communications*, 13(1), p. 2743. doi:10.1038/s41467-022-30187-w.

Meddens, S.F.W., de Vlaming, R., Bowers, P., Burik, C.A.P., Linnér, R.K., Lee, C., Okbay, A., Turley, P., Rietveld, C.A., Fontana, M.A., Ghanbari, M., Imamura, F., McMahon, G., van der Most, P.J., Voortman, T., Wade, K.H., Anderson, E.L., Braun, K.V.E., Emmett, P.M., Esko, T., Gonzalez, J.R., Kiefte-de Jong, J.C., Langenberg, C., Luan, J., Muka, T., Ring, S., Rivadeneira, F., Snieder, H., van Rooij, F.J.A., Wolffenbuttel, B.H.R., 23andMe Research Team, EPIC- InterAct Consortium, Lifelines Cohort Study, Smith, G.D., Franco, O.H., Forouhi, N.G., Ikram, M.A., Uitterlinden, A.G., van Vliet-Ostaptchouk, J.V., Wareham, N.J., Cesarini, D., Harden, K.P., Lee, J.J., Benjamin, D.J., Chow, C.C. and Koellinger, P.D. (2021) 'Genomic analysis of diet composition finds novel loci and associations with health and lifestyle', *Molecular Psychiatry*, 26(6), pp. 2056–2069. doi:10.1038/s41380-020-0697-5.

Meiselman, H.L., King, S.C. and Gillette, M. (2010) 'The demographics of neophobia in a large commercial US sample', *Food Quality and Preference*, 21(7), pp. 893–897. doi:10.1016/j.foodqual.2010.05.009.

Menni, C., Zhai, G., MacGregor, A., Prehn, C., Römisch-Margl, W., Suhre, K., Adamski, J., Cassidy, A., Illig, T., Spector, T.D. and Valdes, A.M. (2013) 'Targeted metabolomics profiles are strongly correlated with nutritional patterns in women', *Metabolomics*, 9(2), pp. 506–514. doi:10.1007/s11306-012-0469-6.

Merino, J., Dashti, H.S., Sarnowski, C., Lane, J.M., Todorov, P.V., Udler, M.S., Song, Y., Wang, H., Kim, J., Tucker, C., Campbell, J., Tanaka, T., Chu, A.Y., Tsai, L., Pers, T.H., Chasman, D.I., Rutter, M.K., Dupuis, J., Florez, J.C. and Saxena, R. (2022) 'Genetic analysis of dietary intake identifies new loci and functional links with metabolic traits', *Nature Human Behaviour*, 6(1), pp. 155–163. doi:10.1038/s41562-021-01182-w.

Michielsen, C.C.J.R., Hangelbroek, R.W.J., Feskens, E.J.M. and Afman, L.A. (2019) 'Disentangling the Effects of Monounsaturated Fatty Acids from Other Components of a Mediterranean Diet on Serum Metabolite Profiles: A Randomized Fully Controlled Dietary Intervention in Healthy Subjects at Risk of the Metabolic Syndrome', *Molecular Nutrition & Food Research*, 63(9), p. 1801095. doi:10.1002/mnfr.201801095.

Neale, E.P., Batterham, M.J. and Tapsell, L.C. (2016) 'Consumption of a healthy dietary pattern results in significant reductions in C-reactive protein levels in adults: a meta-analysis', *Nutrition Research*, 36(5), pp. 391–401. doi:10.1016/j.nutres.2016.02.009.

Nordestgaard, A.T., Thomsen, M. and Nordestgaard, B.G. (2015) 'Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study', *International Journal of Epidemiology*, 44(2), pp. 551–565. doi:10.1093/ije/dyv083.

Noroozi, R., Ghafouri-Fard, S., Pisarek, A., Rudnicka, J., Spólnicka, M., Branicki, W., Taheri, M. and Pośpiech, E. (2021) 'DNA methylation-based age clocks: From age prediction to age reversion', *Ageing Research Reviews*, 68, p. 101314. doi:10.1016/j.arr.2021.101314.

Pierce, B.L. and Burgess, S. (2013) 'Efficient Design for Mendelian Randomization Studies: Subsample and 2-Sample Instrumental Variable Estimators', *American Journal of Epidemiology*, 178(7), pp. 1177–1184. doi:10.1093/aje/kwt084.

Pirastu, N., McDonnell, C., Grzeszkowiak, E.J., Mounier, N., Imamura, F., Merino, J., Day, F.R., Zheng, J., Taba, N., Concas, M.P., Repetto, L., Kentistou, K.A., Robino, A., Esko, T., Joshi, P.K., Fischer, K., Ong, K.K., Gaunt, T.R., Kutalik, Z., Perry, J.R.B. and Wilson, J.F. (2019) *Using genetic variation to disentangle the complex relationship between food intake and health outcomes*. preprint. Genetics. doi:10.1101/829952.

Pliner, P. and Hobden, K. (1992) 'Development of a scale to measure the trait of food neophobia in humans', *Appetite*, 19(2), pp. 105–120. doi:10.1016/0195-6663(92)90014-W.

Poole, R., Kennedy, O.J., Roderick, P., Fallowfield, J.A., Hayes, P.C. and Parkes, J. (2017) 'Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes', *BMJ*, p. j5024. doi:10.1136/bmj.j5024.

Popkin, B.M. (2001) 'The Nutrition Transition and Obesity in the Developing World', *The Journal of Nutrition*, 131(3), pp. 871S-873S. doi:10.1093/jn/131.3.871S.

Predieri, S., Sinesio, F., Monteleone, E., Spinelli, S., Cianciabella, M., Daniele, G.M., Dinnella, C., Gasperi, F., Endrizzi, I., Torri, L., Gallina Toschi, T., Bendini, A., Pagliarini, E., Cattaneo, C., Di Monaco, R., Vitaglione, P., Condelli, N. and Laureati, M. (2020) 'Gender, Age, Geographical Area, Food Neophobia and Their Relationships with the Adherence to the Mediterranean Diet: New Insights from a Large Population Cross-Sectional Study', *Nutrients*, 12(6), p. 1778. doi:10.3390/nu12061778.

Qian, F., Ardisson Korat, A.V., Imamura, F., Marklund, M., Tintle, N., Virtanen, J.K., Zhou, X., Bassett, J.K., Lai, H., Hirakawa, Y., Chien, K.-L., Wood, A.C., Lankinen, M., Murphy, R.A., Samieri, C., Pertiwi, K., de Mello, V.D., Guan, W., Forouhi, N.G., Wareham, N., Hu, I.C., Frank B., Riserus, U., Lind, L., Harris, W.S., Shadyab, A.H., Robinson, J.G., Steffen, L.M., Hodge, A., Giles, G.G., Ninomiya, T., Uusitupa, M., Tuomilehto, J., Lindström, J., Laakso, M., Siscovick, D.S., Helmer, C., Geleijnse, J.M., Wu, J.H.Y., Fretts, A., Lemaitre, R.N., Micha, R., Mozaffarian, D. and Sun, Q. (2021) 'n-3 Fatty Acid Biomarkers and Incident Type 2 Diabetes: An Individual Participant-Level Pooling Project of 20 Prospective Cohort Studies', *Diabetes Care*, 44(5), pp. 1133–1142. doi:10.2337/dc20-2426.

Quach, A., Levine, M.E., Tanaka, T., Lu, A.T., Chen, B.H., Ferrucci, L., Ritz, B., Bandinelli, S., Neuhouser, M.L., Beasley, J.M., Snetselaar, L., Wallace, R.B., Tsao, P.S., Absher, D., Assimes, T.L., Stewart, J.D., Li, Y., Hou, L., Baccarelli, A.A., Whitsel, E.A. and Horvath, S. (2017) 'Epigenetic clock analysis of diet, exercise, education, and lifestyle factors', *Aging*, 9(2), pp. 419–446. doi:10.18632/aging.101168.

Rabadán, A. and Bernabéu, R. (2021) 'A systematic review of studies using the Food Neophobia Scale: Conclusions from thirty years of studies', *Food Quality and Preference*, 93, p. 104241. doi:10.1016/j.foodqual.2021.104241.

Rebello, S.A. and van Dam, R.M. (2013) 'Coffee Consumption and Cardiovascular Health: Getting to the Heart of the Matter', *Current Cardiology Reports*, 15(10), p. 403. doi:10.1007/s11886-013-0403-1.

Revelle, W. (1979) 'Hierarchical Cluster Analysis And The Internal Structure Of Tests', *Multivariate Behavioral Research*, 14(1), pp. 57–74. doi:10.1207/s15327906mbr1401_4.

Revelle, W. and Zinbarg, R.E. (2009) 'Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma', *Psychometrika*, 74(1), pp. 145–154. doi:10.1007/s11336-008-9102-z.

Ronksley, P.E., Brien, S.E., Turner, B.J., Mukamal, K.J. and Ghali, W.A. (2011) 'Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis', *BMJ*, 342(feb22 1), pp. d671–d671. doi:10.1136/bmj.d671.

de Roos, B., Caslake, M.J., Stalenhoef, A.F., Bedford, D., Demacker, P.N., Katan, M.B. and Packard, C.J. (2001) 'The coffee diterpene cafestol increases plasma triacylglycerol by increasing the production rate of large VLDL apolipoprotein B in healthy normolipidemic subjects', *The American Journal of Clinical Nutrition*, 73(1), pp. 45–52. doi:10.1093/ajcn/73.1.45.

Ros, E., Martínez-González, M.A., Estruch, R., Salas-Salvadó, J., Fitó, M., Martínez, J.A. and Corella, D. (2014) 'Mediterranean Diet and Cardiovascular Health: Teachings of the PREDIMED Study', *Advances in Nutrition*, 5(3), pp. 330S-336S. doi:10.3945/an.113.005389.

Rosoff, D.B., Smith, G.D., Mehta, N., Clarke, T.-K. and Lohoff, F.W. (2019) *Evaluating the effects of alcohol and tobacco use on cardiovascular disease using multivariable Mendelian randomization*. preprint. Genomics. doi:10.1101/757146.

Rundblad, A., Holven, K.B., Ottestad, I., Myhrstad, M.C. and Ulven, S.M. (2017) 'High-quality fish oil has a more favourable effect than oxidised fish oil on intermediate-density lipoprotein and LDL subclasses: a randomised controlled trial', *British Journal of Nutrition*, 117(9), pp. 1291–1298. doi:10.1017/S0007114517001167.

Schaefer, E.J. (2002) 'Lipoproteins, nutrition, and heart disease', *The American Journal of Clinical Nutrition*, 75(2), pp. 191–212. doi:10.1093/ajcn/75.2.191.

Schwingshackl, L., Hoffmann, G., Lampousi, A.-M., Knüppel, S., Iqbal, K., Schwedhelm, C., Bechthold, A., Schlesinger, S. and Boeing, H. (2017) 'Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies', *European Journal of Epidemiology*, 32(5), pp. 363–375. doi:10.1007/s10654-017-0246-y.

Shu, X., Xiang, Y.-B., Rothman, N., Yu, D., Li, H.-L., Yang, G., Cai, H., Ma, X., Lan, Q., Gao, Y.-T., Jia, W., Shu, X.-O. and Zheng, W. (2018) 'Prospective study of blood metabolites associated with colorectal cancer risk: Colorectal cancer risk associated blood metabolites', *International Journal of Cancer*, 143(3), pp. 527–534. doi:10.1002/ijc.31341.

Sofi, F., Abbate, R., Gensini, G.F. and Casini, A. (2010) 'Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis', *The American Journal of Clinical Nutrition*, 92(5), pp. 1189–1196. doi:10.3945/ajcn.2010.29673.

Soininen, P., Kangas, A.J., Würtz, P., Tukiainen, T., Tynkkynen, T., Laatikainen, R., Järvelin, M.-R., Kähönen, M., Lehtimäki, T., Viikari, J., Raitakari, O.T., Savolainen, M.J. and Ala-Korpela, M. (2009) 'High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism', *The Analyst*, 134(9), p. 1781. doi:10.1039/b910205a.

Storey, J.D. (2003) 'The positive false discovery rate: a Bayesian interpretation and the q-value', *The Annals of Statistics*, 31(6). doi:10.1214/aos/1074290335.

Suhre, K., Meisinger, C., Döring, A., Altmaier, E., Belcredi, P., Gieger, C., Chang, D., Milburn, M.V., Gall, W.E., Weinberger, K.M., Mewes, H.-W., Hrabé de Angelis, M., Wichmann, H.-E., Kronenberg, F., Adamski, J. and Illig, T. (2010) 'Metabolic Footprint of Diabetes: A Multiplatform Metabolomics Study in an Epidemiological Setting', *PLoS ONE*. Edited by B. Breant, 5(11), p. e13953. doi:10.1371/journal.pone.0013953.

The AIM-HIGH Investigators (2011) 'Niacin in Patients with Low HDL Cholesterol Levels Receiving Intensive Statin Therapy', *New England Journal of Medicine*, 365(24), pp. 2255–2267. doi:10.1056/NEJMoa1107579.

Toth, P.P., Barter, P.J., Rosenson, R.S., Boden, W.E., Chapman, M.J., Cuchel, M., D'Agostino, R.B., Davidson, M.H., Davidson, W.S., Heinecke, J.W., Karas, R.H., Kontush, A., Krauss, R.M., Miller, M. and Rader, D.J. (2013) 'High-density lipoproteins: A consensus statement from the National Lipid Association', *Journal of Clinical Lipidology*, 7(5), pp. 484–525. doi:10.1016/j.jacl.2013.08.001.

Tuorila, H., Lähteenmäki, L., Pohjalainen, L. and Lotti, L. (2001) 'Food neophobia among the Finns and related responses to familiar and unfamiliar foods', *Food Quality and Preference*, 12(1), pp. 29–37. doi:10.1016/S0950-3293(00)00025-2.

Tynkkynen, J., Chouraki, V., Lee, S.J., Hernesniemi, J., Yang, Q., Li, S., Beiser, A., Larson, M.G., Sääksjärvi, K., Shipley, M.J., Singh-Manoux, A., Gerszten, R.E., Wang, T.J., Havulinna, A.S., Würtz, P., Fischer, K., Demirkan, A., Ikram, M.A., Amin, N., Lehtimäki, T., Kähönen, M., Perola, M., Metspalu, A., Kangas, A.J., Soininen, P., Ala-Korpela, M., Vasan, R.S., Kivimäki, M., Duijn, C.M., Seshadri, S. and Salomaa, V. (2018) 'Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts', *Alzheimer's & Dementia*, 14(6), pp. 723–733. doi:10.1016/j.jalz.2018.01.003.

Vitali, C., Khetarpal, S.A. and Rader, D.J. (2017) 'HDL Cholesterol Metabolism and the Risk of CHD: New Insights from Human Genetics', *Current Cardiology Reports*, 19(12), p. 132. doi:10.1007/s11886-017-0940-0.

Vojinovic, D., Kalaoja, M., Trompet, S., Fischer, K., Shipley, M.J., Li, S., Havulinna, A.S., Perola, M., Salomaa, V., Yang, Q., Sattar, N., Jousilahti, P., Amin, N., Satizabal, C.L., Taba, N., Sabayan, B., Vasan, R.S., Ikram, M.A., Stott, D.J., Ala-Korpela, M., Jukema, J.W., Seshadri, S., Kettunen, J., Kivimaki, M., Esko, T. and van Duijn, C.M. (2021) 'Association of Circulating Metabolites in Plasma or Serum and Risk of Stroke: Meta-analysis From 7 Prospective Cohorts', *Neurology*, 96(8), pp. e1110–e1123. doi:10.1212/WNL.0000000000011236.

Wang, T.J., Larson, M.G., Vasan, R.S., Cheng, S., Rhee, E.P., McCabe, E., Lewis, G.D., Fox, C.S., Jacques, P.F., Fernandez, C., O'Donnell, C.J., Carr, S.A., Mootha, V.K., Florez, J.C., Souza, A., Melander, O., Clish, C.B. and Gerszten, R.E. (2011) 'Metabolite profiles and the risk of developing diabetes', *Nature Medicine*, 17(4), pp. 448–453. doi:10.1038/nm.2307.

Würtz, P., Cook, S., Wang, Q., Tiainen, M., Tynkkynen, T., Kangas, A.J., Soininen, P., Laitinen, J., Viikari, J., Kähönen, M., Lehtimäki, T., Perola, M., Blankenberg, S., Zeller, T., Männistö, S., Salomaa, V., Järvelin, M.-R., Raitakari, O.T., Ala-Korpela, M. and Leon, D.A. (2016) 'Metabolic profiling of alcohol consumption in 9778 young adults', *International Journal of Epidemiology*, 45(5), pp. 1493–1506. doi:10.1093/ije/dyw175.

Würtz, P., Havulinna, A.S., Soininen, P., Tynkkynen, T., Prieto-Merino, D., Tillin, T., Ghorbani, A., Artati, A., Wang, Q., Tiainen, M., Kangas, A.J., Kettunen, J., Kaikkonen, J., Mikkilä, V., Jula, A., Kähönen, M., Lehtimäki, T., Lawlor, D.A., Gaunt, T.R., Hughes, A.D., Sattar, N., Illig, T., Adamski, J., Wang, T.J., Perola, M., Ripatti, S., Vasan, R.S., Raitakari, O.T., Gerszten, R.E., Casas, J.-P., Chaturvedi, N., Ala-Korpela, M. and Salomaa, V. (2015) 'Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of 3 Population-Based Cohorts', *Circulation*, 131(9), pp. 774–785. doi:10.1161/CIRCULATIONAHA.114.013116.

Würtz, P., Kangas, A.J., Soininen, P., Lawlor, D.A., Davey Smith, G. and Ala-Korpela, M. (2017) 'Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale Epidemiology: A Primer on -Omic Technologies', *American Journal of Epidemiology*, 186(9), pp. 1084–1096. doi:10.1093/aje/kwx016.

Würtz, P., Wang, Q., Kangas, A.J., Richmond, R.C., Skarp, J., Tiainen, M., Tynkkynen, T., Soininen, P., Havulinna, A.S., Kaakinen, M., Viikari, J.S., Savolainen, M.J., Kähönen, M., Lehtimäki, T., Männistö, S., Blankenberg, S., Zeller, T., Laitinen, J., Pouta, A., Mäntyselkä, P., Vanhala, M., Elliott, P., Pietiläinen, K.H., Ripatti, S., Salomaa, V., Raitakari, O.T., Järvelin, M.-R., Smith, G.D. and Ala-Korpela, M. (2014) 'Metabolic Signatures of Adiposity in Young Adults: Mendelian Randomization Analysis and Effects of Weight Change', *PLoS Medicine*. Edited by N.A. Sheehan, 11(12), p. e1001765. doi:10.1371/journal.pmed.1001765.

Yang, Q., Lin, S.L., Au Yeung, S.L., Kwok, M.K., Xu, L., Leung, G.M. and Schooling, C.M. (2017) 'Genetically predicted milk consumption and bone health, ischemic heart disease and type 2 diabetes: a Mendelian randomization study', *European Journal of Clinical Nutrition*, 71(8), pp. 1008–1012. doi:10.1038/ejcn.2017.8.

Yi, S.-W., Park, H.-B., Jung, M.-H., Yi, J.-J. and Ohrr, H. (2022) 'High-density lipoprotein cholesterol and cardiovascular mortality: a prospective cohort study among 15.8 million adults', *European Journal of Preventive Cardiology*, 29(5), pp. 844–854. doi:10.1093/eurjpc/zwab230.

Zhang, W.-G., Zhu, S.-Y., Bai, X.-J., Zhao, D.-L., Jiang, S.-M., Li, J., Li, Z.-X., Fu, B., Cai, G.-Y., Sun, X.-F. and Chen, X.-M. (2014) 'Select aging biomarkers based on telomere length and chronological age to build a biological age equation', *AGE*, 36(3), p. 9639. doi:10.1007/s11357-014-9639-y.

Zhao, Q., Wang, J., Hemani, G., Bowden, J. and Small, D.S. (2020) 'Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score', *The Annals of Statistics*, 48(3). doi:10.1214/19-AOS1866.

Zou, H. and Hastie, T. (2005) 'Regularization and variable selection via the elastic net', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), pp. 301–320. doi:10.1111/j.1467-9868.2005.00503.x.

ACKNOWLEDGMENTS

This PhD has been quite a rollercoaster and has shaped me as a person in more ways than I can think of. Every journey has a beginning and for the beginning of this one I have to thank my supervisor Andres who invited me to join the amazing team of Genome Center when I was about to graduate my masters. For this I am forever grateful, I cannot imagine a better working environment for pursuing a PhD than the one that you have created for us – a one where every person is amazing and inspiring and always willing to help. Thank you for believing in me from day one and thank you for sharing with me your wisdom and world-view in our long spontaneous conversations.

The first person I had a longer talk with in the Genome Center about my future, is my supervisor Krista. Thinking back of that time now, I feel a tad bit embarrassed about how beginner I was in the science field, but mostly feel happy that you decided to invest your time to thoroughly introduce me to the wonderful world of statistics and calmly explain when I took my time to understand. I am truly grateful for all our amazing conference trips and that you have eagerly helped me to build a science network of top-notch researchers from all over the world. I admire how smart you are and how much you know about so many things while keeping it chill when we chat about random topics during the off-work time! I am happy that you have decided to join me quite a few times on orienteering or skiing tracks and have truly enjoyed the long conversations with a cup of tea or a glass of wine. I hope this friendship continues for long!

Likewise, I hope my friendship with my supervisor Tõnu does not end with defending this thesis. Thank you for introducing me to the world of start-ups and for being so inspirational. I am so happy that you have always believed in me and that you have expressed it every now and then with some nice words (I know it is cheesy, but I do think that everybody needs some nice words from time to time and they do help to get by when times are a bit rough). I admire your superfast thinking and ability to actually multitask. Furthermore, thank you for introducing me enthusiastically to your science and start-up network – it means more to me than I (as a usual Estonian) express it. And, of course, thank you for the amazing conversations about science and life in general.

I have mentioned several times how grateful I am for the scientific network my supervisors have created me. However, I am particularly happy for one person they introduced me to: my supervisor Nicola, who joined halfway through the PhD. Nicola, you truly have ingenious mind and ideas, and I am grateful that you have shared these ideas with me and taken time to teach, explain and discuss even if I have needed three rounds of explanations before I am able to grasp it all. Thank you for motivating me to keep pushing and not give up even when the review-process can seem inconvenient and take ages. You have definitely shown me that perseverance ultimately leads to goals!

However, all these mind-opening conversations with Nicola were possible because I had the amazing opportunity to collaborate with Jim Wilson's research

group. Herewith I am thankful that there are scholarships such as Dora Plus that provide the wonderful opportunity to visit research groups abroad and make the PhD experience so much more fruitful (definitely recommend to fellow PhD students!). My abroad-research-adventure was at the Prof. Jim Wilson's group in The University of Edinburgh and I cannot really think of a more superb fit for me for such visit than these four months were. Thank you, Jim, for such a warm welcome and for introducing me to this group of brilliant people! And my dear Edinburgh friends, even though you made sure to remind me **every day** how "reaaaaaaaally-really old" I am (or maybe because of it), it is hard to express in words how much it meant to me that you made me super-welcome from day one, made me laugh to tears so-so many times and were helpful in every way possible. You are always welcome to visit me again in Estonia and I hope our roads will cross many more times!

Edinburgh group was my second science-family, but the first is and will always be my lovely Genome Center people. How is it even possible that I have been this crazy lucky to be working together with so many inspiring, smart and cool scientists? I have thought about it oh so many times during the past years and am just sincerely happy about it! And, moreover, Genome Center is full of astonishing scientists, who on the same time are extremely fun to spend time with. I can't help but smile when I think of the countless conversations in the "kohvi-nurk" or dancing and laughing the night away on several events. My heartfelt gratitude to all of you for these amazing moments!

In the lovely-lovely Genome Center family there are some people that have been especially impactful during this journey. My first office room-mates Silva and Kristi, thank you for gently introducing me to the science-world and helping me settle in by answering all my questions about statistics, where-what-is, how to get the most out of conferences and how things work in general. You made me feel really welcome and made the settling in process supersmooth! I always have and probably always will look up to you as statisticians and am happy to have such smart women as my friends! I have been lucky to have as well a second set of astonishingly smart statisticians as office room-mates: Merli and Oliver. Thank you for all the fun times and thanks for being always so helpful when I arrive with yet another question about things that I want to understand (I have a hunch that I might continue with these questions in the years to come…I can't help it that you two are so smart)!

The PhD students in the Genome Center are smart (in fact brilliant) for sure, but it is also important that among them it really feels like family! Some of you are not students anymore and are already proud PhDs, but it still feels like you have always been my fellow PhD students whom I have relied to in so many situations. Kreete, Krebsu, Maarja, Anette, Marili and Krigul – thank you for always being there and thank you for sharing these PhD years with me, shoulder by shoulder. It has been awesome with you through all these years: may it be explaining Mendelian Randomization to you in the middle of the night after watching Eurovision together, or hiking in the knee-deep snow in the freezing

cold – it is always with great humour and mutual understanding, and I am grateful to have had you as my fellow PhD students during this journey.

What I really-really like about Genome Center is that it is not only PhD students who support each-other, but every person in the house is super-mega-giga helpful. May it be senior scientists who you turn to, when you have reached the border of your own skills and wisdom (thank you Reedik!); or administrative staff who are always happily helping when you really don't know how things work or where to go; or all our fun colleagues who are always willing to brain-storm or discuss random topics. Herewith I would like to thank Kelli for several motivational speeches and for being in general such a chill and awesome person! Thank you, Lili, for always listening and giving advice when I have felt like some situations are too much to handle. I know you get that a lot, but I want to say it too: you are truly inspirational and a role model! Thank you, Elin, for the discussions and advice during our long walks (we should start running together again!!), it has been of great support and motivation to me. Thank you, Martin and Karoliina and our communication team, for so eagerly supporting my first steps in communicating with media – it has been extremely helpful and calming. I am so grateful for the privilege that we have in Genome Center – somebody actually prepares us for communicating with media (which otherwise would have been quite stressful at least for me). And, what is truly amazing: sometimes you don't even have to ask and lovely colleagues already come and help you. Thank you, Natàlia, Teele and Erik, for all the support during the thesis-writing. I cannot express in words how much it means to me! And Erik, you one special boy: always there, always helping, always know everything, always listening, and always up for a lunch burger! I am superhappy to have found such a talented and amazing friend among my colleagues!

And, of course, my dear friends outside of academia! I am happy to have quite many of you and this has definitely been of great help to keep pursuing my goals. Thank you for being there for me and thank you for keeping my motivation and spirit up! It truly makes me giggle and smile when I think about how enthusiastic some of you have been about my PhD! Kaia for example is completely familiar with the peer-review process by now, because she has been super-eagerly waiting for my PhD party. Or Piret and Dagny who have been so ready to celebrate it that they have already prepared a quiz for the party (and obviously chosen the dresses long ago). Or Siim, who for some reason still has not given up on me and keeps asking after every three or four days "So!?? Is it done now!??" and sending me pictures of all the fun stuff that I can do after I am ready with the thesis. Thank you all for this cheerleading that brings extra bit of motivation! Thank you, Janika, for always being there and for cheering me with the countless hours of phone-calls when I had some rough times and also during the happy times. Aitäh-aitäh, Mari-Ann, for listening and cheering when I needed it, for shutting me up when I started to doubt myself too much, and for making me laugh on a daily basis! Thank you, Sven, for all the long walks and discussions, especially during the time of writing the thesis! Thank you, Liina and Dorel, for always being so interested in what I have found with my research or what I have read in the new

scientific papers – it is amazing and encouraging to see that my work spans interest outside of academia as well! Thank you, Jaksi, for continuously keeping on track of how my PhD is going and for being such an amazing person! And, thank you, Hanna-Kristel, I am amazed by your knowledge in the medical field and am delighted that you have taken the countless hours to share some of it with me. I admire your enthusiasm and devotion and feel superlucky that you have been so invested in helping me pursue my goals.

As I said, every journey has a beginning. I am pretty sure my research interest began way before the PhD. For that I am forever grateful to my mom, who gently provided the entrance to the scientific world by inviting tiny Nele to listen to presentations and lectures (10-year old Nele was rather enthusiastic about listening to the presentations about neurons), having scientific posters on the walls at home, and discussing interesting topics on a daily basis. As the years went by, she included me to actual scientific projects and it has grown into a mutual collaboration. I look up to you in so many ways and am sincerely grateful for all the support you have provided! If tiny Nele was interested in listening to presentations, then supertiny Nele was often with granny. For many my colleagues she does not need any introduction, but for those who don't know: my granny worked in the 90s in the very same house where I work now (that's where 5-year-old me got to play Tetris and Pacman for the first time) and we even share some colleagues! She has always been eager to fulfill my (and other siblings') curiosity by explaining everything, telling stories and demonstrating that if you don't know something, you can always look it up! And, what is also really important, she has kept me well-fed throughout my PhD with her amazing cooking skills! My family is rather large and it would take quite a few pages to express how grateful I am to them and how happy I am about them. So, let me just say that I am superproud to have such an amazing, inspiring and supportive family and I feel really lucky to have had such a powerful and motivating support-system surrounding me. I try to do my best that you could feel the same way! Thank you!

Inimesed minu ümber on nii ilusad ja head, et lausa rõõmsaks teeb! Aitäh!

PUBLICATIONS

CURRICULUM VITAE

Name: Nele Taba
Date of birth: May 12th 1989
Contact: Estonian Genome Center, University of Tartu
Riia 23B, 20101, Tartu, Estonia
E-mail: nele.taba@ut.ee

Education:

2017–…	PhD, Molecular and Cell Biology, University of Tartu, Estonia
2018	Visiting PhD student in Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Scotland, The United Kingdom
2014–2016	Master of Science, Methodology and Statistics for the Behavioural, Biomedical and Social Sciences, Utrecht University, The Netherlands
2011	Exchange student (BA in Economics and Business Administration), Jönköping International Business School, Sweden
2009–2013	Bachelor of Arts, Psychology, University of Tartu, Estonia
2008–2013	Bachelor of Arts, Economics and Business Administration, University of Tartu, Estonia
2005–2008	Hugo Treffner Gymnasium (with specialisation in maths)
2003	Gymnasium Einsiedel (exchange student in Germany for half a year in 8th grade)
1996–2005	Tartu Tamme Gymnasium

Professional employment:

02/2020–…	Institute of Mathematics and Statistics, University of Tartu, lab session instructor of course "LTMS.00.025" Statistical Data Science
09/2016–…	Estonian Genome Center, University of Tartu, specialist in biostatistics
01/2016–02/2016	University of Wisconsin-Madison, Departments of Biostatistics and Medical Informatics, and Population Health Sciences, internship in handling missing data
05/2015–07/2015	University of Tartu, Data Analyst in Research and Innovation Policy Monitoring Programme
11/2013–05/2014	Statoil Fuel & Retail Eesti AS, Finance and Credit Controller
07/2013–11/2013	Statoil Fuel & Retail Eesti AS, Junior Controller
05/2012–06/2012	Ministry of Finance of the Republic of Estonia, Analytics internship in Insurance Policy

Administrative work:

2019 Organizer of a discussion topic in Opinion Festival (topic of discussion: "Genetic data: threat or opportunity?"

2018–… Member of the Nordic-Baltic Region of the International Biometric Society

Publications:

Macdonald-Dunlop, E.; **Taba, N.**; Klarić, L.; Frkatović, A.; Walker, R.; Hayward, C.; Esko, T.; Haley, C.; Fischer, K.; Wilson, J.F.; Joshi, P.K. (2022). A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk. Aging, 623–659. DOI: 10.18632/aging.203847.

Gobeil, É.; Maltais-Payette, I.; **Taba, N.**; Brière, F.; Ghodsian, N.; Abner, E.; Bourgault, J.; Gagnon, E.; Manikpurage, H.D.; Couture, C.; Mitchell, P.L.; Mathieu, P.; Julien, F.; Corbeil, J.; Vohl, M.-C.; Thériault, S.; Esko, T.; Tchernof, A.; Arsenault, B.J. (2022). Mendelian Randomization Analysis Identifies Blood Tyrosine Levels as a Biomarker of Non-Alcoholic Fatty Liver Disease. Metabolites, 12 (5), 440. DOI: 10.3390/metabo12050440.

Pirastu, N.; McDonnell, C.; Grzeszkowiak, E. J.; Mounier, N.; Imamura, F.; Merino, J.; Day, Felix R.; Zheng, J.; **Taba, N.**; Concas, M. P.; Repetto, L.; Kentistou, K. A.; Robino, A.; Esko, T.; Joshi, P. K.; Fischer, K.; Ong, K. K.; Gaunt, T. R.; Kutalik, Z.; Perry, J. R. B. ... Wilson, J. F. (2022). Using genetic variation to disentangle the complex relationship between food intake and health outcomes. PLoS Genetics, 18 (6), e1010162. DOI: 10.1371/journal.pgen.1010162.

Schneider, S.; **Taba, N.**; Saapar, M.; Vibo, R.; Kõrv, J. (2021). Determinants of long-term health-related quality of life in young ischemic stroke patients. Journal of Stroke and Cerebrovascular Diseases, 30, 105499. DOI: 10.1016/j.jstrokecerebrovasdis.2020.105499.

Vojinovic, D.; Kalaoja, M.; Trompet, S.; Fischer, K.; Shipley, M.J.; Li, S.; Havulinna, A.S.; Perola, M.; Salomaa, V.; Yang, Q.; Sattar, N.; Jousilahti, P.; Amin, N.; Satizabal, C.L.; **Taba, N.**; Sabayan, B.; Vasan, R.S.; Ikram, M.A.; Stott, D.J.; Ala-Korpela, M. ... van Duijn, C.M. (2021). Association of Circulating Metabolites in Plasma or Serum and Risk of Stroke. Neurology, 96 (8), e1110–e1123. DOI: 10.1212/WNL.0000000000011236.

Taba, N.; Valge, H.-K.; Metspalu, A.; Esko, T.; Wilson, J.F.; Fischer, K.; Pirastu, N. (2021). Mendelian Randomization Identifies the Potential Causal Impact of Dietary Patterns on Circulating Blood Metabolites. Frontiers in Genetics, 12, ARTN 738265. DOI: 10.3389/fgene.2021.738265.

Ghodsian, N.; Abner, E.; Emdin, C.A.; Gobeil, É.; **Taba, N.**; Haas, M.E.; Perrot, N.; Manikpurage, H.D.; Gagnon, É.; Bourgault, J.; St-Amand, A.; Couture, C.; Mitchell, P.L.; Bossé, Y.; Mathieu, P.; Vohl, M.-C.; Tchernof, A.; Thériault, S.; Khera, A.V.; Esko, T. ... Arsenault, B.J. (2021). Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. Cell Reports Medicine, 2 (11), 100437. DOI: 10.1016/j.xcrm.2021.100437.

Schneider, S.; Vibo, R.; **Taba, N.**; Kõrv, J. (2020). Mortality in young adult patients with acute ischaemic stroke. Acta Neurologica Scandinavica, 141 (3), 242–249. DOI: 10.1111/ane.13217.

Kadastik-Eerme, L.; **Taba, N.**; Asser, T-; Taba, P. (2019). Incidence and Mortality of Parkinson's Disease in Estonia. Neuroepidemiology, 53 (1–2), 63–72. DOI: 10.1159/000499756.

Sarin, H.V.; **Taba, N.**; Fischer, K.; Esko, T.; Kanerva, N.; Moilanen, L.; Saltevo, J.; Joensuu, A.; Borodulin, K.; Männistö, S.; Kristiansson, K.; Perola, M. (2019). Food neophobia associates with poorer dietary quality, metabolic risk factors, and increased disease outcome risk in population-based cohorts in a metabolomics study. American Journal of Clinical Nutrition, 110 (1), 233–245. DOI: 10.1093/ajcn/nqz100.

Kadastik-Eerme L., **Taba N.**, Asser T., Taba P. (2018). The increasing prevalence of Parkinson's disease in Estonia. Acta Neurologica Scandinavica, 138 (3), 251–258. DOI: 10.1111/ane.12948.

Kadastik-Eerme L.; **Taba N.**; Asser T.; Taba P. (2017). Factors associated with motor complications in Parkinson's disease. Brain and Behavior, 7 (10), e00827–e00827. DOI: 10.1002/brb3.837.

Kadastik-Eerme L.; Muldmaa M.; Lilles S.; Rosenthal M;, **Taba N.**; Taba, P. (2016). Non-motor features in Parkinson's disease: What are the most important associated factors? Parkinson's Disease, 9, 1–8. DOI: 10.1155/2016/4370674.

Supervised dissertations:

2018 Co-supervision of the master's thesis of Marili Zimmermann "Survival Analysis of Left Truncated Data and with Time-Dependent Variable, Based on Estonian Genome Center's Cohort" (Mathematics and Statistics)

Awards and stipends:

2021 Best oral presentation at the annual conference of the Institute of Molecular and Cell Biology (University of Tartu, Tartu, Estonia)

2021 Finalist of the lecture competition "Science in 3 minutes" of Estonian Academy of Sciences

2019 Scholarship of the Graduate School in Biomedicine and Biotechnology: oral presentation in The Nordic Society of Human Genetics and Precision Medicine 2019 meeting (Satellite Symposium to ESHG) (Gothenburg, Sweden)

2019 Scholarship of the Graduate School in Biomedicine and Biotechnology: participating in UseR!2019 course (Tolouse, France)

2018 Best poster presentation at the European Mathematical Genetics Meeting (Cagliari, Italy)

2018 Dora Plus short study visits scholarship

2018 Dora Plus PhD student mobility scholarship

2017 Scholarship in smart specialization growth areas from Archimedes Foundation
2017 Best poster presentation in Workshop on Recent & Future Trends in Biostatistics (Cambridge University, Cambridge, United Kingdom)
2017 Best oral presentation at the annual conference of the Institute of Molecular and Cell Biology (University of Tartu, Tartu, Estonia)
2015 Kristjan Jaak formal studies scholarship (year 2015/2016)
2013 Best bachelor thesis of the year 2013 on the management track in the Institute of Business Administration

ELULOOKIRJELDUS

Nimi: Nele Taba
Sünniaeg: 12. mai 1989
Aadress: Tartu Ülikooli Eesti Geenivaramu
Riia 23B, 51010, Tartu, Eesti
E-post: nele.taba@ut.ee

Haridus:

2017–…	Doktoriõpe, molekulaar- ja rakubioloogia, Tartu Ülikool, Eesti
2018	Külalisdoktorant, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Šotimaa, Ühendkuningriik
2014–2016	Magistriõpe (MSc), Methodology and Statistics for the Behavioural, Biomedical and Social Sciences, Utrecht University, Holland
2011	Vahetustudeng (majandusteaduse bakalaureuseõpe), Jönköping International Business School, Rootsi
2009–2013	Bakalaureuseõpe, psühholoogia, Tartu Ülikool, Eesti
2008–2013	Bakalaureuseõpe, majandusteadus, Tartu Ülikool, Eesti
2005–2008	Hugo Treffneri Gümnaasium (reaalharu)
2003	Gymnasium Einsiedel (vahetusõpilane Saksamaal kaheksandas klassis pool aastat)
1996–2005	Tartu Tamme Gümnaasium

Teenistuskäik:

02/2020–…	Matemaatika ja statistika instituut, Tartu Ülikool, praktikumide juhendaja aines "LTMS.00.025" Statistiline andmeteadus
09/2016–…	Tartu Ülikooli Eesti Geenivaramu, spetsialist (biostatistika)
01/2016–02/2016	University of Wisconsin-Madison, Departments of Biostatistics and Medical Informatics, and Population Health Sciences, praktikant (puuduvate andmete meetodid)
05/2015–07/2015	Tartu Ülikool, Teadus- ja Innovatsioonipoliitika seire töögrupp, andmete analüüsija
11/2013–05/2014	Statoil Fuel & Retail Eesti AS, finants- ja krediidikontroller
07/2013–11/2013	Statoil Fuel & Retail Eesti AS, nooremkontroller
05/2012–06/2012	Rahandusministeerium, Kindlustuspoliitika osakonna analüütiku praktikant

Teadusorganisatsiooniline ja- administratiivne tegevus:

2019	Arvamusfestivalil aruteluteema "Geeniandmed: oht või võimalus?"
2018–…	Rahvusvahelise biomeetriaühingu "International Biometric Society" Põhja-Balti regiooni liige

Teaduspublikatsioonid:

Macdonald-Dunlop, E.; **Taba, N.**; Klarić, L.; Frkatović, A.; Walker, R.; Hayward, C.; Esko, T.; Haley, C.; Fischer, K.; Wilson, J.F.; Joshi, P.K. (2022). A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk. Aging, 623–659. DOI: 10.18632/aging.203847.

Gobeil, É.; Maltais-Payette, I.; **Taba, N.**; Brière, F.; Ghodsian, N.; Abner, E.; Bourgault, J.; Gagnon, E.; Manikpurage, H.D.; Couture, C.; Mitchell, P.L.; Mathieu, P.; Julien, F.; Corbeil, J.; Vohl, M.-C.; Thériault, S.; Esko, T.; Tchernof, A.; Arsenault, B.J. (2022). Mendelian Randomization Analysis Identifies Blood Tyrosine Levels as a Biomarker of Non-Alcoholic Fatty Liver Disease. Metabolites, 12 (5), 440. DOI: 10.3390/metabo12050440.

Pirastu, N.; McDonnell, C.; Grzeszkowiak, E. J.; Mounier, N.; Imamura, F.; Merino, J.; Day, Felix R.; Zheng, J.; **Taba, N.**; Concas, M. P.; Repetto, L.; Kentistou, K. A.; Robino, A.; Esko, T.; Joshi, P. K.; Fischer, K.; Ong, K. K.; Gaunt, T. R.; Kutalik, Z.; Perry, J. R. B. ... Wilson, J. F. (2022). Using genetic variation to disentangle the complex relationship between food intake and health outcomes. PLoS Genetics, 18 (6), e1010162. DOI: 10.1371/journal.pgen.1010162.

Schneider, S.; **Taba, N.**; Saapar, M.; Vibo, R.; Kõrv, J. (2021). Determinants of long-term health-related quality of life in young ischemic stroke patients. Journal of Stroke and Cerebrovascular Diseases, 30, 105499. DOI: 10.1016/j.jstrokecerebrovasdis.2020.105499.

Vojinovic, D.; Kalaoja, M.; Trompet, S.; Fischer, K.; Shipley, M.J.; Li, S.; Havulinna, A.S.; Perola, M.; Salomaa, V.; Yang, Q.; Sattar, N.; Jousilahti, P.; Amin, N.; Satizabal, C.L.; **Taba, N.**; Sabayan, B.; Vasan, R.S.; Ikram, M.A.; Stott, D.J.; Ala-Korpela, M. ... van Duijn, C.M. (2021). Association of Circulating Metabolites in Plasma or Serum and Risk of Stroke. Neurology, 96 (8), e1110–e1123. DOI: 10.1212/WNL.0000000000011236.

Taba, N.; Valge, H.-K.; Metspalu, A.; Esko, T.; Wilson, J.F.; Fischer, K.; Pirastu, N. (2021). Mendelian Randomization Identifies the Potential Causal Impact of Dietary Patterns on Circulating Blood Metabolites. Frontiers in Genetics, 12, ARTN 738265. DOI: 10.3389/fgene.2021.738265.

Ghodsian, N.; Abner, E.; Emdin, C.A.; Gobeil, É.; **Taba, N.**; Haas, M.E.; Perrot, N.; Manikpurage, H.D.; Gagnon, É.; Bourgault, J.; St-Amand, A.; Couture, C.; Mitchell, P.L.; Bossé, Y.; Mathieu, P.; Vohl, M.-C.; Tchernof, A.; Thériault, S.; Khera, A.V.; Esko, T. ... Arsenault, B.J. (2021). Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. Cell Reports Medicine, 2 (11), 100437. DOI: 10.1016/j.xcrm.2021.100437.

Schneider, S.; Vibo, R.; **Taba, N.**; Kõrv, J. (2020). Mortality in young adult patients with acute ischaemic stroke. Acta Neurologica Scandinavica, 141 (3), 242–249. DOI: 10.1111/ane.13217.

Kadastik-Eerme, L.; **Taba, N.**; Asser, T-; Taba, P. (2019). Incidence and Mortality of Parkinson's Disease in Estonia. Neuroepidemiology, 53 (1–2), 63–72. DOI: 10.1159/000499756.

Sarin, H.V.; **Taba, N.**; Fischer, K.; Esko, T.; Kanerva, N.; Moilanen, L.; Saltevo, J.; Joensuu, A.; Borodulin, K.; Männistö, S.; Kristiansson, K.; Perola, M. (2019). Food neophobia associates with poorer dietary quality, metabolic risk factors, and increased disease outcome risk in population-based cohorts in a metabolomics study. American Journal of Clinical Nutrition, 110 (1), 233–245. DOI: 10.1093/ajcn/nqz100.

Kadastik-Eerme L., **Taba N.**, Asser T., Taba P. (2018). The increasing prevalence of Parkinson's disease in Estonia. Acta Neurologica Scandinavica, 138 (3), 251–258. DOI: 10.1111/ane.12948.

Kadastik-Eerme L.; **Taba N.**; Asser T.; Taba P. (2017). Factors associated with motor complications in Parkinson's disease. Brain and Behavior, 7 (10), e00827–e00827. DOI: 10.1002/brb3.837.

Kadastik-Eerme L.; Muldmaa M.; Lilles S.; Rosenthal M;, **Taba N.**; Taba P. (2016). Non-motor features in Parkinson's disease: What are the most important associated factors? Parkinson's Disease, 9, 1–8. DOI: 10.1155/2016/4370674.

Juhendatud väitekirjad:

2018 Marili Zimmermanni magistritöö "Elukestusanalüüs vasakult tõkestatud andmete ning ajast sõltuva argumenttunnuse korral TÜ Eesti geenivaramu kohordi näitel" (matemaatika ja statistika õppekava)

Stipendiumid:

2021 Parim suuline ettekanne: Tartu Ülikooli Molekulaar- ja Rakubioloogia Instituudi aastakonverentsil (Tartu Ülikool, Tartu, Eesti)
2021 Eesti Teaduste Akadeemia "Teadus 3 minutiga" loengute konkursi finalist
2019 Biomeditsiini ja Biotehnoloogia doktorikooli stipendium: suuline ettekanne The Nordic Society of Human Genetics and Precision Medicine 2019 konverentsil (ESHG satelliitsümpoosium) (Göteborg, Rootsi)
2019 Biomeditsiini ja Biotehnoloogia doktorikooli stipendium: osalemine UseR!2019 kursusel (Tolouse, Prantsusmaa)
2018 Parim poster: European Mathematical Genetics Meeting (Cagliari, Itaalia)
2018 Dora Pluss lühiajalise õpirände stipendium
2018 Dora Pluss pikaajalise õpirände stipendium
2017 Sihtasutus Archimedes nutika spetsialiseerumise kasvuvaldkondade doktorandistipendium
2017 Parim poster: Workshop on Recent & Future Trends in Biostatistics (Cambridge University, Cambridge, Ühendkuningriik)
2017 Parim suuline ettekanne: Tartu Ülikooli Molekulaar- ja Rakubioloogia instituudi aastakonverentsil (Tartu Ülikool, Tartu, Eesti)
2015 Kristjan Jaagu tasemeõppe stipendium (õppeaasta 2015/2016)
2013 Parim bakalaureusetöö ettevõttemajanduse instituudis juhtimise suunal

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida*. Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand**. The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous grassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan**. Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.
41. **Sulev Ingerpuu**. Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.

42. **Veljo Kisand**. Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa**. Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa**. Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik**. Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo**. Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo**. Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots**. Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero**. Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees**. Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks**. Cholecystokinin (CCK) – induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and serotonin. Tartu, 1999, 123 p.
52. **Ebe Sild**. Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva**. Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna**. Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro**. Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane**. Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm**. Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg**. Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild**. The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu**. Studies of the TOL plasmid transcription factor XylS. Tartu, 2000, 88 p.
61. **Dina Lepik**. Modulation of viral DNA replication by tumor suppressor protein p53. Tartu, 2000, 106 p.
62. **Kai Vellak**. Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu, 2000, 122 p.

63. **Jonne Kotta**. Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu, 2000, 160 p.
64. **Georg Martin**. Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000, 139 p.
65. **Silvia Sepp**. Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira**. On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000, 96 p.
67. **Priit Zingel**. The role of planktonic ciliates in lake ecosystems. Tartu, 2001, 111 p.
68. **Tiit Teder**. Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu, 2001, 122 p.
69. **Hannes Kollist**. Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu, 2001, 80 p.
70. **Reet Marits**. Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu, 2001, 112 p.
71. **Vallo Tilgar**. Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002, 126 p.
72. **Rita Hõrak**. Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002, 108 p.
73. **Liina Eek-Piirsoo**. The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002, 74 p.
74. **Krõõt Aasamaa**. Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002, 110 p.
75. **Nele Ingerpuu**. Bryophyte diversity and vascular plants. Tartu, 2002, 112 p.
76. **Neeme Tõnisson**. Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002, 124 p.
77. **Margus Pensa**. Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003, 110 p.
78. **Asko Lõhmus**. Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003, 168 p.
79. **Viljar Jaks**. p53 – a switch in cellular circuit. Tartu, 2003, 160 p.
80. **Jaana Männik**. Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003, 140 p.
81. **Marek Sammul**. Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003, 159 p
82. **Ivar Ilves**. Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003, 89 p.
83. **Andres Männik**. Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003, 109 p.

84. **Ivika Ostonen**. Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003, 158 p.
85. **Gudrun Veldre**. Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003, 199 p.
86. **Ülo Väli**. The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004, 159 p.
87. **Aare Abroi**. The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004, 135 p.
88. **Tiina Kahre**. Cystic fibrosis in Estonia. Tartu, 2004, 116 p.
89. **Helen Orav-Kotta**. Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004, 117 p.
90. **Maarja Öpik**. Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004, 175 p.
91. **Kadri Tali**. Species structure of *Neotinea ustulata*. Tartu, 2004, 109 p.
92. **Kristiina Tambets**. Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004, 163 p.
93. **Arvi Jõers**. Regulation of p53-dependent transcription. Tartu, 2004, 103 p.
94. **Lilian Kadaja**. Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004, 103 p.
95. **Jaak Truu**. Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004, 128 p.
96. **Maire Peters**. Natural horizontal transfer of the *pheBA* operon. Tartu, 2004, 105 p.
97. **Ülo Maiväli**. Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004, 130 p.
98. **Merit Otsus**. Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004, 103 p.
99. **Mikk Heidemaa**. Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004, 167 p.
100. **Ilmar Tõnno**. The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004, 111 p.
101. **Lauri Saks**. Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004, 144 p.
102. **Siiri Rootsi**. Human Y-chromosomal variation in European populations. Tartu, 2004, 142 p.
103. **Eve Vedler**. Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover**. Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005, 126 p.
105. **Helen Udras**. Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005, 100 p.

106. **Ave Suija**. Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005, 162 p.
107. **Piret Lõhmus**. Forest lichens and their substrata in Estonia. Tartu, 2005, 162 p.
108. **Inga Lips**. Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005, 156 p.
109. **Krista Kaasik**. Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005, 121 p.
110. **Juhan Javoiš**. The effects of experience on host acceptance in ovipositing moths. Tartu, 2005, 112 p.
111. **Tiina Sedman**. Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005, 103 p.
112. **Ruth Aguraiuja**. Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005, 112 p.
113. **Riho Teras**. Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005, 106 p.
114. **Mait Metspalu**. Through the course of prehistory in India: tracing the mtDNA trail. Tartu, 2005, 138 p.
115. **Elin Lõhmussaar**. The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006, 124 p.
116. **Priit Kupper**. Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006, 126 p.
117. **Heili Ilves**. Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006, 120 p.
118. **Silja Kuusk**. Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006, 126 p.
119. **Kersti Püssa**. Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006, 90 p.
120. **Lea Tummeleht**. Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006, 94 p.
121. **Toomas Esperk**. Larval instar as a key element of insect growth schedules. Tartu, 2006, 186 p.
122. **Harri Valdmann**. Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.
123. **Priit Jõers**. Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli**. Gata3 and Gata2 in inner ear development. Tartu, 2007, 123 p.
125. **Kai Rünk**. Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007, 143 p.

126. **Aveliina Helm**. Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007, 89 p.
127. **Leho Tedersoo**. Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007, 233 p.
128. **Marko Mägi**. The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007, 135 p.
129. **Valeria Lulla**. Replication strategies and applications of Semliki Forest virus. Tartu, 2007, 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007, 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007, 171 p.
132. **Tatjana Krama**. Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007, 112 p.
133. **Signe Saumaa**. The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007, 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007, 96 p.
135. **Priit Kilgas**. Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007, 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007, 95 p.
137. **Kärt Padari**. Protein transduction mechanisms of transportans. Tartu, 2008, 128 p.
138. **Siiri-Lii Sandre**. Selective forces on larval colouration in a moth. Tartu, 2008, 125 p.
139. **Ülle Jõgar**. Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008, 99 p.
140. **Lauri Laanisto**. Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008, 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008, 105 p.
142. **Birgot Paavel**. Bio-optical properties of turbid lakes. Tartu, 2008, 175 p.
143. **Kaire Torn**. Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg**. Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd**. Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag**. Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.

147. **Ulvi Karu**. Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm**. Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks**. Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu**. Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš**. Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko**. Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast**. Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats**. Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova**. The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik**. Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2009, 112 p.
158. **Kaja Kiiver**. Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja**. Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast**. Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak**. Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.
162. **Triinu Remmel**. Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe**. Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe**. Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa**. Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.

166. **Pille Säälik**. The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil**. Ribosome assembly factors in *Escherichia coli*. Tartu, 2009, 147 p.
168. **Lea Hallik**. Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark**. Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap**. Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan**. Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe**. Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi**. Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson**. Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts**. Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis**. Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov**. Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster**. Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap**. Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar**. Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül**. Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.
182. **Arto Pulk**. Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü**. Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla**. Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey**. The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.

186. **Katrin Kepp**. Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.
187. **Virve Sõber**. The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro**. The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold**. Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert**. Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu**. Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik**. ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida*. Tartu, 2010, 118 p.
193. **Siim Sõber**. Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper**. Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak**. Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo**. Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel**. Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus**. Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius**. Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv**. Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk**. Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.
202. **Arno Põllumäe**. Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht**. Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus**. Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.

206. **Kessy Abarenkov**. PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.
207. **Marina Grigorova**. Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar**. The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild**. Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag**. Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.
212. **Aleksei Lulla**. Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage. Tartu, 2012, 143 p.
213. **Mari Järve**. Different genetic perspectives on human history in Europe and the Caucasus: the stories told by uniparental and autosomal markers. Tartu, 2012, 119 p.
214. **Ott Scheler**. The application of tmRNA as a marker molecule in bacterial diagnostics using microarray and biosensor technology. Tartu, 2012, 93 p.
215. **Anna Balikova**. Studies on the functions of tumor-associated mucin-like leukosialin (CD43) in human cancer cells. Tartu, 2012, 129 p.
216. **Triinu Kõressaar**. Improvement of PCR primer design for detection of prokaryotic species. Tartu, 2012, 83 p.
217. **Tuul Sepp**. Hematological health state indices of greenfinches: sources of individual variation and responses to immune system manipulation. Tartu, 2012, 117 p.
218. **Rya Ero**. Modifier view of the bacterial ribosome. Tartu, 2012, 146 p.
219. **Mohammad Bahram**. Biogeography of ectomycorrhizal fungi across different spatial scales. Tartu, 2012, 165 p.
220. **Annely Lorents**. Overcoming the plasma membrane barrier: uptake of amphipathic cell-penetrating peptides induces influx of calcium ions and downstream responses. Tartu, 2012, 113 p.
221. **Katrin Männik**. Exploring the genomics of cognitive impairment: whole-genome SNP genotyping experience in Estonian patients and general population. Tartu, 2012, 171 p.
222. **Marko Prous**. Taxonomy and phylogeny of the sawfly genus *Empria* (Hymenoptera, Tenthredinidae). Tartu, 2012, 192 p.
223. **Triinu Visnapuu**. Levansucrases encoded in the genome of *Pseudomonas syringae* pv. tomato DC3000: heterologous expression, biochemical characterization, mutational analysis and spectrum of polymerization products. Tartu, 2012, 160 p.
224. **Nele Tamberg**. Studies on Semliki Forest virus replication and pathogenesis. Tartu, 2012, 109 p.

225. **Tõnu Esko**. Novel applications of SNP array data in the analysis of the genetic structure of Europeans and in genetic association studies. Tartu, 2012, 149 p.
226. **Timo Arula**. Ecology of early life-history stages of herring *Clupea harengus membras* in the northeastern Baltic Sea. Tartu, 2012, 143 p.
227. **Inga Hiiesalu**. Belowground plant diversity and coexistence patterns in grassland ecosystems. Tartu, 2012, 130 p.
228. **Kadri Koorem**. The influence of abiotic and biotic factors on small-scale plant community patterns and regeneration in boreonemoral forest. Tartu, 2012, 114 p.
229. **Liis Andresen**. Regulation of virulence in plant-pathogenic pectobacteria. Tartu, 2012, 122 p.
230. **Kaupo Kohv**. The direct and indirect effects of management on boreal forest structure and field layer vegetation. Tartu, 2012, 124 p.
231. **Mart Jüssi**. Living on an edge: landlocked seals in changing climate. Tartu, 2012, 114 p.
232. **Riina Klais**. Phytoplankton trends in the Baltic Sea. Tartu, 2012, 136 p.
233. **Rauno Veeroja**. Effects of winter weather, population density and timing of reproduction on life-history traits and population dynamics of moose (*Alces alces*) in Estonia. Tartu, 2012, 92 p.
234. **Marju Keis**. Brown bear (*Ursus arctos*) phylogeography in northern Eurasia. Tartu, 2013, 142 p.
235. **Sergei Põlme**. Biogeography and ecology of *alnus*- associated ectomycorrhizal fungi – from regional to global scale. Tartu, 2013, 90 p.
236. **Liis Uusküla**. Placental gene expression in normal and complicated pregnancy. Tartu, 2013, 173 p.
237. **Marko Lõoke**. Studies on DNA replication initiation in *Saccharomyces cerevisiae*. Tartu, 2013, 112 p.
238. **Anne Aan**. Light- and nitrogen-use and biomass allocation along productivity gradients in multilayer plant communities. Tartu, 2013, 127 p.
239. **Heidi Tamm**. Comprehending phylogenetic diversity – case studies in three groups of ascomycetes. Tartu, 2013, 136 p.
240. **Liina Kangur**. High-Pressure Spectroscopy Study of Chromophore-Binding Hydrogen Bonds in Light-Harvesting Complexes of Photosynthetic Bacteria. Tartu, 2013, 150 p.
241. **Margus Leppik**. Substrate specificity of the multisite specific pseudouridine synthase RluD. Tartu, 2013, 111 p.
242. **Lauris Kaplinski**. The application of oligonucleotide hybridization model for PCR and microarray optimization. Tartu, 2013, 103 p.
243. **Merli Pärnoja**. Patterns of macrophyte distribution and productivity in coastal ecosystems: effect of abiotic and biotic forcing. Tartu, 2013, 155 p.
244. **Tõnu Margus**. Distribution and phylogeny of the bacterial translational GTPases and the Mqsr/YgiT regulatory system. Tartu, 2013, 126 p.
245. **Pille Mänd**. Light use capacity and carbon and nitrogen budget of plants: remote assessment and physiological determinants. Tartu, 2013, 128 p.

246. **Mario Plaas**. Animal model of Wolfram Syndrome in mice: behavioural, biochemical and psychopharmacological characterization. Tartu, 2013, 144 p.
247. **Georgi Hudjašov**. Maps of mitochondrial DNA, Y-chromosome and tyrosinase variation in Eurasian and Oceanian populations. Tartu, 2013, 115 p.
248. **Mari Lepik**. Plasticity to light in herbaceous plants and its importance for community structure and diversity. Tartu, 2013, 102 p.
249. **Ede Leppik**. Diversity of lichens in semi-natural habitats of Estonia. Tartu, 2013, 151 p.
250. **Ülle Saks**. Arbuscular mycorrhizal fungal diversity patterns in boreonemoral forest ecosystems. Tartu, 2013, 151 p.
251. **Eneli Oitmaa**. Development of arrayed primer extension microarray assays for molecular diagnostic applications. Tartu, 2013, 147 p.
252. **Jekaterina Jutkina**. The horizontal gene pool for aromatics degradation: bacterial catabolic plasmids of the Baltic Sea aquatic system. Tartu, 2013, 121 p.
253. **Helen Vellau**. Reaction norms for size and age at maturity in insects: rules and exceptions. Tartu, 2014, 132 p.
254. **Randel Kreitsberg**. Using biomarkers in assessment of environmental contamination in fish – new perspectives. Tartu, 2014, 107 p.
255. **Krista Takkis**. Changes in plant species richness and population performance in response to habitat loss and fragmentation.Tartu, 2014, 141 p.
256. **Liina Nagirnaja**. Global and fine-scale genetic determinants of recurrent pregnancy loss. Tartu, 2014, 211 p.
257. **Triin Triisberg**. Factors influencing the re-vegetation of abandoned extracted peatlands in Estonia. Tartu, 2014, 133 p.
258. **Villu Soon**. A phylogenetic revision of the *Chrysis ignita* species group (Hymenoptera: Chrysididae) with emphasis on the northern European fauna. Tartu, 2014, 211 p.
259. **Andrei Nikonov**. RNA-Dependent RNA Polymerase Activity as a Basis for the Detection of Positive-Strand RNA Viruses by Vertebrate Host Cells. Tartu, 2014, 207 p.
260. **Eele Õunapuu-Pikas**. Spatio-temporal variability of leaf hydraulic conductance in woody plants: ecophysiological consequences. Tartu, 2014, 135 p.
261. **Marju Männiste**. Physiological ecology of greenfinches: information content of feathers in relation to immune function and behavior. Tartu, 2014, 121 p.
262. **Katre Kets**. Effects of elevated concentrations of CO_2 and O_3 on leaf photosynthetic parameters in *Populus tremuloides*: diurnal, seasonal and interannual patterns. Tartu, 2014, 115 p.
263. **Külli Lokko**. Seasonal and spatial variability of zoopsammon communities in relation to environmental parameters. Tartu, 2014, 129 p.
264. **Olga Žilina**. Chromosomal microarray analysis as diagnostic tool: Estonian experience. Tartu, 2014, 152 p.

265. **Kertu Lõhmus**. Colonisation ecology of forest-dwelling vascular plants and the conservation value of rural manor parks. Tartu, 2014, 111 p.
266. **Anu Aun**. Mitochondria as integral modulators of cellular signaling. Tartu, 2014, 167 p.
267. **Chandana Basu Mallick**. Genetics of adaptive traits and gender-specific demographic processes in South Asian populations. Tartu, 2014, 160 p.
268. **Riin Tamme**. The relationship between small-scale environmental heterogeneity and plant species diversity. Tartu, 2014, 130 p.
269. **Liina Remm**. Impacts of forest drainage on biodiversity and habitat quality: implications for sustainable management and conservation. Tartu, 2015, 126 p.
270. **Tiina Talve**. Genetic diversity and taxonomy within the genus *Rhinanthus.* Tartu, 2015, 106 p.
271. **Mehis Rohtla**. Otolith sclerochronological studies on migrations, spawning habitat preferences and age of freshwater fishes inhabiting the Baltic Sea. Tartu, 2015, 137 p.
272. **Alexey Reshchikov**. The world fauna of the genus *Lathrolestes* (Hymenoptera, Ichneumonidae). Tartu, 2015, 247 p.
273. **Martin Pook**. Studies on artificial and extracellular matrix protein-rich surfaces as regulators of cell growth and differentiation. Tartu, 2015, 142 p.
274. **Mai Kukumägi**. Factors affecting soil respiration and its components in silver birch and Norway spruce stands. Tartu, 2015, 155 p.
275. **Helen Karu**. Development of ecosystems under human activity in the North-East Estonian industrial region: forests on post-mining sites and bogs. Tartu, 2015, 152 p.
276. **Hedi Peterson**. Exploiting high-throughput data for establishing relationships between genes. Tartu, 2015, 186 p.
277. **Priit Adler**. Analysis and visualisation of large scale microarray data, Tartu, 2015, 126 p.
278. **Aigar Niglas**. Effects of environmental factors on gas exchange in deciduous trees: focus on photosynthetic water-use efficiency. Tartu, 2015, 152 p.
279. **Silja Laht**. Classification and identification of conopeptides using profile hidden Markov models and position-specific scoring matrices. Tartu, 2015, 100 p.
280. **Martin Kesler**. Biological characteristics and restoration of Atlantic salmon *Salmo salar* populations in the Rivers of Northern Estonia. Tartu, 2015, 97 p.
281. **Pratyush Kumar Das**. Biochemical perspective on alphaviral nonstructural protein 2: a tale from multiple domains to enzymatic profiling. Tartu, 2015, 205 p
282. **Priit Palta**. Computational methods for DNA copy number detection. Tartu, 2015, 130 p.
283. **Julia Sidorenko**. Combating DNA damage and maintenance of genome integrity in pseudomonads. Tartu, 2015, 174 p.

284. **Anastasiia Kovtun-Kante**. Charophytes of Estonian inland and coastal waters: distribution and environmental preferences. Tartu, 2015, 97 p.
285. **Ly Lindman**. The ecology of protected butterfly species in Estonia. Tartu, 2015, 171 p.
286. **Jaanis Lodjak**. Association of Insulin-like Growth Factor I and Corticosterone with Nestling Growth and Fledging Success in Wild Passerines. Tartu, 2016, 113 p.
287. **Ann Kraut**. Conservation of Wood-Inhabiting Biodiversity – Semi-Natural Forests as an Opportunity. Tartu, 2016, 141 p.
288. **Tiit Örd**. Functions and regulation of the mammalian pseudokinase TRIB3. Tartu, 2016, 182. p.
289. **Kairi Käiro**. Biological Quality According to Macroinvertebrates in Streams of Estonia (Baltic Ecoregion of Europe): Effects of Human-induced Hydromorphological Changes. Tartu, 2016, 126 p.
290. **Leidi Laurimaa**. *Echinococcus multilocularis* and other zoonotic parasites in Estonian canids. Tartu, 2016, 144 p.
291. **Helerin Margus**. Characterization of cell-penetrating peptide/nucleic acid nanocomplexes and their cell-entry mechanisms. Tartu, 2016, 173 p.
292. **Kadri Runnel**. Fungal targets and tools for forest conservation. Tartu, 2016, 157 p.
293. **Urmo Võsa**. MicroRNAs in disease and health: aberrant regulation in lung cancer and association with genomic variation. Tartu, 2016, 163 p.
294. **Kristina Mäemets-Allas**. Studies on cell growth promoting AKT signaling pathway – a promising anti-cancer drug target. Tartu, 2016, 146 p.
295. **Janeli Viil**. Studies on cellular and molecular mechanisms that drive normal and regenerative processes in the liver and pathological processes in Dupuytren's contracture. Tartu, 2016, 175 p.
296. **Ene Kook**. Genetic diversity and evolution of *Pulmonaria angustifolia* L. and *Myosotis laxa sensu lato* (Boraginaceae). Tartu, 2016, 106 p.
297. **Kadri Peil**. RNA polymerase II-dependent transcription elongation in *Saccharomyces cerevisiae*. Tartu, 2016, 113 p.
298. **Katrin Ruisu**. The role of RIC8A in mouse development and its function in cell-matrix adhesion and actin cytoskeletal organisation. Tartu, 2016, 129 p.
299. **Janely Pae**. Translocation of cell-penetrating peptides across biological membranes and interactions with plasma membrane constituents. Tartu, 2016, 126 p.
300. **Argo Ronk**. Plant diversity patterns across Europe: observed and dark diversity. Tartu, 2016, 153 p.
301. **Kristiina Mark**. Diversification and species delimitation of lichenized fungi in selected groups of the family Parmeliaceae (Ascomycota). Tartu, 2016, 181 p.
302. **Jaak-Albert Metsoja**. Vegetation dynamics in floodplain meadows: influence of mowing and sediment application. Tartu, 2016, 140 p.

303. **Hedvig Tamman**. The GraTA toxin-antitoxin system of *Pseudomonas putida*: regulation and role in stress tolerance. Tartu, 2016, 154 p.
304. **Kadri Pärtel**. Application of ultrastructural and molecular data in the taxonomy of helotialean fungi. Tartu, 2016, 183 p.
305. **Maris Hindrikson**. Grey wolf (*Canis lupus*) populations in Estonia and Europe: genetic diversity, population structure and -processes, and hybridization between wolves and dogs. Tartu, 2016, 121 p.
306. **Polina Degtjarenko**. Impacts of alkaline dust pollution on biodiversity of plants and lichens: from communities to genetic diversity. Tartu, 2016, 126 p.
307. **Liina Pajusalu**. The effect of CO_2 enrichment on net photosynthesis of macrophytes in a brackish water environment. Tartu, 2016, 126 p.
308. **Stoyan Tankov**. Random walks in the stringent response. Tartu, 2016, 94 p.
309. **Liis Leitsalu**. Communicating genomic research results to population-based biobank participants. Tartu, 2016, 158 p.
310. **Richard Meitern**. Redox physiology of wild birds: validation and application of techniques for detecting oxidative stress. Tartu, 2016, 134 p.
311. **Kaie Lokk**. Comparative genome-wide DNA methylation studies of healthy human tissues and non-small cell lung cancer tissue. Tartu, 2016, 127 p.
312. **Mihhail Kurašin**. Processivity of cellulases and chitinases. Tartu, 2017, 132 p.
313. **Carmen Tali**. Scavenger receptors as a target for nucleic acid delivery with peptide vectors. Tartu, 2017, 155 p.
314. **Katarina Oganjan**. Distribution, feeding and habitat of benthic suspension feeders in a shallow coastal sea. Tartu, 2017, 132 p.
315. **Taavi Paal**. Immigration limitation of forest plants into wooded landscape corridors. Tartu, 2017, 145 p.
316. **Kadri Õunap**. The Williams-Beuren syndrome chromosome region protein WBSCR22 is a ribosome biogenesis factor. Tartu, 2017, 135 p.
317. **Riin Tamm**. In-depth analysis of factors affecting variability in thiopurine methyltransferase activity. Tartu, 2017, 170 p.
318. **Keiu Kask**. The role of RIC8A in the development and regulation of mouse nervous system. Tartu, 2017, 184 p.
319. **Tiia Möller**. Mapping and modelling of the spatial distribution of benthic macrovegetation in the NE Baltic Sea with a special focus on the eelgrass *Zostera marina* Linnaeus, 1753. Tartu, 2017, 162 p.
320. **Silva Kasela**. Genetic regulation of gene expression: detection of tissue- and cell type-specific effects. Tartu, 2017, 150 p.
321. **Karmen Süld**. Food habits, parasites and space use of the raccoon dog *Nyctereutes procyonoides*: the role of an alien species as a predator and vector of zoonotic diseases in Estonia. Tartu, 2017, p.
322. **Ragne Oja**. Consequences of supplementary feeding of wild boar – concern for ground-nesting birds and endoparasite infection. Tartu, 2017, 141 p.
323. **Riin Kont**. The acquisition of cellulose chain by a processive cellobiohydrolase. Tartu, 2017, 117 p.

324. **Liis Kasari**. Plant diversity of semi-natural grasslands: drivers, current status and conservation challenges. Tartu, 2017, 141 p.
325. **Sirgi Saar**. Belowground interactions: the roles of plant genetic relatedness, root exudation and soil legacies. Tartu, 2017, 113 p.
326. **Sten Anslan**. Molecular identification of Collembola and their fungal associates. Tartu, 2017, 125 p.
327. **Imre Taal**. Causes of variation in littoral fish communities of the Eastern Baltic Sea: from community structure to individual life histories. Tartu, 2017, 118 p.
328. **Jürgen Jalak**. Dissecting the Mechanism of Enzymatic Degradation of Cellulose Using Low Molecular Weight Model Substrates. Tartu, 2017, 137 p.
329. **Kairi Kiik**. Reproduction and behaviour of the endangered European mink (*Mustela lutreola*) in captivity. Tartu, 2018, 112 p.
330. **Ivan Kuprijanov**. Habitat use and trophic interactions of native and invasive predatory macroinvertebrates in the northern Baltic Sea. Tartu, 2018, 117 p.
331. **Hendrik Meister**. Evolutionary ecology of insect growth: from geographic patterns to biochemical trade-offs. Tartu, 2018, 147 p.
332. **Ilja Gaidutšik**. Irc3 is a mitochondrial branch migration enzyme in *Saccharomyces cerevisiae.* Tartu, 2018, 161 p.
333. **Lena Neuenkamp**. The dynamics of plant and arbuscular mycorrhizal fungal communities in grasslands under changing land use. Tartu, 2018, 241 p.
334. **Laura Kasak**. Genome structural variation modulating the placenta and pregnancy maintenance. Tartu, 2018, 181 p.
335. **Kersti Riibak**. Importance of dispersal limitation in determining dark diversity of plants across spatial scales. Tartu, 2018, 133 p.
336. **Liina Saar**. Dynamics of grassland plant diversity in changing landscapes. Tartu, 2018, 206 p.
337. **Hanna Ainelo**. Fis regulates *Pseudomonas putida* biofilm formation by controlling the expression of *lapA*. Tartu, 2018, 143 p.
338. **Natalia Pervjakova**. Genomic imprinting in complex traits. Tartu, 2018, 176 p.
339. **Andrio Lahesaare**. The role of global regulator Fis in regulating the expression of *lapF* and the hydrophobicity of soil bacterium *Pseudomonas putida*. Tartu, 2018, 124 p.
340. **Märt Roosaare**. *K*-mer based methods for the identification of bacteria and plasmids. Tartu, 2018, 117 p.
341. **Maria Abakumova**. The relationship between competitive behaviour and the frequency and identity of neighbours in temperate grassland plants. Tartu, 2018, 104 p.
342. **Margus Vilbas**. Biotic interactions affecting habitat use of myrmecophilous butterflies in Northern Europe. Tartu, 2018, 142 p.

343. **Liina Kinkar**. Global patterns of genetic diversity and phylogeography of *Echinococcus granulosus* sensu stricto – a tapeworm species of significant public health concern. Tartu, 2018, 147 p.
344. **Teivi Laurimäe**. Taxonomy and genetic diversity of zoonotic tapeworms in the species complex of *Echinococcus granulosus* sensu lato. Tartu, 2018, 143 p.
345. **Tatjana Jatsenko**. Role of translesion DNA polymerases in mutagenesis and DNA damage tolerance in Pseudomonads. Tartu, 2018, 216 p.
346. **Katrin Viigand**. Utilization of α-glucosidic sugars by *Ogataea* (*Hansenula*) *polymorpha*. Tartu, 2018, 148 p.
347. **Andres Ainelo**. Physiological effects of the *Pseudomonas putida* toxin grat. Tartu, 2018, 146 p.
348. **Killu Timm**. Effects of two genes (DRD4 and SERT) on great tit (*Parus major*) behaviour and reproductive traits. Tartu, 2018, 117 p.
349. **Petr Kohout**. Ecology of ericoid mycorrhizal fungi. Tartu, 2018, 184 p.
350. **Gristin Rohula-Okunev**. Effects of endogenous and environmental factors on night-time water flux in deciduous woody tree species. Tartu, 2018, 184 p.
351. **Jane Oja**. Temporal and spatial patterns of orchid mycorrhizal fungi in forest and grassland ecosystems. Tartu, 2018, 102 p.
352. **Janek Urvik**. Multidimensionality of aging in a long-lived seabird. Tartu, 2018, 135 p.
353. **Lisanna Schmidt**. Phenotypic and genetic differentiation in the hybridizing species pair *Carex flava* and *C. viridula* in geographically different regions. Tartu, 2018, 133 p.
354. **Monika Karmin**. Perspectives from human Y chromosome – phylogeny, population dynamics and founder events. Tartu, 2018, 168 p.
355. **Maris Alver**. Value of genomics for atherosclerotic cardiovascular disease risk prediction. Tartu, 2019, 148 p.
356. **Lehti Saag**. The prehistory of Estonia from a genetic perspective: new insights from ancient DNA. Tartu, 2019, 171 p.
357. **Mari-Liis Viljur**. Local and landscape effects on butterfly assemblages in managed forests. Tartu, 2019, 115 p.
358. **Ivan Kisly**. The pleiotropic functions of ribosomal proteins eL19 and eL24 in the budding yeast ribosome. Tartu, 2019, 170 p.
359. **Mikk Puustusmaa**. On the origin of papillomavirus proteins. Tartu, 2019, 152 p.
360. **Anneliis Peterson**. Benthic biodiversity in the north-eastern Baltic Sea: mapping methods, spatial patterns, and relations to environmental gradients. Tartu, 2019, 159 p.
361. **Erwan Pennarun**. Meandering along the mtDNA phylogeny; causerie and digression about what it can tell us about human migrations. Tartu, 2019, 162 p.

362. **Karin Ernits**. Levansucrase Lsc3 and endo-levanase BT1760: characterization and application for the synthesis of novel prebiotics. Tartu, 2019, 217 p.
363. **Sille Holm**. Comparative ecology of geometrid moths: in search of contrasts between a temperate and a tropical forest. Tartu, 2019, 135 p.
364. **Anne-Mai Ilumäe**. Genetic history of the Uralic-speaking peoples as seen through the paternal haplogroup N and autosomal variation of northern Eurasians. Tartu, 2019, 172 p.
365. **Anu Lepik**. Plant competitive behaviour: relationships with functional traits and soil processes. Tartu, 2019, 152 p.
366. **Kunter Tätte**. Towards an integrated view of escape decisions in birds under variable levels of predation risk. Tartu, 2020, 172 p.
367. **Kaarin Parts**. The impact of climate change on fine roots and root-associated microbial communities in birch and spruce forests. Tartu, 2020, 143 p.
368. **Viktorija Kukuškina**. Understanding the mechanisms of endometrial receptivity through integration of 'omics' data layers. Tartu, 2020, 169 p.
369. **Martti Vasar**. Developing a bioinformatics pipeline gDAT to analyse arbuscular mycorrhizal fungal communities using sequence data from different marker regions. Tartu, 2020, 193 p.
370. **Ott Kangur**. Nocturnal water relations and predawn water potential disequilibrium in temperate deciduous tree species. Tartu, 2020, 126 p.
371. **Helen Post**. Overview of the phylogeny and phylogeography of the Y-chromosomal haplogroup N in northern Eurasia and case studies of two linguistically exceptional populations of Europe – Hungarians and Kalmyks. Tartu, 2020, 143 p.
372. **Kristi Krebs**. Exploring the genetics of adverse events in pharmacotherapy using Biobanks and Electronic Health Records. Tartu, 2020, 151 p.
373. **Kärt Ukkivi**. Mutagenic effect of transcription and transcription-coupled repair factors in *Pseudomonas putida.* Tartu, 2020, 154 p.
374. **Elin Soomets**. Focal species in wetland restoration. Tartu, 2020, 137 p.
375. **Kadi Tilk**. Signals and responses of ColRS two-component system in *Pseudomonas putida.* Tartu, 2020, 133 p.
376. **Indrek Teino**. Studies on aryl hydrocarbon receptor in the mouse granulosa cell model. Tartu, 2020, 139 p.
377. **Maarja Vaikre**. The impact of forest drainage on macroinvertebrates and amphibians in small waterbodies and opportunities for cost-effective mitigation. Tartu, 2020, 132 p.
378. **Siim-Kaarel Sepp**. Soil eukaryotic community responses to land use and host identity. Tartu, 2020, 222 p.
379. **Eveli Otsing**. Tree species effects on fungal richness and community structure. Tartu, 2020, 152 p.
380. **Mari Pent**. Bacterial communities associated with fungal fruitbodies. Tartu, 2020, 144 p.

381. **Einar Kärgenberg**. Movement patterns of lithophilous migratory fish in free-flowing and fragmented rivers. Tartu, 2020, 167 p.
382. **Antti Matvere**. The studies on aryl hydrocarbon receptor in murine granulosa cells and human embryonic stem cells. Tartu, 2021, 163 p.
383. **Jhonny Capichoni Massante**. Phylogenetic structure of plant communities along environmental gradients: a macroecological and evolutionary approach. Tartu, 2021, 144 p.
384. **Ajai Kumar Pathak.** Delineating genetic ancestries of people of the Indus Valley, Parsis, Indian Jews and Tharu tribe. Tartu, 2021, 197 p.
385. **Tanel Vahter.** Arbuscular mycorrhizal fungal biodiversity for sustainable agroecosystems. Tartu, 2021, 191 p.
386. **Burak Yelmen.** Characterization of ancient Eurasian influences within modern human genomes. Tartu, 2021, 134 p.
387. **Linda Ongaro.** A genomic portrait of American populations. Tartu, 2021, 182 p.
388. **Kairi Raime.** The identification of plant DNA in metagenomic samples. Tartu, 2021, 108 p.
389. **Heli Einberg.** Non-linear and non-stationary relationships in the pelagic ecosystem of the Gulf of Riga (Baltic Sea). Tartu, 2021, 119 p.
390. **Mickaël Mathieu Pihain.** The evolutionary effect of phylogenetic neighbourhoods of trees on their resistance to herbivores and climatic stress. Tartu, 2022, 145 p.
391. **Annika Joy Meitern.** Impact of potassium ion content of xylem sap and of light conditions on the hydraulic properties of trees. Tartu, 2022, 132 p.
392. **Elise Joonas.** Evaluation of metal contaminant hazard on microalgae with environmentally relevant testing strategies. Tartu, 2022, 118 p.
393. **Kreete Lüll.** Investigating the relationships between human microbiome, host factors and female health. Tartu, 2022, 141 p.
394. **Triin Kaasiku.** A wader perspective to Boreal Baltic coastal grasslands: from habitat availability to breeding site selection and nest survival. Tartu, 2022, 141 p.
395. **Meeli Alber.** Impact of elevated atmospheric humidity on the structure of the water transport pathway in deciduous trees. Tartu, 2022, 170 p.
396. **Ludovica Molinaro.** Ancestry deconvolution of Estonian, European and Worldwide genomic layers: a human population genomics excavation. Tartu, 2022, 138 p.
397. **Tina Saupe.** The genetic history of the Mediterranean before the common era: a focus on the Italian Peninsula. Tartu, 2022, 165 p.
398. **Mari-Ann Lind.** Internal constraints on energy processing and their consequences: an integrative study of behaviour, ornaments and digestive health in greenfinches. Tartu, 2022, 137 p.
399. **Markus Valge.** Testing the predictions of life history theory on anthropometric data. Tartu, 2022, 171 p.
400. **Ants Tull.** Domesticated and wild mammals as reservoirs for zoonotic helminth parasites in Estonia. Tartu, 2022, 152 p.

401. **Saleh Rahimlouye Barabi.** Investigation of diazotrophic bacteria association with plants. Tartu, 2022, 137 p.
402. **Farzad Aslani.** Towards revealing the biogeography of belowground diversity. Tartu, 2022, 124 p.