

Mendelian randomization identifies the potential causal impact of dietary patterns on circulating blood metabolites.

Nele Taba^{1,2}, Hanna-Kristel Valge³, Andres Metspalu^{1,2}, Tõnu Esko^{1,4}, James F. Wilson^{5,6}, Krista Fischer^{1,7}, Nicola Pirastu⁵

¹ Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Riia 23b, 51010, Estonia

² Institute of Molecular and Cell Biology, University of Tartu, Tartu, Riia 23, 51010, Estonia

³ Faculty of Medicine, University of Tartu, Tartu, Ravila 19, 50411, Estonia

⁴ Program in Medical and Population Genetics, Broad Institute, 415 Main St, Cambridge, MA 02142, United States

⁵ Centre for Global Health Research, Usher Institute, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, Scotland

⁶ MRC Human Genetics Unit, Institute of Genetic and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, Scotland

⁷ Institute of Mathematics and Statistics, University of Tartu, Tartu, Narva mnt 18, 51009, Estonia

Correspondence: nele.tab@ut.ee ; nicola.pirastu@ed.ac.uk

Abstract: Nutrition plays an important role in the development and progress of several health conditions, but the exact mechanism is often still unclear. Blood metabolites are likely candidates to be mediating these relationships, as their levels are strongly dependent on the frequency of consumption of several foods/drinks. Understanding the causal effect of food on metabolites is thus of extreme importance. To establish these effects we utilized Two-sample Mendelian randomization using the genetic variants associated with dietary traits as instrumental variables. The estimates of single-nucleotide polymorphisms' effects on exposures were obtained from a recent genome-wide association study (GWAS) of 25 individual and 15 principal-component dietary traits, whereas the ones for outcomes were obtained from a GWAS of 123 blood metabolites measured by nuclear magnetic resonance spectroscopy. We identified 417 potentially causal links between food and metabolites, replicating previous findings, such as the association between increased oily fish consumption and higher DHA, and highlighting several novel associations. Most of the associations were related to very-low-density, intermediate-density (IDL) and low-density lipoproteins (LDL). For example, we found that constituents of IDL particles and large LDL particles were raised by coffee and alcohol while lowered by an overall healthier diet and fruit consumption. Our results represent one of the first examples of the estimates of long-term causal effects of diet on metabolites and start bridging the gap in the mechanistic understanding linking food consumption to its health consequences.

Keywords: nutrition, metabolomics, two-sample mendelian randomization

Introduction

Nutrition plays an important role in the development and progress of several diseases, such as obesity [1], type II diabetes (T2D) [2], cardiovascular diseases (CVD) [3–5] and cancer [6,7]. These in turn create a high burden for individuals, society, the economy and health-care and thus prevention is of great importance. In many cases the mechanism by which food consumption acts on health is still unclear. Blood metabolites are promising candidates for filling this gap. Metabolites have been shown to be important in the onset of a wide range of diseases such as type II diabetes [8,9], incident cardiovascular events [10,11], dementia [12,13] and colorectal cancer [14,15], and indicative of mortality [16], and are thus likely mediators for at least some of the food-health relationships.

In recent years, the progress in quantifying metabolites has allowed investigation of the relationship between food and blood or urine metabolite levels. Studies have shown associations between metabolomic profile and intake of fruit and vegetables [17], coffee [14], alcohol [18] and a wide range of dietary patterns. A more detailed overview of the current state of the field is summarized by Guasch-Ferré, Bhupathiraju and Hu [19] and Brennan and Hu [20]. Most studies in this field are observational and are thus limited by the typical biases which affect nutritional epidemiology (i.e. reporting bias, strong correlation between the studied variables etc.) and are therefore unfit to detect causal relationships. Moreover, even in feeding studies conducted under very strict and controlled conditions, the effects could be measured only on a short-term basis and on limited sample size, which in turn limits statistical power.

Nevertheless, the feeding studies provide convincing evidence that dietary intake has causal effects on the metabolic profile, highlighting the potential to assess dietary intake via investigation of metabolomic profiles. Such studies have been applied to: the percentage of dietary intake coming from carbohydrates/fat/protein [21], the effects of low-glycaemic index diet [22] and orange juice on proline betaine [23]. Change in biomarkers can be caused by consumption or non-consumption of various dietary items and therefore variations in blood metabolites are likely candidates to be mediating the effects of food on health. Thus, detecting causal relationships between dietary choices and biomarkers might reveal more insight into the mechanism by which food affects health.

A possible alternative to approaches of observational- and feeding studies is the method of Mendelian Randomization (MR) [24]. MR exploits the natural randomization of the alleles associated with exposure in the population to measure the long term effects of the trait on the outcome of interest. MR assumes that people with the allele increasing the exposure will be exposed to this difference for their whole life and thus by comparing the effect of the allele on the exposure and the outcome it is possible to derive the effect of the exposure on the outcome. MR relies on the results coming from genome-wide association studies (GWAS) which are generally publicly available and does not require the direct involvement of the participants. It is thus extremely cost-effective and it is possible to apply it to contexts where randomized controlled trials would be unethical (for example alcohol consumption). If carefully conducted, MR is exempt from the biases that are typical of observational studies. MR has been successfully used in many

different contexts including nutritional epidemiology in the cases, for example, of milk [25,26], alcohol [27,28], and coffee consumption [29,30]. Nevertheless, there are no studies using MR on a broader range of dietary items due to the lack of single nucleotide polymorphisms (SNP) strongly associated with food consumption to be used as instrumental variables.

By virtue of the availability of the data from UK Biobank, we have recently been able to broaden the number of foods for which genetic instruments are available, identifying several causal food-health relationships [31]. We thus decided to use MR to investigate the causal effect of 40 foods/dietary patterns on 123 blood metabolites measured by nuclear magnetic resonance spectroscopy (NMR) available from a previous large GWAS by Kettunen *et al.* [32]. We detected 417 potentially causal links between food and metabolites, replicating previous findings and bringing novel insights, and discuss how these may be mediating the effect of food on health.

Materials and Methods

To infer the causal relationships between food and metabolites we used Two-sample MR. In contrast to conventional MR, which would require instrument, exposure and outcome to be available as individual-level data for the same samples, the same analysis can be performed using summary statistics from genome-wide association studies. In this case, instead of directly estimating the effect of the SNP on the exposure and on the outcome, the parameter estimates from previous association studies for the two variables are used. MR can thus be performed, even if sample sizes of the two studies are different and there is no sample overlap - in the latter case the method is called Two-sample MR [33]. Apart from the obvious advantage of using the existing summary statistics, Two-sample MR minimizes the potential residual genetic confounding. We performed Two-sample MR using MR-base via R-package TwoSampleMR [34] (See detailed description: <https://mrcieu.github.io/TwoSampleMR/>).

For our study, exposure instruments were selected from the previous GWAS on food consumption conducted in UK Biobank data (up to N=445,799) [31] which included 25 traits for which valid instruments were available: consumption of beef, beer, bread, champagne/white wine, cheese, cooked vegetables, decaffeinated coffee, dried fruit, fresh fruit, ground coffee, instant coffee, lamb, non-oily fish, oily fish, pork, poultry, processed meat, red wine, salad, salt, spirits, tea, water adjusted for coffee, and vegetarianism and drink temperature.

In order to be able to distinguish between the independent effects of single food items and those arising due to the effect of dietary patterns, we defined 15 “Principal Component traits” (PC-traits) by first clustering the single food items using the iCLUST algorithm [35,36]. The defining and calculation of PC-traits is not a part of this paper and was done for our previous manuscript - therefore only a brief description is included here and more detailed information can be found in Pirastu *et al.* [31]. After clustering the single food items we split the resulting tree dendrogram into different layers depending on the items in each cluster and the degree of similarity. For example, Oily-Fish and Non-Oily fish were first grouped in an overall fish

consumption variable (*Fish-PC1*) and then in a more general healthy foods measure together with *Fruit-PC1* and *Vegetables-PC1*. Finally, they were all used to estimate a measure of overall dietary pattern (*All-PC1-3*). Figure 1 shows the Sankey plot of the relationships between the different defined traits, and Figure 2 represents the loadings of each single food/drink on each of the main PC-traits (a full table of the PC-loadings can be viewed in Table S1 and is visualized in Figure S2). Higher values of *All-PC1* correspond to what is generally considered a “healthy diet”: higher consumption of vegetables, fruit and fish and lower consumption of meat, coffee and alcohol. *All-PC2* separates foods from drinks (alcohol containing beverages and coffee) with higher values corresponding to higher consumption of coffee and alcohol, and lower consumption of the rest of the foods. *Fruit-PC1* corresponds to higher consumption of dried and fresh fruit whereas higher value on *Psychoactive-PC1* corresponds to higher consumption of coffee and alcohol.

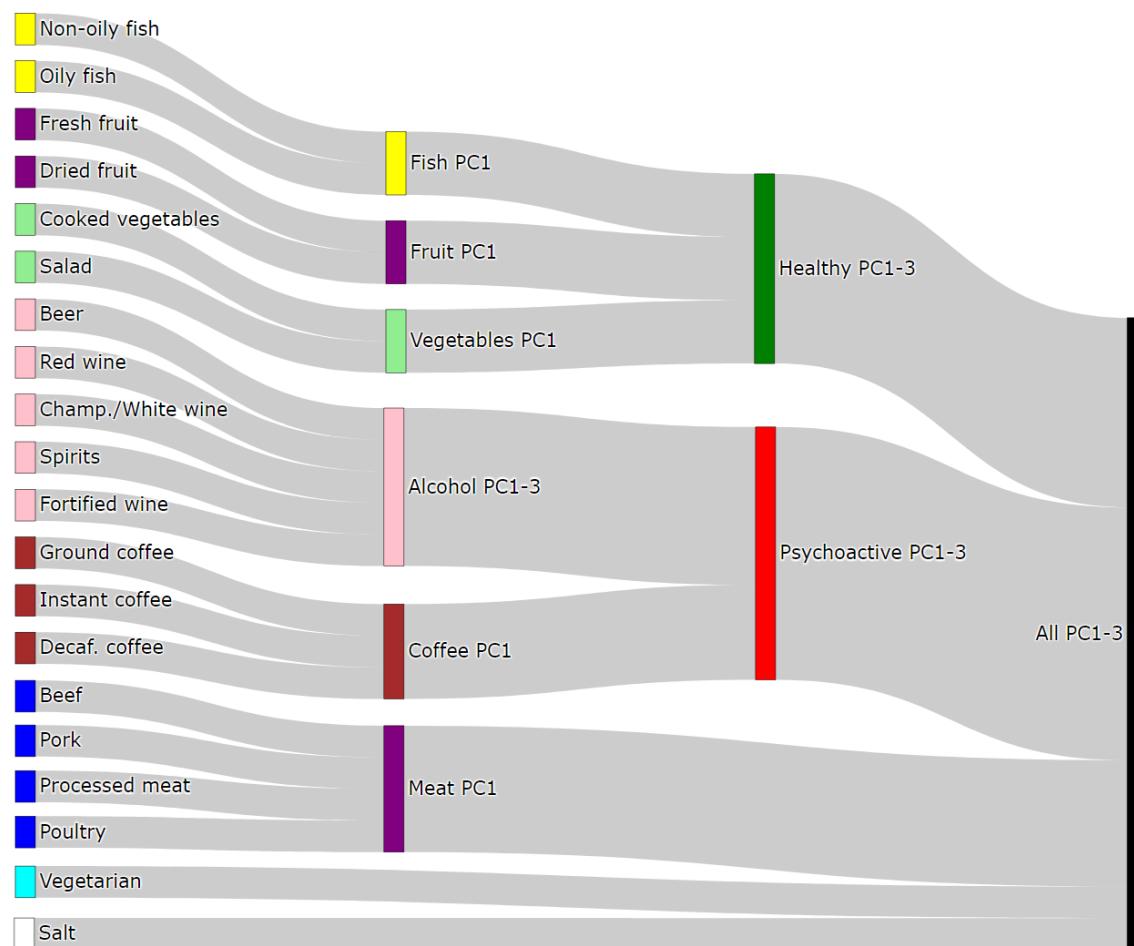


Figure 1 - Sankey diagram of the relationships between dietary items and the principal components traits.

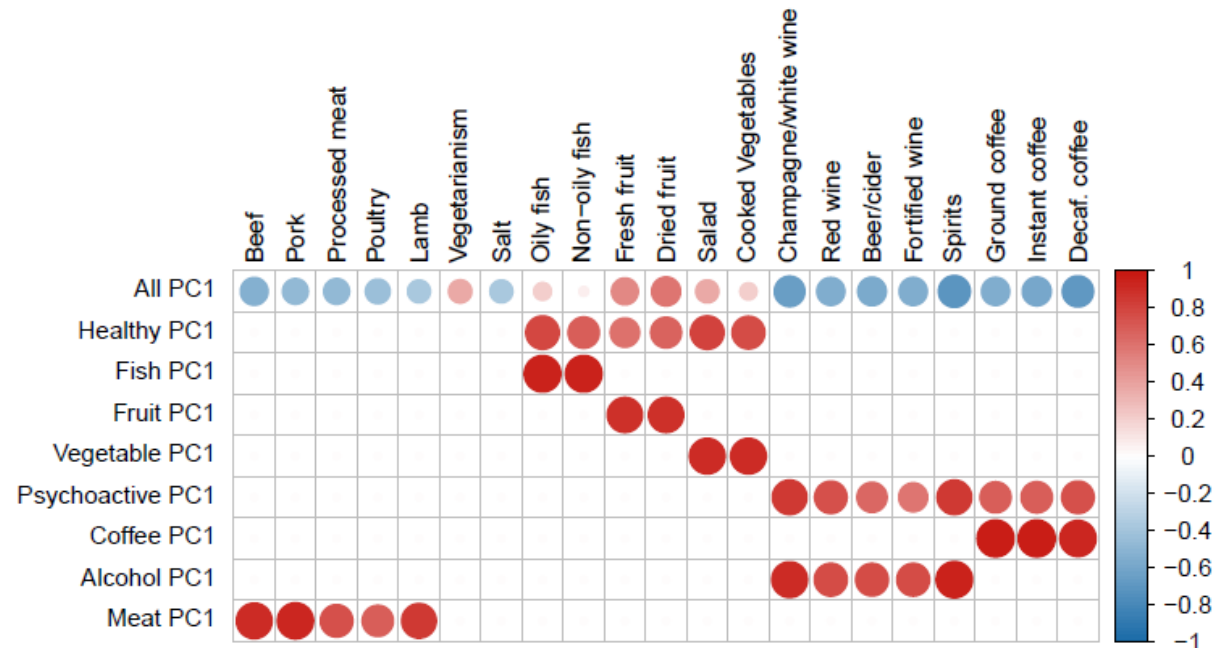


Figure 2 - Loadings of PC traits. The plot represents the loadings of separate dietary items on each of the main PC-traits. Blank squares indicate that the corresponding item is not a component of the PC-trait. The size of the dots corresponds to the magnitude of the loadings whereas the color of the dots describe in addition the direction of the loadings.

The SNP effect estimates for the PC-traits were estimated by applying principal components analysis to the genetic correlation matrix of the items included in each group and by using the resulting rotation matrix to project the estimated effects of the SNPs on each of the items onto the PC space. Standard errors for the PC-traits effects were estimated using the single items SEs and the phenotypic correlation matrix.

As instruments we used independent ($r^2 < 0.001$) SNPs significantly associated with the traits in the exposure GWAS ($p\text{-value} < 5 \times 10^{-8}$) for which the effect was not mediated through other confounders or health-related traits. In case of the PC-traits, SNPs were selected if they were associated ($p\text{-value} < 5 \times 10^{-8}$) with at least one of the items which participated in the trait definitions. Once extracted, we assigned to each selected SNP the lowest $p\text{-value}$ amongst the traits of interest. Finally, we applied LD pruning ($r < 0.001$) to the selected SNPs. We have previously shown that food frequency GWAS results are strongly affected by educational attainment (as a proxy of socioeconomic status) and by health-related traits such as body mass index, blood pressure and cholesterol, for which dietary advice is generally given [31]. This results in either bias by indication (where the behaviour is determined by health advice or belief, e.g. lower fat consumption in people with high cholesterol) or reporting bias (where people under- or over-report food consumption due to their health status, e.g. obese people under-report true fat consumption). If the biasing trait is heritable, this leads to spurious results in the GWAS, which can bias the MR results. In order to distinguish which variables are likely directly associated with the food of interest (and not mediated by health

conditions) we have previously developed a method called Corrected to Uncorrected ratio filtering (CUR), which is based on the idea that if the SNP is directly associated with food preferences, then its effect should not change when adjusted for education status or health conditions [31]. In short, CUR corresponds to the ratio between the effect corrected for the aforementioned covariates and the uncorrected effect, and CUR substantially deviating from 1 is an indication of an invalid instrument. We thus selected only those SNPs which showed $CUR = 1 \pm 0.05$, which we have previously shown maximises the chances of selecting the correct SNPs. The SNPs effects for the outcomes were obtained from a previous GWAS on 123 plasma metabolites in 24,925 individuals [32].

As the main MR method we used the Inverse Variance Weighted method, with random effect standard error if heterogeneity p-value was less than 0.05/123. One of the problems of MR is when SNPs are associated with the outcome through causal paths which do not pass through the exposure of interest, also referred to as horizontal pleiotropy. In order to remove the SNPs with the highest heterogeneity (and thus likely are subject to horizontal pleiotropy) we used the method called MR-Radial [37]. All analyses were thus run on the instruments selected with this method. As sensitivity analyses, we used MR-Median [38], MR-RAPS [39] and MR-Egger [40]. These methods have been thoroughly described elsewhere and are all sensitive to the breaking of different MR assumptions. When only one instrument was available the Wald ratio method was used. Finally, we defined as significant the food-metabolite relationships where Storey's q-value [41] was less than 0.05. The analyses were run using R version 3.6.1 [42].

Results

After correcting for multiple testing by using false discovery rate ($FDR < 0.05$) via Storey's q-values, 417 potentially causal relationships remained statistically significant. Most of these are associations related to atherogenic lipoproteins: very-low density lipoproteins (VLDL), intermediate density lipoproteins (IDL) and low-density lipoproteins (LDL), which all contain Apolipoprotein B (ApoB). Figures 3 and 4 are heatmaps reporting the relationships between food-items and -groups with the metabolites, Figure 3 depicts atherogenic lipoproteins and related metabolites, and Figure 4 depicts all other food-metabolite relationships of interest. For readability some metabolites which had only single significant associations are left out of these graphs. A full list of all significant food-metabolite relationships that we detected, is found in Table S2.

Out of the 417 significant results, 19 were based on the Wald ratio and 398 were based on the Inverse Variance Weighted method. For the latter, we used MR-Median, MR-RAPS and MR-Egger as sensitivity analyses to detect potential violation of assumptions. We found that the sensitivity analyses broadly agreed with the results of the main analysis with the exception of one case (*Alcohol-PC2* on "Description of average fatty acid chain length, not actual carbon number"), where the direction of the effect estimate from MR-Median was opposite to the one from the main analysis and thus this result cannot be considered reliable. In 16 cases (mostly when *Alcohol-PC2* was the exposure), the betas from MR-median were smaller

(>50% effect difference) than the ones in the main analysis, which indicates that in these cases the effect sizes might be inflated. For some food-metabolite pairs, we detected some indication of heterogeneity. This was mostly related to PC-traits which reflect the effects of many different exposures and it is thus not unexpected.

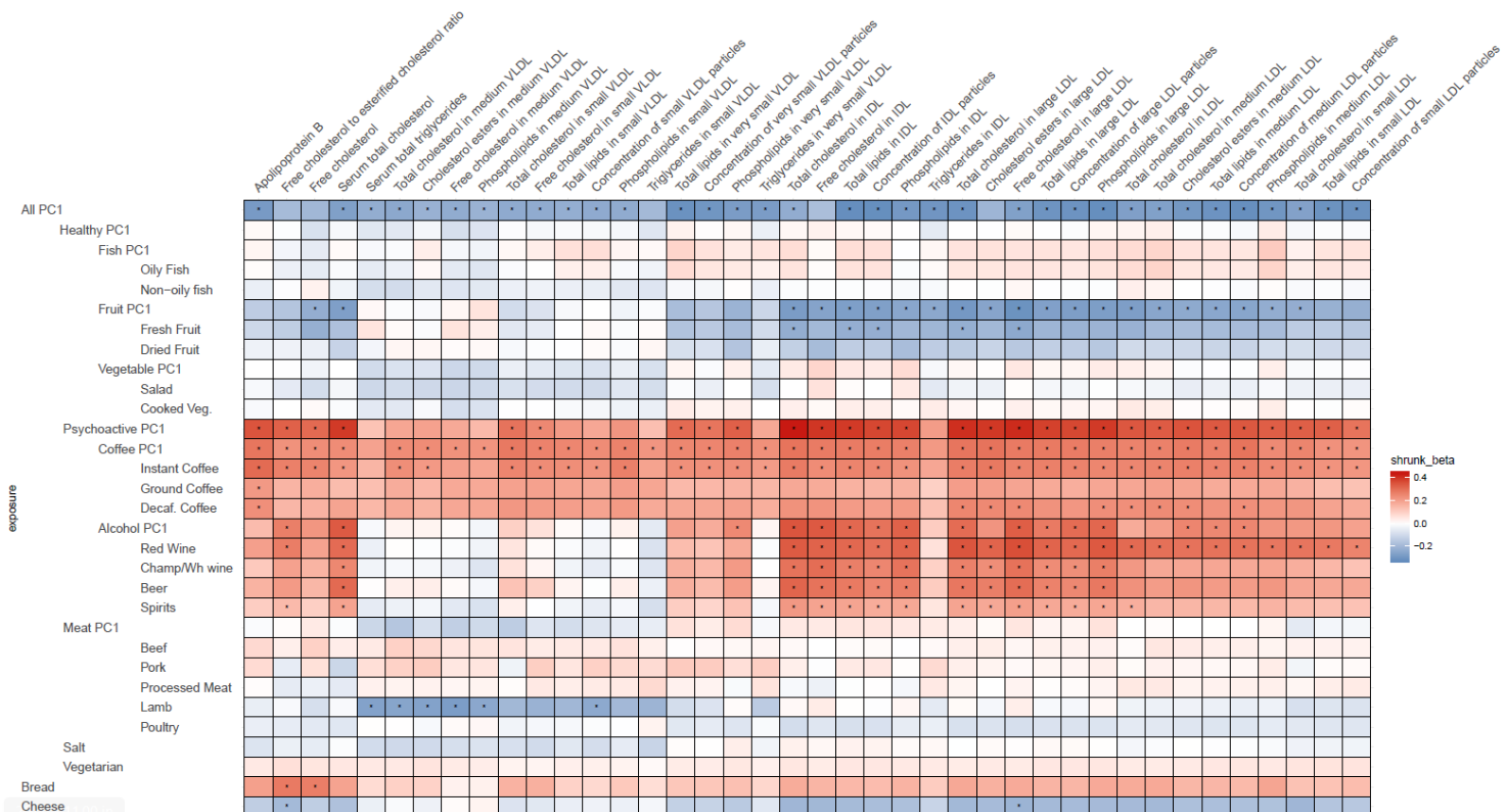


Figure 3. Heatmap of the relations of food traits with atherogenic lipoproteins and related metabolites (VLDL, IDL, LDL and related). Depicted are only the metabolites which showed significant association with the food-items or -groups. To facilitate meaningful visualisation and maximise the appearance of signal rather than noise, we applied a shrinkage method - imposing a bayesian prior assumption on the distribution of beta (mean 0, SD 0.1), and conjugating that with the likelihood of our results and then taking mean beta from the resulting distribution, thus shrinking estimates with larger SEs more towards 0. The color of the squares indicates the size and direction of betas after a shrinking procedure described in the Methods. The FDR-significant results are marked with “*” in the middle of the square.

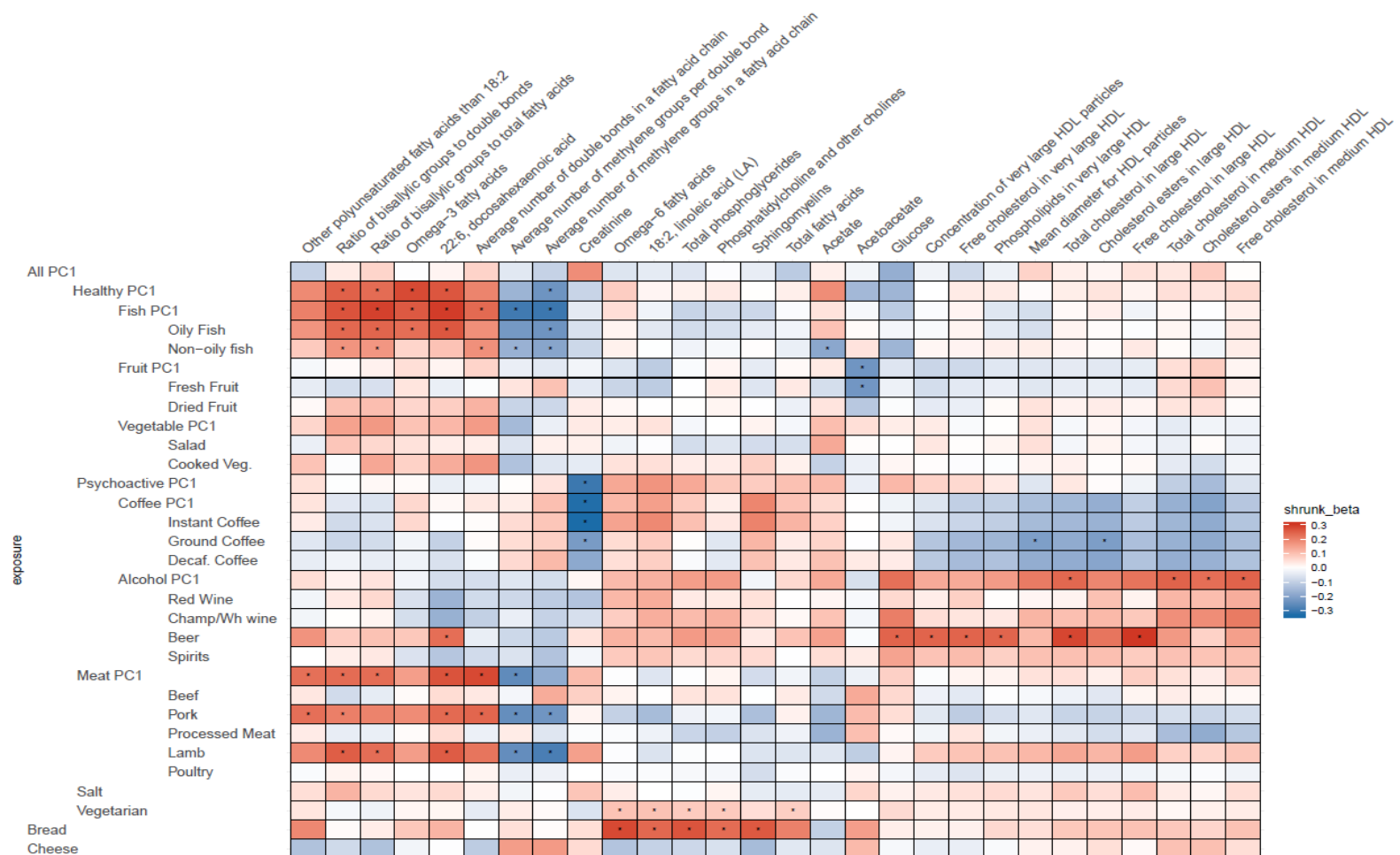


Figure 4. Heatmap of the relations of food traits and all other metabolites (not VLDL, IDL, LDL related). Depicted are only the metabolites which showed significant association with the food-items or -groups. To facilitate meaningful visualisation and maximise the appearance of signal rather than noise, we applied a shrinkage method - imposing a bayesian prior assumption on the distribution of beta (mean 0, SD 0.1), and conjugating that with the likelihood of our results and then taking mean beta from the resulting distribution, thus shrinking estimates with larger SEs more towards 0. The color of the squares indicates the size and direction of betas after a shrinking procedure described in the Methods. The FDR-significant results are marked with “*” in the middle of the square.

Most significant relationships were found with the overall “healthy diet” trait (*All-PC1*), higher consumption of coffee (*Coffee-PC1*) and consumption of alcoholic beverages and coffee (*Psychoactive-PC1* and *-PC2*). The majority of the effects of these traits were on the same metabolites (VLDL, IDL and LDL related) and with opposite directions: namely showing negative correlation for *All-PC1*, where higher values correspond to a generally healthier diet, and positive correlations for *Coffee-PC1* and *Psychoactive-PC1*. These effects were very similar to each other in size, with betas ranging from -0.480 to -0.315 for *All-PC1*, from 0.327 to 0.436 for *Coffee-PC1* and from 0.426 to 0.778 for *Psychoactive-PC1*. The latter is likely partly due to the strong correlation between these lipid measurements.

All-PCI had significant associations with 43 of the measured 123 metabolites, which was, as expected, the highest number of associations we saw for any of the tested foods or PC's. Interestingly these effects were mostly related to atherogenic lipoproteins: the components of VLDL, IDL and LDL of various sizes, and ApoB; whereas surprisingly there were no notable effects on any of the components of high density lipoprotein (HDL) particles or other metabolites measured (for example omega-3 or omega-6 fatty acids). Additionally, *All-PCI* had significant effects on serum total cholesterol and serum total triglycerides. When a significant causal relationship was found for a metabolite, the associations of this metabolite with each dietary item were examined. Often the effects on PC-traits were driven only by a subset of dietary items. For example, in the case of total lipids in IDL, although we see a clear effect of overall diet (*All-PCI*), it seems that this is driven primarily by the components of *Fruit-PCI* and *Psychoactive-PCI* while the remaining foods do not seem to play any role. This observation may explain some of the significant heterogeneity we detected.

Some strikingly clear patterns can be noted, when examining the results that are depicted on Figure 3 - for example looking at the significant positive effects that *Psychoactive-PCI* has on the components of IDL particles, large LDL particles, and serum total cholesterol. In the case of these results both subcategories - coffee and alcohol - clearly contribute to the overall effect of *Psychoactive-PCI*. More importantly, all the four subcategories of alcohol show independently significant effects. All the significant effects that *Psychoactive-PCI* has on the components of IDL particles and large LDL particles are accompanied by a contrary effect from *Fruit-PCI*, which shows significant negative effects. The latter indicates that higher consumption of fruits has a lowering effect on several IDL and large LDL components. In these cases, the effects of *Fruit-PCI* are with the same direction as the effects of *All-PCI*, indicating the healthier diet. Furthermore, this association pattern is also partly followed by the components of medium LDL particles, but in this case, when looking into the components of *Psychoactive-PCI*, alcohol seems to play less important role compared to coffee.

In most of the cases alcohol and coffee both contributed to the overall effect of *Psychoactive-PCI* and we cannot distinguish between the effect of the components (although alcohol seems to play a slightly larger role in the case of IDL particles). Therefore, in most of the cases *Alcohol-PCI*, *Coffee-PCI* and the items comprising these behave in the same way. Nevertheless, there is a group of clear and notable counterexamples: the components of medium VLDL and small VLDL particles. In these cases, *Coffee-PCI* shows clear positive effects and all the betas of all the coffee subgroups agree with the direction of these effects, whereas *Alcohol-PCI* and all of its subgroups show no clear effects and do not seem to be associated with the components of medium and small VLDL particles. Another interesting example that behaves differently from other items in the group of *Psychoactive-PCI* is beer, which has significant positive effects on glucose ($\beta = 0.43$, 95% CI: 0.15 ; 0.70), concentration of very large HDL particles ($\beta = 0.47$, 95% CI: 0.16 ; 0.77), free cholesterol in very large HDL ($\beta = 0.45$, 95% CI: 0.16 ; 0.75), phospholipids in very large HDL ($\beta = 0.46$, 95% CI: 0.16 ; 0.76), total cholesterol in large HDL ($\beta = 0.53$, 95% CI: 0.23 ; 0.83), free cholesterol in large HDL ($\beta = 0.57$, 95% CI: 0.27 ; 0.86), and on 22:6 docosahexaenoic acid (DHA, a

subgroup of omega-3 fatty acids; $\beta = 0.84$, 95% CI: 0.32 ; 1.37). The latter is an example, where beer is clearly the odd-one-out compared to the effects of other alcohol subgroups, indicating that the effect comes from other ingredients in beer rather than alcohol.

The effects of *Meat-PCI* do not have much contribution from beef, processed meat and poultry and are driven mostly by pork (in the case of polyunsaturated fatty acids than other 18:2, and average number of double bonds in a fatty acid chain) or lamb (in the case of ratio of bisallylic groups to total fatty acids) or both (ratio of bisallylic groups to double bonds, DHA, and average number of methylene groups per double bond). Furthermore, it is noteworthy that lamb as a separate item had significant negative effects on several components of medium, large, very large and largest VLDL particles, namely on total cholesterol, cholesterol esters, free cholesterol and phospholipids in medium and large VLDL; total lipids and triglycerides in very large and largest VLDL; phospholipids in very large VLDL; concentration of small and large VLDL particles, and mean diameter for VLDL particles. Thus, lamb has significant negative effects on the components of those lipoproteins that are largest and with lowest density, whereas showing no notable effects on the components of any of the lipoproteins that are smaller and with higher density than small VLDL particles.

As expected, *Fish-PCI* had significant effects on omega-3 fatty acids and DHA and these effects were clearly driven by oily fish. Furthermore, looking at Figure 4, the results regarding creatinine notably stand out, namely there is a lowering effect of *Psychoactive-PCI* on creatinine, which is clearly driven by coffee (which in turn shows significant negative effect on creatinine). Surprisingly vegetarianism and bread share largely the structure of effects - they both have significant effects on omega-6 fatty acids ($\beta = 3.73$, 95% CI: 1.75 ; 5.70; $\beta = 0.67$ 95% CI: 0.30 ; 1.04 respectively), 18:2 linoleic acid (LA, a subgroup of omega-6 fatty acids; $\beta = 3.83$, 95% CI: 1.86 ; 5.80; $\beta = 0.55$ 95% CI: 0.18 ; 0.91), phosphatidylcholine and other choline ($\beta = 3.27$, 95% CI: 1.31 ; 5.24; $\beta = 0.53$ 95% CI: 0.17 ; 0.89) and total phosphoglycerides ($\beta = 3.19$, 95% CI: 1.23 ; 5.16; $\beta = 0.63$ 95% CI: 0.25 ; 1.00). Of note, the results of vegetarianism are based only on one instrument and the method of Wald ratio was used.

Discussion

In this study we have assessed the effect of long term exposure to single foods and food groups on blood metabolite profiles using Mendelian randomization. We have in general found that in many cases these changes are not due to specific food items but are related to general dietary patterns. This could be due to the fact that most of the metabolites assayed are linked to lipid profile and it does not exclude more specific biomarkers being discovered for single items.

Some of the significant results that we found between healthier diet (represented by higher values on *All-PCI*) and metabolites (such as total lipids in large VLDL, small VLDL, large LDL, medium LDL, small LDL and IDL) replicated the findings of a recent observational study investigating association between metabolites and healthy diet [43], while for others (ApoB, LDL cholesterol (LDL-C), IDL cholesterol (IDL-

C), serum total cholesterol, total lipids in very small VLDL) they showed a non-significant association with the same direction. Many of the results that we found elaborate the lipid profiles of lipoprotein subclasses in more detail than the previous studies. Our results regarding the effects of alcohol on IDL and large LDL related lipids conflicted with some observational studies' findings [18,44], but agreed with another MR-study [45], whereas in our study we showed that the same relationships hold for each of the alcohol subgroups as well. Our finding of vegetarianism raising 18:2 linoleic acid replicated a previous finding from a randomized trial conducted on subjects with T2D [46]. Furthermore the finding of vegetarianism raising omega-6 replicated a previous observational finding [47]. We also saw a significant positive effect of oily fish on omega-3 fatty acids and DHA, which are well known causal relationships [48]. The fact that our results conflicted with some observational findings, but aligned with known causal relationships and with a randomized trial, highlights the strength of MR studies as an intermediate step between observational studies and clinical trials and validates the utility of our approach.

One of the most notable patterns in our results was the one regarding IDL particles and large LDL particles, where higher consumption of coffee and alcohol had elevating effects, whereas higher consumption of fruits and higher value on *All-PCI* (i.e. healthier overall diet) had lowering effects. Higher levels of atherogenic particles and their components is part of a less desirable blood lipoprotein profile due to elevated risk for several cardiovascular diseases [49]. For example, IDL-C has been shown to be associated with the degree and frequency of CAD independent of LDL-C [50]. Furthermore, different properties of LDL, like size, the amount of cholesterol esters and cholesterol, and fatty acid composition are all considered to be aspects of its CVD-causing capability [51]. Large LDL particles have been found to be predictive of coronary artery disease in numerous studies, also when accompanied by low or normal triglyceride levels [52]. IDL and LDL may have slightly different mechanisms by which they advance the progress of coronary artery disease, while still having a direct association with CVD [52]. The clear pattern of associations we saw on IDL and large LDL indicate that these metabolites are likely affected by dietary habits. Since these metabolites are largely shown to affect cardiovascular health, and Mendelian randomization results indicate possible causal pathways, there is considerable scope for further investigation of these results. Furthermore, our results indicate that the harmful effects alcohol and coffee have on cardiovascular health are at least partly mediated by IDL and large LDL lipoproteins.

On the other hand, as we saw from the results, in addition to harmful effects, alcoholic beverages can have somewhat beneficial effects as well, namely elevating effects on the components of large and medium HDL. The effects on large HDL were largely driven by beer or seen only for the beer subgroup. Nevertheless, it is important to note that even though there is epidemiological evidence of HDL having a beneficial role for cardiovascular health, the extent and specific mechanism is still not fully understood or proven [53], and the studies have failed to show a causal link [54]. Furthermore, there was a surprising positive effect of beer on DHA, which has been shown to have a cardioprotective effect, as concluded in a recent large meta-analysis of randomized control trials [55]. This effect on DHA was in the opposite direction compared to other alcoholic beverages, indicating that the beneficial effect is likely not due to alcohol itself but some

other ingredient in beer. We propose that the beer-DHA relationship is worth further investigation and it would be especially interesting to compare regular beer with non-alcoholic beer to detect whether the beneficial effects remain. Furthermore, even if alcoholic beverages can have some beneficial effects, one can not underestimate the negative impact of elevated IDL and LDL levels on health coming from alcohol itself.

Many of our significant results were related to an overall healthier diet (*All-PCI*), which seems to affect mostly LDL, VLDL, IDL and related subclasses. The fact that the effect sizes were very similar, can be partly explained by the strong correlations between the different metabolites explained in the results section. Another explanation might be the effect that *All-PCI* has on ApoB. Namely, healthier values on *All-PCI* have a lowering effect on ApoB. The latter is the protein part of all the lipoproteins formed in the liver that carry cholesterol and triglycerides to the cells of the body. Those lipoproteins are VLDL, IDL, LDL. As there is only one ApoB protein per each such lipoprotein [56], one can assume that the total amount of ApoB will reflect the whole amount of atherogenic lipoprotein particles. This is important, because it shows that ApoB, being a component of all atherogenic lipoprotein particles, might reflect the actual CVD risk better than the amount of cholesterol in any lipoprotein particle type on their own [57]. The latter has been shown in some studies, for example Walldius and Jungner [58] found that ApoB levels predicted the risk of CVD better than LDL-C and especially in patients who didn't have elevated LDL-C levels. Ference *et al.* [59] came to the conclusion that the risk of CVD posed by LDL particles is more determined by the concentration of LDL particles measured by ApoB compared to the mass of cholesterol in LDL particles. Because of the lowering effect of *All-PCI* on ApoB, we can assume that ApoB might be one of factors playing a substantial role in the relationship between diet and health, and that healthy diet changes that contribute to lowering ApoB levels might have the potential to actually reduce the risk of CVD.

We saw a significant increasing effect from *Coffee-PCI* on several VLDL, IDL and LDL lipoprotein subclasses and their components and in addition a negative effect on the mean HDL diameter and cholesterol esters in large HDL. Effects on other HDL parameters were not significant, but there is a notable trend towards a negative correlation. Overall, a higher value on *Coffee-PCI* results in a more unhealthy lipid profile, raising ApoB, serum total cholesterol, VLDL, IDL and LDL levels and their constituents. Nevertheless, in the case of coffee, the results regarding cardiovascular health are controversial and mostly show either beneficial or neutral effects [60] or that moderate consumption is unlikely to have adverse effects [61]. A recent review article concluded that filtered coffee consumption up to 6 cups a day did not increase the risk of cardiovascular diseases, but in fact moderate coffee consumption has been associated with a reduced disease risk [62]. Because coffee contains various biologically active substances besides caffeine, it is hard to pinpoint the different effects we see, compared to the ones in previous studies, to any specific compound. Raised ApoB levels indicate that the overall atherogenic particle amount is higher, which, as mentioned before, is a good predictor of CVD risk. Raised VLDL levels might be due to cafestol, a common ingredient in coffee, having an effect of increased VLDL particle assembly rate in the liver [63]. Raised serum cholesterol and LDL levels associated with coffee consumption have been documented before

and are in line with our results [64,65]. Overall, we did not see any beneficial effect of coffee on lipid profile in our results, in fact coffee turned out to have a negative effect on a larger variety of atherogenic lipoproteins than other dietary items. Hence, the total effect of coffee on health remains still unclear, but we propose that the harmful effect of coffee on health might be mediated by ApoB and thus via VLDL, IDL and LDL.

We saw significant increasing effects of vegetarianism on omega-6 fatty acids and on its subgroup 18:2 linoleic acid. The omega-6 result replicated a previous finding by Kornsteiner, Singer and Elmadfa [47], however our MR-analysis showed the potential for a causal relationship. The effect of vegetarian diet on 18:2 linoleic acid replicated a previous result of a randomized controlled trial in T2D patients [46], whereas our results show that this finding applies for the general population in a larger sample as well. These results are important in the context of omega-6/omega-3 ratio, since too high omega-6/omega-3 ratio is associated with multiple diseases, including cardiovascular disease, cancer, inflammatory and autoimmune diseases [66]. The recommended ratio between omega-6 and omega-3 is about 4:1 [66], but vegetarian diet has been shown to have a ratio of 10:1 [47]. Thus, a higher amount of omega-6 may be driving the ratio to an undesired direction and consequently the individuals on a vegetarian diet might, in addition to adding omega-3 rich foods to their menu, need to also pay extra care in reducing the intake of omega-6 rich foods in order to maintain a healthy omega-3/omega-6 ratio.

Our study has several limitations: for some items only a few SNPs were available to use as an instrument in MR; self-reported dietary data is a difficult trait to investigate since it encompasses several biases - we aimed to mitigate this issue by using Corrected-to-Uncorrected Ratio [31]; MR analysis can suffer from horizontal pleiotropy, we tried to mitigate this issue by using sensitivity analyses. Nevertheless, our study has several strengths: to our best knowledge, we are the first ones to perform MR analysis between dietary items and blood metabolites with such a large amount of dietary SNP instruments available; we found several distinct patterns that shed more light on how dietary changes might affect cardiovascular health; we found multiple interesting associations that are worth further investigation via feeding studies or randomized trials. We investigated the effects on metabolites profiled with NMR spectroscopy, which encompass mostly lipid profiles. Future studies with proteomics and metabolite data from mass spectrometry might give more detailed insight about the mechanisms by which food affects health.

In conclusion, we aimed to investigate the relationships between dietary items and blood metabolites in order to gain more insight into the mechanisms by which food affects health. Mendelian randomization proved a useful method for fulfilling this aim. Moreover, we replicated several previous findings and known associations, which validates the method used. We did not detect any reported causal relationships in the literature conflicting with our results, however occasionally our results conflicted with previous observational studies. This demonstrates the strength of MR studies and indicates that some of the previously reported findings might be confounded through other unobserved factors. Nevertheless, in order to give actual dietary intervention suggestions, additional thorough investigations should be carried out via

feeding studies or randomized trials. We believe that many of the potentially causal relationships that have been described here, have promising potential for further investigation.

Supplementary Materials

Figure S1: Loadings of all principal component traits (The plot represents the loadings of separate dietary items on each of the PC-traits. Blank squares indicate that the corresponding item is not a component of the PC-trait. The size of the dots corresponds to the magnitude of the loadings whereas the color of the dots describe in addition the direction of the loadings.), **Table S1:** Loadings of all separate dietary items on each of the principal component traits., **Table S2:** All FDR-significant results between dietary traits and blood metabolites.

Funding

This research was funded by the European Union through the European Regional Development Fund (N.T.) and SP1GI18045T (T.E., N.T.), grants PUT1660 (T.E., N.T.) and PUT1665 (K.F., N.T.) of the Estonian Research Council, and Gentransmed EU RDF No. 2014-2020.4.01.15-0012 (A.M.). We acknowledge support from the MRC Human Genetics Unit programme grant, “Quantitative traits in health and disease” (U. MC_UU_00007/10).

References

1. Popkin, B.M. The Nutrition Transition and Obesity in the Developing World. *J. Nutr.* **2001**, *131*, 871S-873S, doi:10.1093/jn/131.3.871S.
2. Schwingshackl, L.; Hoffmann, G.; Lampousi, A.-M.; Knüppel, S.; Iqbal, K.; Schwedhelm, C.; Bechthold, A.; Schlesinger, S.; Boeing, H. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. *Eur. J. Epidemiol.* **2017**, *32*, 363–375, doi:10.1007/s10654-017-0246-y.
3. Schaefer, E.J. Lipoproteins, nutrition, and heart disease. *Am. J. Clin. Nutr.* **2002**, *75*, 191–212, doi:10.1093/ajcn/75.2.191.
4. Boeing, H.; Bechthold, A.; Bub, A.; Ellinger, S.; Haller, D.; Kroke, A.; Leschik-Bonnet, E.; Müller, M.J.; Oberritter, H.; Schulze, M.; et al. Critical review: vegetables and fruit in the prevention of chronic diseases. *Eur. J. Nutr.* **2012**, *51*, 637–663, doi:10.1007/s00394-012-0380-y.
5. Dilis, V.; Katsoulis, M.; Lagiou, P.; Trichopoulos, D.; Naska, A.; Trichopoulou, A. Mediterranean diet and CHD: the Greek European Prospective Investigation into Cancer and Nutrition cohort. *Br. J. Nutr.* **2012**, *108*, 699–709, doi:10.1017/S0007114512001821.
6. Key, T.J.; Schatzkin, A.; Willett, W.C.; Allen, N.E.; Spencer, E.A.; Travis, R.C. Diet, nutrition and the prevention of cancer. *Public Health Nutr.* **2004**, *7*, 187–200, doi:10.1079/PHN2003588.
7. Johnson, I.T.; Lund, E.K. Review article: nutrition, obesity and colorectal cancer. *Aliment. Pharmacol. Ther.* **2007**, *26*, 161–181, doi:10.1111/j.1365-2036.2007.03371.x.
8. Suhre, K.; Meisinger, C.; Döring, A.; Altmaier, E.; Belcredi, P.; Gieger, C.; Chang, D.; Milburn, M.V.; Gall, W.E.; Weinberger, K.M.; et al. Metabolic Footprint of Diabetes: A Multiplatform Metabolomics Study in an

- Epidemiological Setting. *PLoS ONE* **2010**, 5, e13953, doi:10.1371/journal.pone.0013953.
9. 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.; et al. Metabolite profiles and the risk of developing diabetes. *Nat. Med.* **2011**, 17, 448–453, doi:10.1038/nm.2307.
 10. Würtz, P.; Havulinna, A.S.; Soininen, P.; Tynkkynen, T.; Prieto-Merino, D.; Tillin, T.; Ghorbani, A.; Artati, A.; Wang, Q.; Tiainen, M.; et al. Metabolite Profiling and Cardiovascular Event Risk: A Prospective Study of 3 Population-Based Cohorts. *Circulation* **2015**, 131, 774–785, doi:10.1161/CIRCULATIONAHA.114.013116.
 11. Holmes, M.V.; Millwood, I.Y.; Kartsonaki, C.; Hill, M.R.; Bennett, D.A.; Boxall, R.; Guo, Y.; Xu, X.; Bian, Z.; Hu, R.; et al. Lipids, Lipoproteins, and Metabolites and Risk of Myocardial Infarction and Stroke. *J. Am. Coll. Cardiol.* **2018**, 71, 620–632, doi:10.1016/j.jacc.2017.12.006.
 12. Tynkkynen, J.; Chouraki, V.; van der Lee, S.J.; Hernesniemi, J.; Yang, Q.; Li, S.; Beiser, A.; Larson, M.G.; Sääksjärvi, K.; Shipley, M.J.; et al. 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. *Alzheimers Dement.* **2018**, 14, 723–733, doi:10.1016/j.jalz.2018.01.003.
 13. van der 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.; et al. Circulating metabolites and general cognitive ability and dementia: Evidence from 11 cohort studies. *Alzheimers Dement.* **2018**, 14, 707–722, doi:10.1016/j.jalz.2017.11.012.
 14. 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.; et al. Serum biomarkers of habitual coffee consumption may provide insight into the mechanism underlying the association between coffee consumption and colorectal cancer. *Am. J. Clin. Nutr.* **2015**, 101, 1000–1011, doi:10.3945/ajcn.114.096099.
 15. Shu, X.; Xiang, Y.-B.; Rothman, N.; Yu, D.; Li, H.-L.; Yang, G.; Cai, H.; Ma, X.; Lan, Q.; Gao, Y.-T.; et al. Prospective study of blood metabolites associated with colorectal cancer risk: Colorectal cancer risk associated blood metabolites. *Int. J. Cancer* **2018**, 143, 527–534, doi:10.1002/ijc.31341.
 16. Fischer, K.; Kettunen, J.; Würtz, P.; Haller, T.; Havulinna, A.S.; Kangas, A.J.; Soininen, P.; Esko, T.; Tammesoo, M.-L.; Mägi, R.; et al. Biomarker Profiling by Nuclear Magnetic Resonance Spectroscopy for the Prediction of All-Cause Mortality: An Observational Study of 17,345 Persons. *PLoS Med.* **2014**, 11, e1001606, doi:10.1371/journal.pmed.1001606.
 17. Menni, C.; Zhai, G.; MacGregor, A.; Prehn, C.; Römisch-Margl, W.; Suhre, K.; Adamski, J.; Cassidy, A.; Illig, T.; Spector, T.D.; et al. Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. *Metabolomics* **2013**, 9, 506–514, doi:10.1007/s11306-012-0469-6.
 18. Würtz, P.; Cook, S.; Wang, Q.; Tiainen, M.; Tynkkynen, T.; Kangas, A.J.; Soininen, P.; Laitinen, J.; Viikari, J.; Kähönen, M.; et al. Metabolic profiling of alcohol consumption in 9778 young adults. *Int. J. Epidemiol.* **2016**, 45, 1493–1506, doi:10.1093/ije/dyw175.
 19. Guasch-Ferré, M.; Bhupathiraju, S.N.; Hu, F.B. Use of Metabolomics in Improving Assessment of Dietary Intake. *Clin. Chem.* **2018**, 64, 82–98, doi:10.1373/clinchem.2017.272344.
 20. Brennan, L.; Hu, F.B. Metabolomics-Based Dietary Biomarkers in Nutritional Epidemiology-Current Status and Future Opportunities. *Mol. Nutr. Food Res.* **2019**, 63, 1701064, doi:10.1002/mnfr.201701064.
 21. Esko, T.; Hirschhorn, J.N.; Feldman, H.A.; Hsu, Y.-H.H.; Deik, A.A.; Clish, C.B.; Ebbeling, C.B.; Ludwig, D.S. Metabolomic profiles as reliable biomarkers of dietary composition. *Am. J. Clin. Nutr.* **2017**, 105, 547–554, doi:10.3945/ajcn.116.144428.
 22. Hernández-Alonso, P.; Giardina, S.; Cañueto, D.; Salas-Salvadó, J.; Cañellas, N.; Bulló, M. Changes in Plasma Metabolite Concentrations after a Low-Glycemic Index Diet Intervention. *Mol. Nutr. Food Res.* **2019**, 63, 1700975, doi:10.1002/mnfr.201700975.
 23. Gibbons, H.; Michielsen, C.J.R.; Rundle, M.; Frost, G.; McNulty, B.A.; Nugent, A.P.; Walton, J.; Flynn, A.; Gibney, M.J.; Brennan, L. Demonstration of the utility of biomarkers for dietary intake assessment; proline betaine as an example. *Mol. Nutr. Food Res.* **2017**, 61, 1700037, doi:10.1002/mnfr.201700037.

24. Davey Smith, G.; Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int. J. Epidemiol.* **2003**, *32*, 1–22, doi:10.1093/ije/dyg070.
25. Bergholdt, H.K.; Nordestgaard, B.G.; Ellervik, C. Milk intake is not associated with low risk of diabetes or overweight-obesity: a Mendelian randomization study in 97,811 Danish individuals. *Am. J. Clin. Nutr.* **2015**, *102*, 487–496, doi:10.3945/ajcn.114.105049.
26. Yang, Q.; Lin, S.L.; Au Yeung, S.L.; Kwok, M.K.; Xu, L.; Leung, G.M.; Schooling, C.M. Genetically predicted milk consumption and bone health, ischemic heart disease and type 2 diabetes: a Mendelian randomization study. *Eur. J. Clin. Nutr.* **2017**, *71*, 1008–1012, doi:10.1038/ejcn.2017.8.
27. Chen, L.; Davey Smith, G.; Harbord, R.M.; Lewis, S.J. Alcohol Intake and Blood Pressure: A Systematic Review Implementing a Mendelian Randomization Approach. *PLoS Med.* **2008**, *5*, e52, doi:10.1371/journal.pmed.0050052.
28. Andrews, S.J.; Goate, A.; Anstey, K.J. Association between alcohol consumption and Alzheimer's disease: A Mendelian randomization study. *Alzheimers Dement.* **2020**, *16*, 345–353, doi:10.1016/j.jalz.2019.09.086.
29. Nordestgaard, A.T.; Thomsen, M.; Nordestgaard, B.G. Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study. *Int. J. Epidemiol.* **2015**, *44*, 551–565, doi:10.1093/ije/dyv083.
30. Lee, Y.H. Investigating the possible causal association of coffee consumption with osteoarthritis risk using a Mendelian randomization analysis. *Clin. Rheumatol.* **2018**, *37*, 3133–3139, doi:10.1007/s10067-018-4252-6.
31. Pirastu, N.; McDonnell, C.; Grzeszkowiak, E.J.; Mounier, N.; Imamura, F.; Day, F.R.; Zheng, J.; Taba, N.; Concas, M.P.; Repetto, L.; et al. Using genetics to disentangle the complex relationship between food choices and health status. *bioRxiv* **2019**, 829952, doi:10.1101/829952.
32. Kettunen, J.; Demirkan, A.; Würtz, P.; Draisma, H.H.M.; Haller, T.; Rawal, R.; Vaarhorst, A.; Kangas, A.J.; Lyytikäinen, L.-P.; Pirinen, M.; et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat. Commun.* **2016**, *7*, 11122, doi:10.1038/ncomms11122.
33. Pierce, B.L.; Burgess, S. Efficient Design for Mendelian Randomization Studies: Subsample and 2-Sample Instrumental Variable Estimators. *Am. J. Epidemiol.* **2013**, *178*, 1177–1184, doi:10.1093/aje/kwt084.
34. Hemani, G.; Zheng, J.; Elsworth, B.; Wade, K.H.; Haberland, V.; Baird, D.; Laurin, C.; Burgess, S.; Bowden, J.; Langdon, R.; et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **2018**, *7*, e34408, doi:10.7554/eLife.34408.
35. Revelle, W. Hierarchical Cluster Analysis And The Internal Structure Of Tests. *Multivar. Behav. Res.* **1979**, *14*, 57–74, doi:10.1207/s15327906mbr1401_4.
36. Revelle, W.; Zinbarg, R.E. Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika* **2009**, *74*, 145–154, doi:10.1007/s11336-008-9102-z.
37. Bowden, J.; Spiller, W.; Del Greco M, F.; Sheehan, N.; Thompson, J.; Minelli, C.; Davey Smith, G. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the Radial plot and Radial regression. *Int. J. Epidemiol.* **2018**, *47*, 1264–1278, doi:10.1093/ije/dyy101.
38. Bowden, J.; Davey Smith, G.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314, doi:10.1002/gepi.21965.
39. Zhao, Q.; Wang, J.; Hemani, G.; Bowden, J.; Small, D.S. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *ArXiv180109652 Math Stat* **2019**.
40. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525, doi:10.1093/ije/dyv080.
41. Storey, J.D. The positive false discovery rate: a Bayesian interpretation and the q -value. *Ann. Stat.* **2003**, *31*, 2013–2035, doi:10.1214/aos/1074290335.
42. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing.

2019.

43. Akbaraly, T.; Würtz, P.; Singh-Manoux, A.; Shipley, M.J.; Haapakoski, R.; Lehto, M.; Desrumaux, C.; Kähönen, M.; Lehtimäki, T.; Mikkilä, V.; et al. Association of circulating metabolites with healthy diet and risk of cardiovascular disease: analysis of two cohort studies. *Sci. Rep.* **2018**, *8*, 8620, doi:10.1038/s41598-018-26441-1.
44. Du, D.; Bruno, R.; Blizzard, L.; Venn, A.; Dwyer, T.; Smith, K.J.; Magnussen, C.G.; Gall, S. The metabolomic signatures of alcohol consumption in young adults. *Eur. J. Prev. Cardiol.* **2020**, *27*, 840–849, doi:10.1177/2047487319834767.
45. Rosoff, D.B.; Smith, G.D.; Mehta, N.; Clarke, T.-K.; Lohoff, F.W. *Evaluating the effects of alcohol and tobacco use on cardiovascular disease using multivariable Mendelian randomization*; Genomics, 2019;
46. Kahleova, H.; Matoulek, M.; Bratova, M.; Malinska, H.; Kazdova, L.; Hill, M.; Pelikanova, T. Vegetarian diet-induced increase in linoleic acid in serum phospholipids is associated with improved insulin sensitivity in subjects with type 2 diabetes. *Nutr. Diabetes* **2013**, *3*, e75–e75, doi:10.1038/nutd.2013.12.
47. Kornsteiner, M.; Singer, I.; Elmadfa, I. Very Low n–3 Long-Chain Polyunsaturated Fatty Acid Status in Austrian Vegetarians and Vegans. *Ann. Nutr. Metab.* **2008**, *52*, 37–47, doi:10.1159/000118629.
48. Horrocks, L.A.; Yeo, Y.K. Health benefits of docosahexaenoic acid (DHA). *Pharmacol. Res.* **1999**, *40*, 211–225, doi:10.1006/phrs.1999.0495.
49. Carmena, R.; Duriez, P.; Fruchart, J.-C. Atherogenic Lipoprotein Particles in Atherosclerosis. *Circulation* **2004**, *109*, III-2-III–7, doi:10.1161/01.CIR.0000131511.50734.44.
50. Tatami, R.; Mabuchi, H.; Ueda, K.; Ueda, R.; Haba, T.; Kametani, T.; Ito, S.; Koizumi, J.; Ohta, M.; Miyamoto, S.; et al. Intermediate-density lipoprotein and cholesterol-rich very low density lipoprotein in angiographically determined coronary artery disease. *Circulation* **1981**, *64*, 1174–1184, doi:10.1161/01.CIR.64.6.1174.
51. Lada, A.T.; Rudel, L.L. Associations of low density lipoprotein particle composition with atherogenicity: *Curr. Opin. Lipidol.* **2004**, *15*, 19–24, doi:10.1097/00041433-200402000-00005.
52. Berneis, K.K.; Krauss, R.M. Metabolic origins and clinical significance of LDL heterogeneity. *J. Lipid Res.* **2002**, *43*, 1363–1379, doi:10.1194/jlr.R200004-JLR200.
53. Asztalos, B.F.; Niisuke, K.; Horvath, K.V. High-density lipoprotein: our elusive friend. *Curr. Opin. Lipidol.* **2019**, *30*, 314–319, doi:10.1097/MOL.0000000000000612.
54. Millwood, I.Y.; Bennett, D.A.; Holmes, M.V.; Boxall, R.; Guo, Y.; Bian, Z.; Yang, L.; Sansome, S.; Chen, Y.; Du, H.; et al. Association of *CETP* Gene Variants With Risk for Vascular and Nonvascular Diseases Among Chinese Adults. *JAMA Cardiol.* **2018**, *3*, 34, doi:10.1001/jamacardio.2017.4177.
55. Bernasconi, A.A.; Wiest, M.M.; Lavie, C.J.; Milani, R.V.; Laukkanen, J.A. Effect of Omega-3 Dosage on Cardiovascular Outcomes: An Updated Meta-Analysis and Meta-Regression of Interventional Trials. *Mayo Clin. Proc.* **2020**, doi:https://doi.org/10.1016/j.mayocp.2020.08.034.
56. Elovson, J.; Chatterton, J.E.; Bell, G.T.; Schumaker, V.N.; Reuben, M.A.; Puppione, D.L.; Reeve, J.R.; Young, N.L. Plasma very low density lipoproteins contain a single molecule of apolipoprotein B. *J. Lipid Res.* **1988**, *29*, 1461–1473.
57. Leslie, M. Is it time to retire cholesterol tests? *Science* **2017**, doi:10.1126/science.aar6854.
58. Walldius, G.; Jungner, I. Apolipoprotein B and apolipoprotein A-I: risk indicators of coronary heart disease and targets for lipid-modifying therapy. *J. Intern. Med.* **2004**, *255*, 188–205, doi:10.1046/j.1365-2796.2003.01276.x.
59. 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.; et al. Association of Triglyceride-Lowering *LPL* Variants and LDL-C-Lowering *LDLR* Variants With Risk of Coronary Heart Disease. *JAMA* **2019**, *321*, 364, doi:10.1001/jama.2018.20045.
60. Chrysant, S.G. Coffee Consumption and Cardiovascular Health. *Am. J. Cardiol.* **2015**, *116*, 818–821, doi:10.1016/j.amjcard.2015.05.057.

61. Rebello, S.A.; van Dam, R.M. Coffee Consumption and Cardiovascular Health: Getting to the Heart of the Matter. *Curr. Cardiol. Rep.* **2013**, *15*, 403, doi:10.1007/s11886-013-0403-1.
62. van Dam, R.M.; Hu, F.B.; Willett, W.C. Coffee, Caffeine, and Health. *N. Engl. J. Med.* **2020**, *383*, 369–378, doi:10.1056/NEJMr1816604.
63. de Roos, B.; Caslake, M.J.; Stalenhoef, A.F.; Bedford, D.; Demacker, P.N.; Katan, M.B.; Packard, C.J. The coffee diterpene cafestol increases plasma triacylglycerol by increasing the production rate of large VLDL apolipoprotein B in healthy normolipidemic subjects. *Am. J. Clin. Nutr.* **2001**, *73*, 45–52, doi:10.1093/ajcn/73.1.45.
64. Williams, P.T.; Wood, P.D.; Vranizan, K.M.; Albers, J.J.; Garay, S.C.; Taylor, C.B. Coffee intake and elevated cholesterol and apolipoprotein B levels in men. *JAMA* **1985**, *253*, 1407–1411.
65. Kokjohn, K.; Graham, M.; McGregor, M. The effect of coffee consumption on serum cholesterol levels. *J. Manipulative Physiol. Ther.* **1993**, *16*, 327–335.
66. Simopoulos, A.P. The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomed. Pharmacother.* **2002**, *56*, 365–379, doi:10.1016/S0753-3322(02)00253-6.