Clinical Nutrition 41 (2022) 441–451

Contents lists available at ScienceDirect

Clinical Nutrition

journal homepage: http://www.elsevier.com/locate/clnu

Randomized Control Trials

Analysis of the SYSDIET Healthy Nordic Diet randomized trial based on metabolic profiling reveal beneficial effects on glucose metabolism and blood lipids

Gözde Gürdeniz [a, b, *], Matti Uusitupa [c], Kjeld Hermansen [d], Markku J. Savolainen [e], Ursula Schwab [c, f], Marjukka Kolehmainen [c], Lea Brader [d], Lieselotte Cloetens [g], Karl-Heinz Herzig [h, i], Janne Hukkanen [e], Fredrik Rosqvist [j], Stine Marie Ulven [k], Ingibjörg Gunnarsdóttir [l, m], Inga Thorsdottir [l, m], Matej Oresic [n, o], Kaisa S. Poutanen [c, o], Ulf Risérus [j], Björn Åkesson [g, p], Lars Ove Dragsted [a]

[a] Department of Nutrition, Exercise and Sports, University of Copenhagen, Frederiksberg, Denmark
[b] Department of Food Science, University of Copenhagen, Frederiksberg, Denmark
[c] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
[d] Department of Endocrinology and Internal Medicine, Aarhus University Hospital, Aarhus, Denmark
[e] Institute of Clinical Medicine, Department of Internal Medicine, University of Oulu, Oulu, Finland
[f] Department of Medicine, Endocrinology and Clinical Nutrition, Kuopio University Hospital, Kuopio, Finland
[g] Biomedical Nutrition, Pure and Applied Biochemistry, Lund University, Lund, Sweden
[h] Institute of Biomedicine and Biocenter of Oulu, University of Oulu, Finland
[i] Department of Psychiatry, Kuopio University Hospital, Kuopio, Finland
[j] Department of Public Health and Caring Sciences, Uppsala University, Sweden
[k] Department of Nutrition, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway
[l] Unit for Nutrition Research, Faculty of Food Science and Nutrition, University of Iceland, Reykjavík, Iceland
[m] Unit for Nutrition Research, Landspitali National University Hospital, Reykjavik, Iceland
[n] Turku Center for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland
[o] VTT Technical Research Centre of Finland, Espoo, Finland
[p] Department of Clinical Nutrition, Skåne University Hospital, Lund, Sweden

ARTICLE INFO

Article history:
Received 11 April 2021
Accepted 20 December 2021

Keywords:
LC–MS metabolomics
Healthy Nordic diet
Randomized controlled trial
Glucose and lipid metabolism
Plasma metabolite scores

SUMMARY

Background & aims: Intake assessment in multicenter trials is challenging, yet important for accurate outcome evaluation. The present study aimed to characterize a multicenter randomized controlled trial with a healthy Nordic diet (HND) compared to a Control diet (CD) by plasma and urine metabolic profiles and to associate them with cardiometabolic markers.

Methods: During 18–24 weeks of intervention, 200 participants with metabolic syndrome were advised at six centres to eat either HND (*e.g.* whole-grain products, berries, rapeseed oil, fish and low-fat dairy) or CD while being weight stable. Of these 166/159 completers delivered blood/urine samples. Metabolic profiles of fasting plasma and 24 h pooled urine were analysed to identify characteristic diet-related patterns. Principal components analysis (PCA) scores (*i.e.* PC1 and PC2 scores) were used to test their combined effect on blood glucose response (primary endpoint), serum lipoproteins, triglycerides, and inflammatory markers.

Results: The profiles distinguished HND and CD with AUC of 0.96 ± 0.03 and 0.93 ± 0.02 for plasma and urine, respectively, with limited heterogeneity between centers, reflecting markers of key foods. Markers of fish, whole grain and polyunsaturated lipids characterized HND, while CD was reflected by lipids containing palmitoleic acid. The PC1 scores of plasma metabolites characterizing the intervention is associated with HDL ($\beta = 0.05$; 95% CI: 0.02, 0.08; $P = 0.001$) and triglycerides ($\beta = -0.06$; 95% CI: −0.09, −0.03; $P < 0.001$). PC2 scores were related with glucose metabolism (2 h Glucose, $\beta = 0.1$; 95% CI: 0.05, 0.15; $P < 0.001$), LDL ($\beta = 0.06$; 95% CI: 0.01, 0.1; $P = 0.02$) and triglycerides ($\beta = 0.11$; 95% CI: 0.06, 0.15; $P < 0.001$). For urine, the scores were related with LDL cholesterol.

Abbreviations: Glu, glucose; APO, apolipoprotein; BP, blood pressure; hsCRP, high-sensitivity C-reactive protein; HMW, high molecular weight; Ins, insulin; ISI, insulin sensitivity index.

* Corresponding author. University of Copenhagen, Frederiksberg Campus, Rolighedsvej 30, 1958 Frederiksberg, Denmark.
E-mail address: gozde.gurdeniz@food.ku.dk (G. Gürdeniz).

https://doi.org/10.1016/j.clnu.2021.12.031

Conclusions: Plasma and urine metabolite profiles from SYSDIET reflected good compliance with dietary recommendations across the region. The scores of metabolites characterizing the diets associated with outcomes related with cardio-metabolic risk. Our analysis therefore offers a novel way to approach a per protocol analysis with a balanced compliance assessment in larger multicentre dietary trials.
The study was registered at clinicaltrials.gov with NCT00992641.

Abbreviations

AR(Cn:m)	Alkylresorcinol
AUC	area under the curve
CD	control diet
CMPF	3-Carboxy-4-methyl-5-propyl-2-furanpropanoic acid
CMHPF	3-Carboxy-4-methyl-5-hydroxypropyl-2-furanpropanoic acid
CN(Cn:m)	acylcarnitine
CPF	3-Carboxy-5-propyl-2-furanpropanoic acid
CMPentylF	3-Carboxy-4-methyl-5-pentyl-2-furanpropanoic acid
DHA	docosahexaenoic acid
DHBA	3,5-dihydroxybenzoic acid
DHPPA	3,5-dihydroxyphenylpropionic acid
EPA	eicosapentaenoic acid
HND	healthy Nordic diet
LPC(n:m)	lysophosphatidylcholine
PE(n:m)	Phosphatidylethanolamine
PC(n:m)	phosphatidylcholine
PCA	principal component analysis
PLS-DA	Partial least squares discriminant analysis
PUFA	polyunsaturated fatty acid
SR	selectivity ratio
QTOF	quadrupole time-of-flight mass spectrometer
SM(n:m)	Sphingomyelin
UI	unidentified Brackets
UPLC	ultra-performance liquid chromatography
VIP	variable importance in projection
(Cn:m)	include number of carbons, n, and number of double bonds, m

1. Introduction

Observational studies suggest that consumption of diets containing fruits, vegetables, fish, nuts and whole grain reduce the risk of cardiovascular disease and type 2 diabetes [1–3]. Adherence to the Mediterranean or the DASH diets, which consist of many of these healthy food components, substantially improves several risk factors such as insulin resistance, blood lipid profile and blood pressure [4,5]. However, implementing *e.g.* the Mediterranean diet in populations elsewhere is challenging owing to regional and cultural differences. Therefore, the concept of healthy regional diets was developed. The healthy Nordic diet (HND) is in accordance with the Nordic Nutrition Recommendations [6] and characterized by food items easily available and traditionally consumed in the Nordic region, *e.g.* whole-grain rye, oat and barley products, Nordic berries, fish, root vegetables, and rapeseed oil [7]. Dietary interventions with Nordic diets have consistently provided favourable effects on weight loss, blood pressure, inflammation and blood lipid profiles [8–11]. In addition, observational studies point to a lower risk of cardiovascular diseases [12], type 2 diabetes [13] as well as of lower mortality [14] in individuals following the recommendations for a healthy Nordic diet.

Nutritional research is challenged by the lack of objective dietary assessment tools and relies largely on subjective tools such as food frequency questionnaires. Biomarkers, instead, provide objective estimates of dietary intake and they can be used in combination with questionnaires or food records as a measure of compliance [15]. Metabolic profiling has emerged as a powerful tool, facilitating discovery of biomarkers, *i.e.* proline betaine for citrus fruit intake [16] and alkylresorcinols (ARs) for whole grain wheat and rye intake [17]. Combinations of food intake biomarkers have been used to investigate compliance to a diet following the Nordic dietary guidelines [18].

A randomized, controlled, multicenter intervention study (SYSDIET) was performed to investigate whether a HND would have beneficial effects on glucose metabolism (primary outcome) or blood lipid profiles, and inflammatory markers at a weight stable condition. Intention-to-treat analysis revealed that HND improved the lipid profile and had a beneficial effect on low grade inflammation yet did not have an influence on the primary outcome [10]. For the same study, a "biomarker score" combining targeted dietary biomarkers previously showed that the most compliant individuals had a more favourable effect of the HND [19]. In the current substudy, we applied untargeted metabolic profiling (a secondary trial outcome) to identify the metabolites differentiating HND from the control diet (CD). The changes in metabolic profiles characterizing the actual dietary intakes of the volunteers were used to objectively assess the change in the clinical outcomes. This approach aims to approach *per protocol* analysis of this multicentre study based on compliance while allowing for centre differences in food culture, and therefore represents a novel way to re-analyse dietary intervention studies.

2. Materials and methods

2.1. Study design

A randomised dietary parallel multicentre intervention study was carried out in six centers [10]. In total, 200 overweight participants with features of metabolic syndrome at a mean age of 55 years and an average BMI of 31.6 kg/m^2 (67% women) were recruited. After a four-week run-in period, during which participants consumed their habitual diet, they were randomized by the lead centre (Kuopio) to balance centre, age, sex, BMI and fasting plasma glucose (normal/abnormal) between groups to follow a CD or to HND. The HND was based on traditional foods originating

from Nordic countries providing a nutrient composition in line with the Nordic nutrition recommendations [6] whereas CD was defined by the mean nutrient intake in each of the Nordic countries and therefore CD represents an average Nordic diet.

The volunteers in the HND group were recommended to increase their intake of products high in whole grain – *i.e.* rye, barley and oat – berries, fruit and vegetables, whereas the participants in the CD received advice to eat common low-fibre wheat cereal products like refined wheat bread and pasta, rice and butter and not to moderate their intake of fruit and vegetables. Most of the food items according to the calculated needs to keep the volunteer weight stable were provided on a weekly basis to the volunteers in both groups but for logistical reasons staples and frozen foods were sometimes provided in the amounts needed during longer intervals (up to one month). Due to differences in food cultures, these foods differed by center, see Supplementary Table 1. Both diets were isocaloric to keep volunteers weight stable during the intervention. The intervention period was 18 weeks for centers in Aarhus, Uppsala, Reykjavik and Oulu and 24 weeks for Lund and Kuopio. In addition to visits and sampling of biological specimens at baseline and at the end of the intervention, participants also visited the study clinics at week 12 for sampling. The study flow chart is given in Supplementary Fig. 1 and other details of the study has been previously published [10]. The number of participants included in the present metabolic profiling study was 98 in the HND and 71 in the CD group. These included all participants who delivered plasma or urine samples at baseline and at least once after the baseline visit (*i.e.* weeks 12, 18 or 24).

The main outcome measure was a change in glucose metabolism (conglomerate outcome) and the study was originally dimensioned to find a difference in HOMA by intention-to-treat analysis (alpha 0.05, beta 0.80) with 120 subjects per group and 20% attrition. No per-protocol analysis was intended in the original protocol.

All study participants provided written informed consent, and local ethical committees of all the participating centers approved the study protocol. Samples were pseudonymized to remove any person-identifiable information before transfer from each center for metabolic profiling in Copenhagen. The study was registered in the clinicaltrials.gov database under NCT00992641.

2.2. Sample preparation and UPLC-QTOF analysis

The metabolic profiles were acquired for fasting blood from 166 participants and 24-h pooled urine samples from 159 participants.

Plasma protein precipitation was performed as described earlier [20]. The urine was diluted 1:1 with aqueous 5% 30:70 (v/v) acetonitrile–methanol (Optima grade LC–MS, Fisher Scientific, USA). An ultra-performance liquid chromatography (UPLC) system coupled to a quadrupole time-of-flight mass spectrometer (Premier QTOF, Waters Corporation, Manchester, UK) was used for sample analysis. Sample analysis is based on previously described analytical methodology [21] and further details are provided in Supplementary Methods 1. Blanks, external metabolomics standards mixtures, and pooled samples were injected initially, after every 30 samples, and at the end of each analytical batch. Samples from each center were randomized among batches, yet samples of each participant were randomized within each batch.

The flow chart (Supplementary Fig. 1) provides the number of participants included in the analysis and the number of participants at each visit used in the analysis is given in Supplementary Table 2.

2.3. Metabolite identification

The fidelity of chemical identification was defined at four levels according to the published guidelines for metabolomics [22]. 3-Carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF), 3,5-ihydroxybenzoic acid (3,5-DHBA), and 3,5-dihydroxyphenylpropionic acid (3,5-DHPPA) were annotated using an in-house database, which includes retention time and MS/MS spectra of authentic standards. Some of the lysophosphatidylcholines (LPCs) and phosphatidylcholines (PCs) that were not present in the in-house standard database were identified using a previously described Matlab function [23]. The authentic standards, trans-2-octenoylcarnitine (CN(C8:1)), 5-Heneicosylresorcinol (AR(C21:0)), nonadecylresorcinol (AR(C19:0)), and human liver S9 and microsomes were purchased from Sigma Aldrich. MS/MS spectra were compared with comparable spectra in available databases, such as METLIN, HMDB, mzCloud and produced by in-silico fragmentation tools such as CFM-ID, FingerID and lipidmaps. Glucuronide conjugates of CMPF, AR(C19:0), AR(C21:1), 3,5-DHBA and 3,5-DHPPA were produced using human liver S9 as described previously [24]. CMHPF, a hydroxylated metabolite of CMPF, was produced by incubating CMPF with liver microsomes [25,26] and the identity confirmed as described in Supplementary Fig. 2, whereas the annotation CPF and CMPentylF was based on MS/MS matches with CMPF.

2.4. Data preprocessing

XCMS [27] was used to preprocess raw LC–MS data (MZdata) while detected ions were grouped by CAMERA [28]. The XCMS-preprocessed data was imported into Matlab (MathWorks, MA, US). The data acquired in positive and negative ionization modes were concatenated. Noisy, irrelevant and redundant features were filtered as described in Supplementary Methods 2.

Urine samples that were collected at baseline were excluded since they represent the recent, uncontrolled 24 h food intake for each volunteer, whereas fasting plasma sample baseline measurements were subtracted from the corresponding measurements at weeks 12 and 18–24 measurements to reduce between-subject differences in the metabolic profiles. A plasma dataset of weeks 12 and 18–24 without baseline subtraction was also produced for combined analysis of plasma and urine. The use of both 12 weeks and 18–24 weeks in the analysis was used to improve the coverage of metabolite changes after the two diets. Urine samples were normalized using probabilistic quotient normalization [29].

2.5. Clinical measurements

Measurement of glucose, blood lipids, apolipoproteins, inflammatory markers and anthropometric parameters were performed as described in detail previously [10].

2.6. Statistical analysis

2.6.1. Selection of the metabolites discriminating the diet groups

PLS_Toolbox (version 6.5, Eigenvector Research, Inc., MA, US) was used for principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA). Auto-scaled data was subjected to PLS-DA to differentiate the plasma and urine metabolic patterns associated with the diet (HND *vs.* CD). The selection of discriminative variables is based on their variable importance in projection (VIP) and selectivity ratio (SR). To validate the PLS-DA models, the datasets were randomly divided into training sets and test sets, including 80% and 20% of the total sample metabolic profiles, respectively. Training samples were used for feature selection and model development while test set samples were solely used to evaluate the performance of the model with selected features. Training samples were used to iteratively build PLS-DA models where features with low VIP and SR were removed, and

then a new model was built based on the remaining features. The feature reduction process was repeated as long as the remaining selected features improved the performance of the model in terms of classification accuracy of the cross-validation set (evaluated by the error rate and AUC). When reduction of the features was no longer able to improve the performance of the cross-validation set, the PLS-DA models with a final number of features were evaluated using the classification performance of the test set by error rate and AUC. This procedure was repeated with 100 different randomly selected training/test set pairs. The variables that were selected in 80% or more of the randomly selected training/test set pairs formed the final discriminative feature set.

2.6.2. PCA of plasma and urine metabolites associated with the intervention

To get an overview of the differences between the two diets we used the combined discriminant features from PLS-DA analysis of urine and plasma from weeks 12 and 18–24 without baseline correction. These plasma and urine features were summarized using PCA, (1) to investigate whether the dietary differences after intervention can be visually captured by the selected features (scores) and (2) to identify the contribution of each diet and the relationship between the signature foods represented by the selected features (loadings). The combined set of plasma and urine features discriminating the diet groups was decomposed by two-component PCA. The resulting PC1 and PC2 scores are labelled by treatment group (scores plot) and the contributing features were labelled by their dietary source (loadings plot).

An additional PCA was performed for features selected by PLS-DA, *i.e.* discriminating the diets of baseline subtracted plasma combining weeks 12 and 18–24. The six participants who gained or lost more than 4 kg of weight were excluded from this analysis to replicate the previously published per protocol analysis [10]. All annotated features were diet-related and the PC1 and PC2 scores for each volunteer were recorded for further use as individual diet-related scores. The same procedure was used for PCA analysis of features identified as selective for HND and CD in urine, combining weeks 12 and 18–24 without baseline correction, and PC1 and PC2 urine scores were recorded for each participant.

2.6.3. Testing the relation between PCA scores reflecting diet and clinical measurements

PC scores characterizing the individuals' dietary similarity to HND or CD were tested to see whether they could explain the primary outcome, glucose metabolism, as well as other SYSDIET clinical outcomes [10], *i.e.*, serum lipids and apolipoproteins, inflammatory markers and blood pressure, using linear mixed effects models in R (version 3.6.1, lme4 package). The linear mixed model data analysis protocol published by Uusitupa et al. (2013) for an *intention-to-treat* analysis of SYSDIET was applied but PC1 and PC2 scores were used as independent variables instead of the intervention groups to provide a *per-protocol* analysis considering the dietary patterns of HND and CD. All clinical outcomes were corrected for baseline levels to reflect their change as a consequence of the dietary exposures. The subject was included as a random effect with body weight, age, gender, study center, study group, and visit as covariates. Analyses of systolic and diastolic blood pressures included antihypertensive treatment as an additional covariate, and analyses of lipids and inflammation markers included statin usage as an additional covariate. To account of the multiple testing issues, *P* values were adjusted for the false discovery rate, provided as q-values. This statistical analysis was applied separately on the set of urine and plasma features.

To perform group-based comparisons (HND *vs.* CD), we performed PLS-DA with the identical metabolite measurements used to calculate PCA scores. The subjects who were misclassified by PLS-DA are considered as 'non-compliant' and are removed. The group-based analysis of the outcomes was repeated with the remaining samples.

Table 1
Baseline characteristics of participants within control and healthy Nordic diet (HND) intervention groups with metabolomics measurements.

	CD (n = 71)	HND (n = 98)	*P* value
Age	54.7 ± 8.48	54.5 ± 8.09	0.86
Sex (female, n%)	43 (61%)	68 (69%)	0.30[1]
BMI, kg/m^2	31.8 ± 2.83	31.6 ± 3.50	0.73
0 h Glucose, mmol/L	5.7 ± 0.69	5.7 ± 0.60	0.69
2 h Glucose, mmol/L	6.7 ± 2.1	6 ± 1.47	0.03
HDL cholesterol, mmol/L	1.4 ± 0.43	1.4 ± 0.33	0.84
LDL cholesterol, mmol/L	3.2 ± 0.92	3.3 ± 0.79	0.52
Total cholesterol, mmol/L	5.2 ± 1	5.3 ± 0.88	0.44
Triglycerides, mmol/L	1.5 ± 0.64	1.5 ± 0.76	0.75
APO A1, g/L	1.4 ± 0.24	1.4 ± 0.21	0.38
APO B, g/L	1 ± 0.26	1.1 ± 0.26	0.44
Diastolic BP, mmHg	80.1 ± 10.93	82.2 ± 10.51	0.21
Systolic BP, mmHg	129.5 ± 15.92	130.3 ± 14.97	0.73
Smoking, n	5 (7%)	4 (4.1%)	0.62[1]
Statin use, n	20 (28.2%)	20 (20.4%)	0.32[1]
Antihypertensive use, n	34 (47.9%)	59 (60.2%)	0.15[1]
Metabolic Syndrome, n	55 (77.5%)	73 (74.5)	0.79[1]

Values are given as mean ± standard deviation for continuous measurements. *P* values based on unpaired Student *t* test, CD (control diet) *vs.* HND (healthy Nordic diet), [1]Values are given as the total numbers for categorical variables n (%) and *P* values are based on Pearson χ^2 test.
Abbreviations: APO, apolipoprotein; BP, blood pressure.

3. Results

3.1. Metabolic profiles reflecting HND vs. CD dietary patterns

Baseline characteristics of the participants included in this study (n = 169) are shown in Table 1 and a comparison with the baseline characteristics from the intention-to-treat analysis is shown in Supplementary Table 3. Of the included subjects 166 delivered blood samples and 159 delivered urine samples for analysis (Supplementary Table 2) at either week 12 or week 18/24. After the data preprocessing steps, the LC–MS profiles of urine samples resulted in a samples-by-features matrix of 313 × 7090 (only including sampling at 12 weeks (154 samples) and 18–24 weeks (159 samples)) whereas the size of baseline subtracted plasma profiles was 323 × 3643 (160 matched samples at 12 weeks and 163 at 18/24 weeks), both sets with merged positive and negative mode data. The description of data analysis workflow and principal findings is given in Supplementary Fig. 3. Initially, to investigate whether the effect of intervention on the metabolome is similar at 12 weeks and at 18–24 weeks, we performed separate PLS-DA for the samples at each of these time points and for both combined. PLS-DA models performed the best when 12 week and 18–24 weeks were pooled (Supplementary Table 4) with average test set AUC of 0.96 ± 0.03 and 0.93 ± 0.02 (based on 100 different training/test set pairs) for plasma and urine, respectively. Furthermore, the selected metabolites from the pooled models covered >90% of the metabolites selected by the 12 week and 18–24 weeks therefore we continued further analysis with the pooled time points. The PLS-DA model scores plots of the two first components and AUC built with the selected variables from this pooled analysis using an average performing training/test set pair are illustrated in Supplementary Fig. 4.

The variable selection procedure resulted in 67 and 81 metabolites for plasma and urine, respectively (Supplementary Table 5). Among these compounds, the list of annotated discriminating

Table 2
Annotated plasma and urine metabolites discriminating the healthy Nordic and Control diets.

Suggested compound	Molecular formula[a]	Rt. time	Observed *m/z*	Annotation	Fragments (MS2 or MS1)	Diet[b]	Biological fluid	Suggested food source	Centers with P < 0.05[c]
CMPF(1)[d]	$C_{12}H_{16}O_5$	4.04	239.091	$[M-H]^-$	221.1,195.1,151.1,177.1,135.1,123.1	HND	Plasma Urine	Fish	AOKLR OAKLUR
CMHPF(2)	$C_{12}H_{16}O_6$	3.54	255.095	$[M-H]^-$	237.1,211.1,193.1,167.1,139.1,153.05, 109.1,107.1	HND	Plasma	Fish	OKLA
CPF(2)	$C_{11}H_{14}O_5$	3.88	225.075	$[M-H]^-$	121.1,135.1,137.1, 163.1,181.1,207.1, 107.1	HND	Plasma	Fish	KOALU
CMPentylF(2)	$C_{14}H_{20}O_5$	4.36	267.121	$[M-H]^-$	209.1,223.1	HND	Plasma	Fish	OA
CMPF glucuronide(1)	$C_{18}H_{24}O_{11}$	3.75	415.125	$[M-H]^-$	239.1,175,195.1,151.1, 113,85	HND	Urine	Fish	OLKRUA
AR(C21:0) glucuronide(1)	$C_{33}H_{56}O_8$	5.03	579.389	$[M-H]^-$	561.5,505.3,403,75, 113, 95,85	HND	Plasma	Whole grain	OLAK
AR(C21:1) glucuronide(2)	$C_{33}H_{54}O_8$	4.92	577.375	$[M-H]^-$	508.4, 01.3,113,175, 95,85	HND	Plasma	Whole grain	OLKA
AR(C19:0) glucuronide(1)	$C_{31}H_{52}O_8$	4.92	551.360	$[M-H]^-$	508.3,492.3,113,175, 95,85	HND	Plasma	Whole grain	OLKA
3,5-DHPPA glucuronide(1)	$C_{15}H_{18}O_{10}$	1.98	357.09	$[M-H]^-$	181.0175,113	HND	Urine	Whole grain	OAKLU
3,5-DHBA glucuronide(2)	$C_{13}H_{14}O_{10}$	0.93	329.051	$[M-H]^-$	329.0175,113,109, 65	HND	Urine	Whole grain	OKA
3,5-DHPPA(1)	$C_9H_{10}O_4$	2.66	181.041	$[M-H]^-$	137.0,95.0	HND	Urine	Whole grain	OAKLU
3,5-DHBA(1)	$C_7H_6O_4$	1.94	153.018	$[M-H]^-$	109.0,65.0	HND	Urine	Whole grain	OAULK
DHBA sulphate(2)	$C_7H_6O_7S$	2.91	232.977	$[M-H]^-$	134.1,138.0,153.0 (DHBA),191.1	HND	Urine	Whole grain	LOKA
AR metabolite sulphate(3)	$C_8H_8O_8S$	0.78	262.9874	$[M-H]^-$	97.0,139.0,111.0, 183.0	HND	Urine	Whole grain	LOKA
UI conjugate of DHBA glucuronide(3)	$C_{11}H_{12}O_7$	1.19	511.0971	$[M-H]^-$	113.0,153.0,175.0, 181.1(DHPPA),197.1,219.0,221.0, 329.1 (DHBA glu)	HND	Urine	Whole grain	OKLUA
3,5-DHBA glycine(3)	$C_9H_9O_5N$	1.63	210.041	$[M-H]^-$		HND	Urine	Whole grain	OKAURL
CN(C10:3)(3)	$C_{17}H_{28}NO_4$	3.67	310.202	$[M+H]^+$	85.0,251.1,121.1,79.0, 93.1,149.1	HND	Plasma	Unknown	KLUO
CN(C8:1)(1)	$C_{15}H_{29}NO_4$	3.57	286.202	$[M+H]^+$	227.1,85,143.1,125.1, 79,60.1	HND	Plasma	Unknown	LKOUA
CN(C13:0)(3)	$C_{20}H_{39}NO_4$	4.29	358.297	$[M+H]^+$	85,299.2,60.1,197.2,144.1	HND	Plasma	Fish	KALO
CN(C13:1)(3)	$C_{20}H_{37}NO_4$	4.21	356.281	$[M+H]^+$	85,297.2,60.1,97.1,195.2, 177.2	HND	Plasma	Fish	A
DHA(1)	$C_{22}H_{32}O_2$	4.78	327.220	$[M-H]^-$		HND	Plasma	Fish	OA
EPA(1)	$C_{20}H_{30}O_2$	4.75	301.210	$[M-H]^-$		HND	Plasma	Fish	OAKL
LPC(22:6)(1)	$C_{30}H_{50}NO_7P$	4.62	568.339	$[M+H]^+$		HND	Plasma	Fish	A
			590.323	$[M+Na]^+$					
			612.331	$[M+FA]^-$					
LPC(20:5)(1)	$C_{28}H_{48}NO_7P$	4.57	542.326	$[M+H]^+$		HND	Plasma	Fish	AKOLU
			564.307	$[M+Na]^+$					
			586.315	$[M+FA]^-$					
			654.303	$[M+FA+HCOONa]^-$					
LPC(16:1)(1)	$C_{24}H_{48}NO_7P$	4.61	494.319	$[M+H]^+$		CD	Plasma	Unknown	KL
			516.307	$[M+Na]^+$					
			538.315	$[M+Fa]^-$					
			606.303	$[M+FA+HCOONa]^-$					
PC(36:5)	$C_{44}H_{78}NO_8P$	5.18	824.546	$[M+FA]^-$	255.2(C16:0), 301.2(EPA), 764.5(M-FA-CH_3), 526.3(M-C16:0-FA-CH_3), 480.3(M-C20:5-FA-CH_3), 526.3(M-C16:0-FA-CH_3)	HND	Plasma	Fish	OKAL
PC(16:0/20:5 or 20:5/16:0)(1)			802.534	$[M+Na]^+$					
PC(P40:7) PE(22:6/P18:1) or PE(P18:1/22:6)(1)	$C_{45}H_{76}NO_7P$	5.40	772.530	$[M-H]^-$	755.3(M-NH_2), 462.3(LPE(P18:1)), 327.3(C22:6), 444.3(M-22:6), 283.2(C22:6-CO_2), 140($PO_4C_2H_7N$)	HND	Plasma	Fish	AKLUO
SM(34:2)(1)	$C_{39}H_{77}N_2O_6P$	5.20	745.551	$[M+FA]^-$		CD	Plasma	Unknown	KOU
PC(40:6)	$C_{48}H_{84}NO_8P$	5.46	878.594	$[M+FA]^-$	283.2(C18:0), 327.2(C22:6), 508.3(M-C22:6-FA-CH_3), 818.6(M-FA-CH_3)	HND	Plasma	Fish	OAKU
PC(18:0/22:6 or 22:6/18:0)(1)			946.582	$[M+FA+HCOONa]^-$					
PC(40:4) PC(18:0/22:4 or 22:4/18:0)(1)	$C_{48}H_{88}NO_8P$	5.81	882.622	$[M+FA]^-$	283.3(C18:0), 331.3(C22:4),490.3(M-C22:4-FA-CH_3), 538.3(M-C18:0-FA-CH_3)	CD	Plasma	Unknown	KOARU
PC(32:1) PC(16:0/16:1 or 16:1/16:0)(2)	$C_{40}H_{78}NO_8P$	5.34	776.546	$[M+FA]^-$	253.2 (C16:1), 255.2(C16:0),716.5(M-FA-CH_3), 508.3(M-C22:6-FA-CH_3)	CD	Plasma	Unknown	LKU
PC(P36:5)(2)	$C_{44}H_{78}NO_7P$	5.26	808.552	$[M+FA]^-$		HND	Plasma	Unknown	OK
PC(P38:6)(2)	$C_{46}H_{80}NO_7P$	5.31	834.566	$[M+FA]^-$		HND	Plasma	Unknown	OAK
UI glucuronide(3)		3.80	399.131	$[M-H]^-$		HND	Urine	Berry	OKLRA
UI(4)		3.78	221.189	$[M+H]^+$		HND	Urine	Berry	OKLR

Abbreviations: O: Oulu, K: Kuopio, R: Reykjavik, L: Lund, U: Uppsala, A: Aarhus. AR, Alkylresorcinol; CMPF, 3-Carboxy-4-methyl-5-propyl-2-furanpropanoic acid; CMHPF, 3-Carboxy-4-methyl-5-hydroxypropyl-2-furanpropanoic acid; CN(Cn:m), acylcarnitine; CPF, 3-Carboxy-5-propyl-2-furanpropanoic acid; CMPentylF, 3-Carboxy-4-methyl-5-pentyl-2-furanpropanoic acid; DHA, docosahexaenoic acid; DHBA, 3,5-dihydroxybenzoic acid; DHPPA, 3,5-dihydroxy phenylpropionic acid; EPA, eicosapentaenoic acid; LPC(Cn:m), lysophosphatidylcholine; MS: mass spectrometry; PE(Cn:m), Phosphatidylethanolamine; PC(Cn:m), phosphatidylcholine; SM(Cn:m), Sphingomyelin; UI; unidentified, Brackets (Cn:m) include number of carbons, n, and number of double bonds, m; glu, glucuronide.

[a] Molecular formula corresponds to the neutral compound.

[b] The diet column shows which diet increases the metabolite.

[c] The centers providing a significant difference between the diet groups ($P < 0.05$) were sorted in ascending order according to P-value. P-values were calculated with a nonparametric Mann–Whitney test.

[d] Level of identification according to Sumner et al. [22] is given in the brackets.

metabolites is provided in Table 2. Plasma and urine HND metabolic patterns reflected the contrasting food components in HND compared with CD, *i.e.* whole grain, fish, berries and fats. HND metabolites associated with whole grain intake were alkylresorcinol (AR(C19:0), AR(C21:0) and AR(C21:1)) glucuronides in plasma whereas their metabolites (3,5-DHBA, 3,5-DHPPA and their sulphate, glucuronide and glycine conjugates) were observed in urine.

Furan fatty acid metabolites (CMPF, CMPF glucuronide, CPF and CMHPF) and plasma lipid profiles *e.g.* docosahexaenoic acid (DHA; n-3 C22:6), eicosapentaenoic acid (EPA; n-3 C20:5), LPC(20:5), LPC(22:6), PC(18:0/22:6) and PC(16:0/20:5) characterized fish consumption (Table 2). The HND also associated with four specific acyl carnitines (CNCs), *i.e.* CN(C8:1), CN(C10:3), CN(C13:0) and CN(C13:1) whereas CD lead to higher levels of plasma LPC(16:1), PC(32:1), PC(40:4) and SM(34:2). Berry intake for HND was characterized by two unknown metabolites associated with recent strawberry intake in a previous meal study [24].

The associations between urine and plasma metabolites reflecting intervention were shown in Fig. 1. The two major clusters explained the differences between the diet groups with metabolites related to Nordic signature foods separating HND from the phospholipids signifying CD. HND dietary pattern provides subclusters, one of which is comprised of various lipids, including plasmalogens containing C20:5 and C22:6 fatty acyls as well as

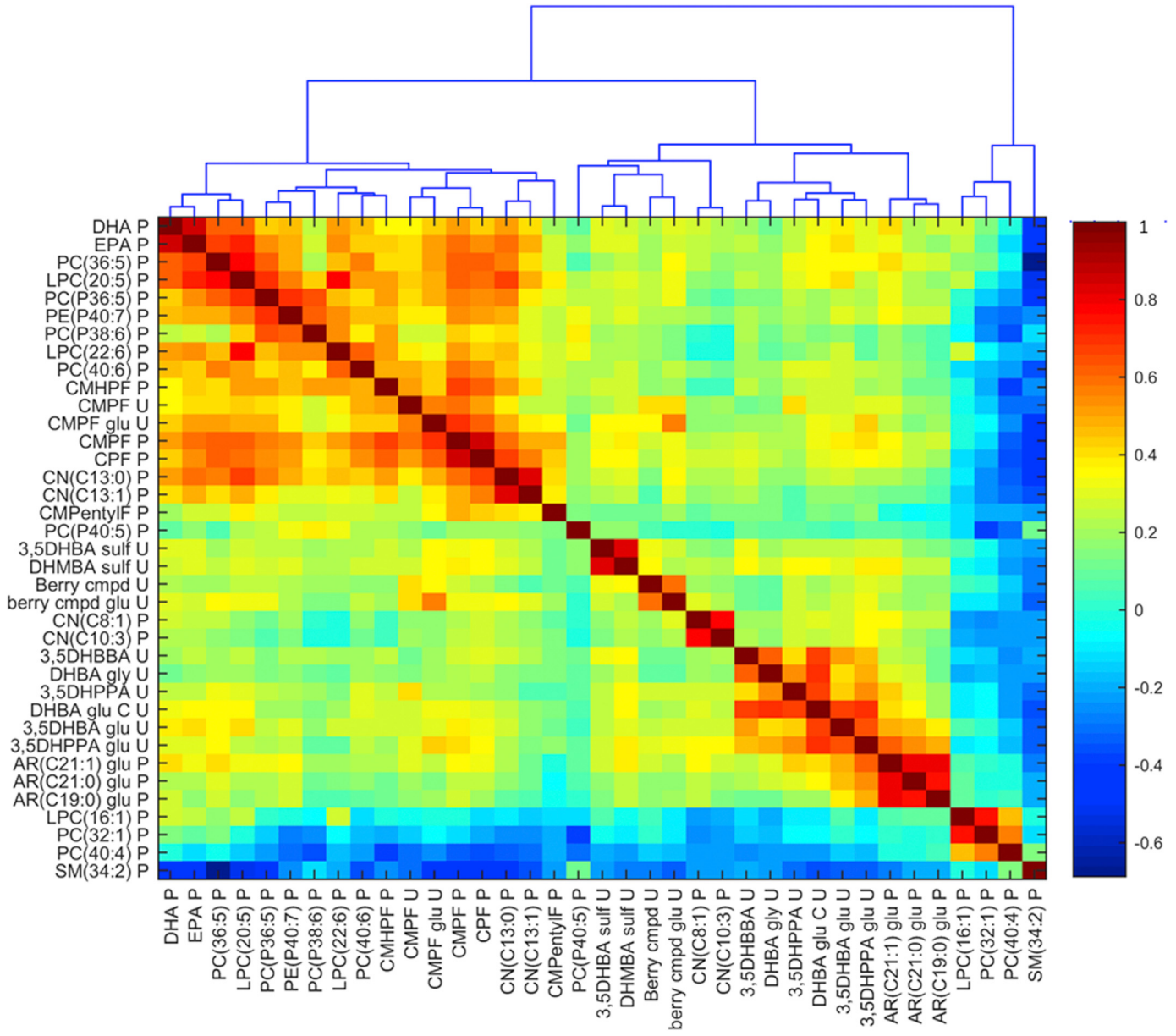

Fig. 1. Heatmap describing the relationship between the identified blood (P) and urine (U) metabolites reflecting the Healthy Nordic Diet or control diet based on Spearman rank correlation. The order of blood and urine metabolites is defined by hierarchical cluster analysis using the Ward method. Colour legend provides the correlation coefficients corresponding to the colours in the heatmap. Abbreviations: AR, Alkylresorcinol; CMPF, 3-Carboxy-4-methyl-5-propyl-2-furanpropanoic acid; CMHPF, 3-Carboxy-4-methyl-5-hydroxypropyl-2-furanpropanoic acid; CN(Cn:m), acylcarnitine; CPF, 3-Carboxy-5-propyl-2-furanpropanoic acid; CMPentylF, 3-Carboxy-4-methyl-5-pentyl-2-furanpropanoic acid; DHA, docosahexaenoic acid; DHBA, 3,5-dihydroxybenzoic acid; DHPPA, 3,5-dihydroxyphenylpropionic acid; EPA, eicosapentaenoic acid; LPC(Cn:m), lysophosphatidylcholine; PE(Cn:m), Phosphatidylethanolamine; PC(Cn:m), phosphatidylcholine; SM(Cn:m), Sphingomyelin; Brackets (Cn:m) provides the number of carbons, n, and the number of double bonds, m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

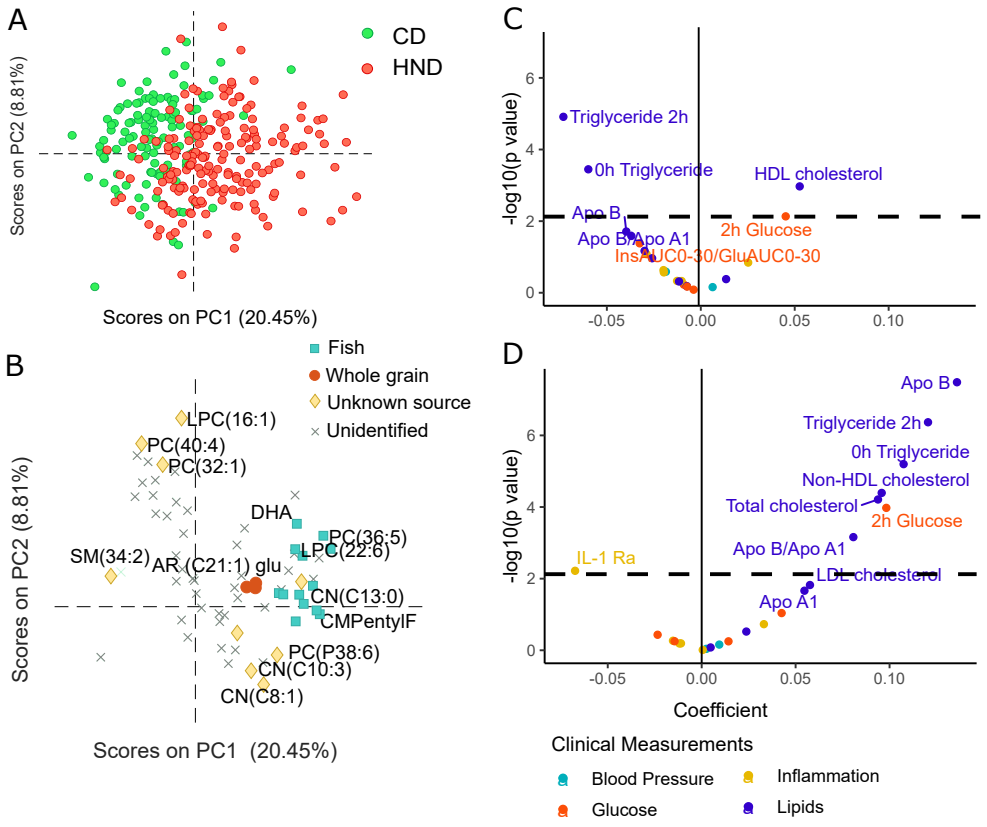

Fig. 2. PCA scores (A) and loadings (B) plot of 67 baseline corrected changes in plasma metabolites, reflecting the contrast between dietary intervention groups, control diet (CD) or Healthy Nordic Diet (HND). Sample measurements at 12 and 18–24 weeks during the intervention with CD or HND were concatenated for this analysis. (C) and (D) Volcano plot representing the relationship between PC1- and PC2-scores with clinical measurement outcomes, respectively. The *P* values and estimates were calculated using a linear mixed model. The clinical measurements that had *P* values <0.05 are annotated and the dashed line represents a q-value of 0.1.AR, Alkylresorcinol; CN(Cn:m), acylcarnitine; CMPentylF, 3-Carboxy-4-methyl-5-pentyl-2-furanpropanoic acid; DHA, docosahexaenoic acid; LPC(Cn:m), lysophosphatidylcholine; PC(Cn:m), phosphatidylcholine; SM(Cn:m), Sphingomyelin; Brackets (Cn:m) include number of carbons, n, and number of double bonds, m.

furan fatty acid metabolites, likely reflecting fish intake. Plasma AR glucuronides are strongly correlated with each other ($r > 0.7$) and cluster together with most of their metabolites in urine (DHBA, DHPPA and their conjugates). Carnitines are divided into two groups in terms of their correlation; CN(C13:0) and CN(C13:1) correlate strongly with each other and also with several other metabolites, in particular CN(C13:0) with fish-related lipids. Other carnitines, CN(C8:1) and CN(C10:3) are only associated with each other but are closer to the whole grain cluster than to fish. Some lipids reflecting CD, *e.g.* PC(40:4), LPC(16:1) and PC(32:1), correlate with each other, whereas sphingomyelin SM(34:2) is negatively associated with most other metabolites, including lipids characterizing HND.

We investigated whether the differences between food cultures between the centers has an influence on the selected metabolites. PCA of combined urine and plasma profiles explains well the differences between the diet groups and contributing metabolites (Supplementary Fig. 5). However, when the same scores plot is annotated with diet and center (Supplementary Fig. 6), HND samples from specific centers are clustered closer. The Finnish centers (Kuopio and Oulu) were characterized by higher whole grain and berry-associated compounds whereas fish intake was more pronounced for the Danish center (Aarhus). For each metabolite in Table 2, the centers providing a significant difference between HND *vs.* CD are shown in Supplementary Table 6. Interestingly, only three compounds, *i.e.*, DHBA-glycine, CMPF and CMPF glucuronide were different between HND and CD at all six centers while most others showed a trend. However, the sample size differed between the centers, hence these results should be interpreted with caution.

3.2. Associations between clinical measurements and PCA diet scores

Separate PCA-analyses were made for changes in plasma metabolites from baseline and for the urine metabolites after intervention. The plasma and urine metabolic profiles reflecting the intervention diets were tested as predictors of the major clinical endpoints using PC1 and PC2 scores from each PCA (dietary pattern) analysis.

The 67 contrasting plasma metabolites decomposed into two PCA components (Fig. 2A and B). Larger PC1 and lower PC2 scores are associated with HND, whereas CD is represented by mostly larger PC2 and lower PC1 scores (Fig. 2A). The relationship between PC1 and PC2 scores and clinical measurements are shown in Table 3, while volcano plots, Fig. 2C and D, display the associations. PC1 explains mainly the metabolites originating from the HND, *e.g.*, ARs, furan fatty acid metabolites, lipids containing n-3 fatty acids (Fig. 2B) and is positively related with HDL and negatively with fasting triglycerides (Fig. 2C).

As shown in Fig. 2B, PC2 for changes in the plasma metabolite pattern is mainly described by contrasting effects of PC(40:4), PC(32:1) and PC(16:1) with higher scores representing CD, and the

Table 3
Associations between clinical measurements and plasma metabolites explained by PCA scores. Clinical measurements and metabolites levels were corrected for baseline measurements.

	PC1 scores			PC2 scores		
	Coefficient estimate [95% CI]	*P* value	q value	Coefficient estimate [95% CI]	*P* value	q value
0 h Glucose, mmol/L	−0.004 [−0.04, 0.03]	0.82	1	0.01 [−0.03, 0.06]	0.56	1
2 h Glucose, mmol/L[1]	0.05 [0.01, 0.08]	0.007	0.08	0.1 [0.05, 0.15]	<0.001	0.001
HOMA-IR	−0.009 [−0.04, 0.02]	0.59	1	0.04 [−0.01, 0.09]	0.09	1
Matsuda ISI	−0.007 [−0.04, 0.03]	0.67	1	−0.02 [−0.07, 0.03]	0.37	1
InsAUC0-30/GluAUC0-30	−0.03 [−0.06, 0]	0.04	0.48	−0.01 [−0.06, 0.03]	0.56	1
HMW adiponectin, μg/L	−0.02 [−0.05, 0.01]	0.23	1	−0.01 [−0.06, 0.04]	0.64	1
IL-10, ng/L	−0.02 [−0.05, 0.02]	0.27	1	−0.01 [−0.06, 0.04]	0.65	1
IL-6, ng/L	−0.01 [−0.04, 0.02]	0.47	1	0.001 [−0.04, 0.04]	0.98	1
IL-1 beta, ng/L	0.03 [−0.01, 0.06]	0.14	1	−0.02 [−0.06, 0.03]	0.54	1
IL-1 Ra, ng/L[2]	−0.03 [−0.06, 0]	0.08	0.96	−0.07 [−0.12, −0.02]	0.006	0.07
hsCRP, mg/L	−0.01 [−0.05, 0.02]	0.47	1	0.03 [−0.02, 0.08]	0.19	1
APO A1/APO B, g/L[1]	−0.04 [−0.07, 0]	0.03	0.3	0.08 [0.03, 0.13]	0.001	0.008
APO A1, g/L[2]	0.01 [−0.02, 0.05]	0.42	1	0.05 [0.01, 0.1]	0.02	0.25
APO B, g/L[1]	−0.04 [−0.07, −0.01]	0.02	0.22	0.14 [0.09, 0.18]	<0.001	<0.001
HDL cholesterol, mmol/L[1]	0.05 [0.02, 0.08]	0.001	0.01	0.005 [−0.04, 0.05]	0.83	1
LDL-C/HDL-C ratio	−0.03 [−0.06, 0.01]	0.11	1	0.02 [−0.02, 0.07]	0.3	1
LDL cholesterol, mmol/L[2]	−0.007 [−0.04, 0.02]	0.65	1	0.06 [0.01, 0.1]	0.02	0.17
Non-HDL cholesterol, mmol/L[1]	−0.03 [−0.06, 0]	0.07	0.78	0.1 [0.05, 0.14]	<0.001	<0.001
Total cholesterol, mmol/L[1]	−0.01 [−0.04, 0.02]	0.48	1	0.09 [0.05, 0.14]	<0.001	0.001
Triglycerides, fasting, mmol/L[1]	−0.06 [−0.09, −0.03]	<0.001	0.004	0.11 [0.06, 0.15]	<0.001	<0.001
Triglycerides, 2 h, mmol/L[1]	−0.07 [−0.11, −0.04]	<0.001	<0.001	0.12 [0.07, 0.17]	<0.001	<0.001
Diastolic BP, mmHg	−0.02 [−0.05, 0.01]	0.26	1	0.009 [−0.04, 0.06]	0.69	1
Systolic BP, mmHg	0.006 [−0.03, 0.04]	0.7	1	0.003 [−0.04, 0.05]	0.91	1

Estimates are changes from baseline values (see Table 1).
P values are calculated by linear mixed model adjusted for random effects body weight, age, gender, study center and visit as covariates. Systolic and diastolic blood pressures included antihypertensive treatment as an additional covariate, and analyses of lipids and inflammation markers included statin usage as an additional covariate.
P values were adjusted for the multiple comparisons by using false discovery rate which are provided as q-values. [1]q < 0.05, [2]P < 0.05.
Abbreviations: APO, apolipoprotein; BP, blood pressure; hsCRP, high-sensitivity C-reactive protein; HMW, high-molecular weight; Ins, insulin; ISI, Insulin sensitivity index; Glu, glucose.

medium chain unsaturated carnitines, CN(C10:3) and CN(C8:1), and a polyunsaturated plasmalogen, PC(P38:6), reflecting HND by lower PC2 scores. A marker of glucose metabolism, *i.e.* 2 h glucose, is associated with the PC2 scores (Fig. 2A and D). Positive relationships with PC2 scores are found for Apo B, non-HDL cholesterol, Apo B/APO A1, total cholesterol and plasma triglycerides, while PC2 has a borderline negative effect on IL1 Ra.

Our results suggest more pronounced associations between intervention reflected by PC scores and clinical outcomes compared to the previously published intention-to-treat analysis [10] To investigate whether this difference is resulting from the data analysis strategy, we repeated the same analysis by replacing PC scores with the intervention groups. The associations were similar to intention-to-treat analysis [10] indicating the intervention reflected by PC scores predicts better the changes in the clinical outcomes (Supplementary Table 7). An additional group-based analysis after removing all participants that did not classify into their assigned group (n = 16), did not performed better than the PC scores (Supplementary Table 7). Removing participants with weight change did not affect the findings (data not shown).

The relationships of PC1 and PC2 scores from PCA of urine metabolites measured at 12 and 18–24 weeks (Supplementary Fig. 7) with clinical measurements was much less pronounced as shown in Supplementary Table 8 and only LDL cholesterol was associated with PC1 scores.

4. Discussion

In the randomized, controlled SYSDIET intervention study, consumption of HND compared to CD under isocaloric conditions gave distinct plasma and urine metabolic profiles characterized by the consumption of signature HND food products. More importantly, 2 h glucose response, lipoproteins, triglycerides and low-grade inflammation are decreased by plasma metabolite scores, reflecting mostly contrasts of plasma dietary fat metabolites in HND versus CD. An improved lipid profile is also associated with higher levels of metabolites well-known to reflect intake of fish [30], berries [24] and whole grains [19]; these are mainly metabolites, increased after HND.

In this study, the majority of identified metabolites reflected higher whole grain *i.e.* glucuronides and beta-oxidation products of AR, or fish intake *i.e.* (a) furan fatty acid metabolites and their conjugates and (b) n−3 fatty acids, in free form or incorporated into LPCs, PCs, and plasmalogens in HND. Both ARs and n−3 fatty acids were previously applied as biomarkers of whole grain [19,31] and fish intake [30], respectively, while CMPF recently emerging as a potential biomarker of fatty fish consumption, at least in northern Europe [32–34]. We identified additional furan fatty acid metabolites, *i.e.* CPF and CMHPF, which were correlated with CMPF, suggesting similar dietary sources for these compounds; however, further studies are needed to elucidate whether CMHPF is a human phase I metabolite of CMPF or possibly reflects a distinct aquatic food [35].

Investigation of metabolite patterns in differently located centers showed moderate heterogeneity, reflecting differences in the selected foods within the framework of the general Nordic dietary recommendations. For instance, ARs and their metabolites did not differ between the two diet groups in Iceland except for DHBA glycine. This is consistent with the fibre intake from whole grain reported previously [10] since barley provided to participants in Iceland is lower in AR [36] than rye provided at the other centers. The fish intake biomarkers, plasma EPA and DHA, were not different between diet groups within all centers. EPA and DHA are very prone to oxidation [37], therefore differences in cooking procedures [38] and/or seasonal variations [39] may play a role. The levels of DHA and EPA in fish also differ according to species;

freshwater fish or lean fish, consumed *e.g.* in Finland and central Sweden have lower polyunsaturated fatty acids (PUFA) levels [40] and where leaner fish from the sea were preferred, *e.g.* in Iceland, levels would also be expected to be lower. Lack of compliance is not a very likely explanation, since urinary CMPF was associated with HND at all centers. The HND was also associated with the odd-chain acyl-carnitines, CN(C13:1) and CN(C13:3), which were also correlated with EPA and DHA. Furthermore, PUFA containing PCs and plasmalogens were higher for HND, possibly related to an increased dietary intake of n-3 and n-6 PUFAs consistent with previous findings [41]. Plasmalogens negatively associate with metabolic syndrome and type 2 diabetes [42], suggesting healthier lipid profiles after HND. CD was reflected by higher amounts of C16:1 containing phospholipids. Palmitoleic acid was not increased overall in plasma lipids after the CD [41], suggesting that C16:1 containing PCs may reflect relatively recent intake of terrestrial fats.

In this multi-center study, metabolites associated with intake of recommended foods in the HND, such as fish, whole grain and berries, could be viewed as a reflection of individual compliance [19]. The Nordic Nutrition Recommendations cover all the Nordic countries although their traditional food cultures differ somewhat with additional regional variations. SYSDIET therefore represents a realistic representation of how a set of national or international dietary recommendations are interpreted, as reflected in the metabolic profiles from different regions. Potential differences in genetics or in gut microbiota between centers might explain part of the heterogeneity, yet none of the identified metabolites were of microbial origin. Due to this heterogeneity, the pattern of dietary metabolites should represent a better score for compliance than single food intake biomarkers.

The previously published intention-to-treat analysis demonstrated that HND improved the lipid profile, while inflammation markers increased in CD; however, by intention-to-treat analysis the trial did not influence the primary outcome, glucose metabolism [10]. The limited impact of HND on health-related outcomes could be due to lack of compliance and/or to the variation introduced by the multi-center design. Indeed, using the six relevant biomarkers of intake known in 2014, the most compliant individuals had a slightly improved effect of the HND on their lipid profile, but glucose metabolism was not affected [19]. In this current study, we combined 67 diet-discriminating plasma metabolites, accounting for both compliance and differences between the centers, into two PCA components. Interestingly, our findings demonstrated that the component described by consumed dietary fat metabolites (PC2) was related with a marker of glucose metabolism, *i.e.* 2 h glucose. This suggests that specific lipid components of HND, explained by contribution of PCs containing C22:6 and medium-chain unsaturated carnitines, possibly led to an improved glucose tolerance. The medium- and even-chain acyl carnitines, CN(C8:1) and CN(C10:3), reflecting HND may result from some non-aquatic HND fat source, potentially rapeseed oil alpha-linoleate, but further studies are needed to identify their source. HND was also reflecting an improved cholesterol profile by both orthogonal PCA components. PC1, composed mainly of polar metabolites originating from whole grain, fish and berries, is directly related to increased HDL, whereas PC2 explains decreases in LDL and non-HDL cholesterol. Polar urine metabolites also explain a lipoprotein, in this case LDL. In accordance with our findings a trial of Nordic foods where saturated fat was replaced by mainly n-6 PUFA led to improved cholesterol metabolism [43]. Furthermore, serum triglycerides were explained by both components, suggesting that not only intake of fish, berries and whole grain in HND, but also altered sources of dietary fats has an impact on markers associated with cardiometabolic risk.

A weakness may be that untargeted metabolomics is explorative and limited by the analytical platform, so additional metabolites might be found by other profiling techniques. In the current study, the limited number of samples from some centres attenuates the power to observe detailed centre differences. In addition, several of the discriminative metabolites were unidentified, yet identified metabolites (level 1–2) represented all signature food groups in HND.

Hereby, the novel approach aims per-protocol analysis; it handles post-randomization factors by the repeated sampling and measurement of diet-related metabolites, potential selection bias by comparing baseline of the completers (Supplementary Table 3), and also the exclusion restriction criterion [44], *i.e.* that the outcome is mediated solely by the intervention itself. This latter criterion is a consequence of using scores based on food metabolites for analysis of the outcome. An additional group-based analysis after removal of non-complaints showed weaker association compared to the scores. This is not surprising since, (1) we decrease the power by reducing the number of samples and (2) we lack the information on “the level of compliance within the groups” obtained from the scores based on per-protocol analysis. Particulary in longer term dietary trials, like SYSDIET, it is hard to achieve the same level of intake for each participant. Therefore, our proposed methodology should be better than group based per-protocol analysis. Still, there are also some remaining issues; this per-protocol analytical strategy was not preplanned in the trial protocol although biomarker analysis was planned; not all Nordic foods or refined foods are covered adequately in a balanced way by our approach and this should ideally be achieved for a per-protocol analysis of a diet; moreover, some of the unknown biomarkers might be a consequence of biological response and this could cause confounding.

Untargeted metabolomics allowed discriminant analysis discerning the two diets; food-group related biomarkers were clearly mapped with the key food sources, including fish, whole grain, berries and unsaturated fats. This approach provides a practical tool to monitor participants' adherence to a healthy diet such as HND [45]; our analysis therefore offers a novel way approaching *per protocol* analyses with a balanced compliance assessment in larger dietary trials like SYSDIET. In conclusion, this untargeted metabolomics analysis provides a multivariate compliance measure allowing a novel outcome analysis based on actual food choices within a common set of dietary recommendations across geographical regions. Metabolite scores reflecting the dietary patterns affect the primary and other clinical outcomes in SYSDIET by changes mainly in metabolites of dietary fat but also by changes in polar metabolites of signature foods such as fish, whole grain and berries.

Author contributions

M.U. was the principal investigator of the SYSDIET consortium; M.U., U.S., K.H., M.J.S., J.H., I.T., K.S.P., U.R., and B.A. had the main responsibility for conducting the study in each trial center; US, MK, LB, LC and IG planned the study diets, trained the study personnel and conducted the study; MO had the responsibility of building and managing of the database; MU, KH, MJS, US, MK, LB, LC, K-HH, JH, FR, SMU, IG, IT, MO, KSP, UR, BA, LOD substantially contributed to the study design, conducting the study. LOD and GG were responsible for metabolomics analysis; GG performed the statistical analysis and metabolite identification; GG and LOD wrote the manuscript. All authors have read and approved the final version of the manuscript.

Funding

Systems biology in controlled dietary interventions and cohort studies (SYSDIET) is one of the three projects in the Nordic Centre of

Excellence Programme on Food, Nutrition and Health nominated and funded by NordForsk for years 2007–2012 (SYSDIET; 070014). The SYSDIET intervention was also funded by the Academy of Finland, Finnish Diabetes Research Foundation, Finnish Foundation for Cardiovascular Research, The Sigfrid Juselius Foundation, and EVO funding from Kuopio University Hospital (Finland); the Druvan Foundation, ESPEN, Skåne County Council Research and Development Foundation, The Heart-Lung Foundation, Diabetesfonden, and Foundation Cerealia (Sweden), The Danish Obesity Research Centre (DanORC, www.danorc.dk), The Danish Council for Strategic Research (DairyHealth, BioFunCarb, FoodBAll) (Denmark), The Agricultural Productivity Fund, The Research Fund of the University of Iceland (Iceland). The metabolomics analysis was supported by a Semper Ardens grant on biomarkers from the Carlsberg Foundation to LOD.

Conflict of interest

The authors declare they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clnu.2021.12.031.

References

[1] Deng C, Lu Q, Gong B, Li L, Chang L, Fu L, et al. Stroke and food groups: an overview of systematic reviews and meta-analyses. Public Health Nutr 2018;21:766–76.
[2] Bechthold A, Boeing H, Schwedhelm C, Hoffmann G, Knüppel S, Iqbal K, et al. Food groups and risk of coronary heart disease, stroke and heart failure: a systematic review and dose-response meta-analysis of prospective studies. Crit Rev Food Sci Nutr 2019;59:1071–90.
[3] Schwingshackl L, Hoffmann G, Lampousi A-M, Knüppel S, Iqbal K, Schwedhelm C, et al. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. Eur J Epidemiol 2017;32:363–75.
[4] Kastorini C-M, Milionis HJ, Esposito K, Giugliano D, Goudevenos JA, Panagiotakos DB. The effect of Mediterranean diet on metabolic syndrome and its components. J Am Coll Cardiol 2011;57:1299–313.
[5] Siervo M, Lara J, Chowdhury S, Ashor A, Oggioni C, Mathers JC. Effects of the Dietary Approach to Stop Hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. Br J Nutr 2015;113:1–15.
[6] Nordic Nutrition Recommendations 2012. Integrating nutrition and physical activity. 5th ed. Nord: Nord Counc Minist.; 2014.
[7] Bere E, Brug J. Towards health-promoting and environmentally friendly regional diets – a Nordic example. Public Health Nutr 2009;12:91–6.
[8] Poulsen SK, Due A, Jordy AB, Kiens B, Stark KD, Stender S, et al. Health effect of the New Nordic Diet in adults with increased waist circumference: a 6-mo randomized controlled trial. Am J Clin Nutr 2014;99:35–45.
[9] Adamsson V, Reumark A, Fredriksson I-B, Hammarström E, Vessby B, Johansson G, et al. Effects of a healthy Nordic diet on cardiovascular risk factors in hypercholesterolaemic subjects: a randomized controlled trial (NORDIET). J Intern Med 2011;269:150–9.
[10] Uusitupa M, Hermansen K, Savolainen MJ, Schwab U, Kolehmainen M, Brader L, et al. Effects of an isocaloric healthy Nordic diet on insulin sensitivity, lipid profile and inflammation markers in metabolic syndrome - a randomized study (SYSDIET). J Intern Med 2013;274:52–66.
[11] de Mello VDF, Schwab U, Kolehmainen M, Koenig W, Siloaho M, Poutanen K, et al. A diet high in fatty fish, bilberries and wholegrain products improves markers of endothelial function and inflammation in individuals with impaired glucose metabolism in a randomised controlled trial: the Sysdimet study. Diabetologia 2011;54:2755–67.
[12] Gunge VB, Andersen I, Kyrø C, Hansen CP, Dahm CC, Christensen J, et al. Adherence to a healthy Nordic food index and risk of myocardial infarction in middle-aged Danes: the diet, cancer and health cohort study. Eur J Clin Nutr 2017;71:652–8.
[13] Lacoppidan S, Kyrø C, Loft S, Helnæs A, Christensen J, Hansen C, et al. Adherence to a healthy Nordic food index is associated with a lower risk of type-2 diabetes—the Danish diet, cancer and health cohort study. Nutrients 2015;7:8633–44.
[14] Olsen A, Egeberg R, Halkjær J, Christensen J, Overvad K, Tjønneland A. Healthy aspects of the Nordic diet are related to lower total mortality. J Nutr 2011;141:639–44.
[15] Gao Q, Praticò G, Scalbert A, Vergères G, Kolehmainen M, Manach C, et al. A scheme for a flexible classification of dietary and health biomarkers. Genes Nutr 2017;12:34.
[16] Heinzmann SS, Brown IJ, Chan Q, Bictash M, Dumas M-E, Kochhar S, et al. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am J Clin Nutr 2010;92:436–43.
[17] Landberg R, Kamal-Eldin A, Andersson A, Vessby B, Aman P. Alkylresorcinols as biomarkers of whole-grain wheat and rye intake: plasma concentration and intake estimated from dietary records 1-3 [Internet]. 2008. p. 832–40. Available from: https://academic.oup.com/ajcn/article-abstract/87/4/832/4633279.
[18] Andersen M-BS, Rinnan Å, Manach C, Poulsen SK, Pujos-Guillot E, Larsen TM, et al. Untargeted metabolomics as a screening tool for estimating compliance to a dietary pattern. J Proteome Res 2014;13:1405–18.
[19] Marklund M, Magnusdottir OK, Rosqvist F, Cloetens L, Landberg R, Kolehmainen M, et al. A dietary biomarker approach captures compliance and cardiometabolic effects of a healthy Nordic diet in individuals with metabolic syndrome. J Nutr 2014;144:1642–9.
[20] Gürdeniz G, Kristensen M, Skov T, Dragsted LO. The effect of LC-MS data preprocessing methods on the selection of plasma biomarkers in fed vs. fasted rats. Metabolites 2012;2:77–99.
[21] Andersen M-BS, Reinbach HC, Rinnan Å, Barri T, Mithril C, Dragsted LO. Discovery of exposure markers in urine for Brassica-containing meals served with different protein sources by UPLC-qTOF-MS untargeted metabolomics. Metabolomics 2013;9:984–97.
[22] Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 2007;3:211–21.
[23] Gürdeniz G, Rago D, Bendsen NT, Savorani F, Astrup A, Dragsted LO. Effect of trans fatty acid intake on LC-MS and NMR plasma profilesOresic M, editor. PLoS One 2013;8:e69589.
[24] Cuparencu CS, Andersen M-BS, Gürdeniz G, Schou SS, Mortensen MW, Raben A, et al. Identification of urinary biomarkers after consumption of sea buckthorn and strawberry, by untargeted LC–MS metabolomics: a meal study in adult men. Metabolomics 2016;12:31.
[25] Berthou F, Ratanasavanh D, Riche C, Picart D, Voirin T, Guillouzo A, et al. Comparison of caffeine metabolism by slices, microsomes and hepatocyte cultures from adult human liver. Xenobiotica 1989;19:401–17.
[26] Slaughter D, Takenaga N, Lu P, Assang C, Walsh DJ, Arison BH, et al. Metabolism of rofecoxib in vitro using human liver subcellular fractions. Drug Metab Dispos [Internet] 2003:31. Available from: http://dmd.aspetjournals.org/content/dmd/31/11/1398.full.pdf.
[27] Smith Colin A, Want Elizabeth J, O'Maille Grace, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Available from: https://pubs.acs.org/doi/abs/10.1021/ac051437y; 2006.
[28] Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Anal Chem 2012;84:283–9.
[29] Dieterle Frank, Ross Alfred, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR Metabonomics. Anal Chem 2006;78(13):4281–90. https://doi.org/10.1021/ac051632c.
[30] Andersen LF, Solvoll K, Drevon CA. Very-long-chain n–3 fatty acids as biomarkers for intake of fish and n–3 fatty acid concentrates. Am J Clin Nutr 1996;64:305–11.
[31] Magnusdottir OK, Landberg R, Gunnarsdottir I, Cloetens L, Åkesson B, Önning G, et al. Plasma alkylresorcinols reflect important whole-grain components of a healthy Nordic diet. J Nutr 2013;143:1383–90.
[32] Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, et al. Nontargeted metabolite profiling discriminates diet-specific biomarkers for consumption of whole grains, fatty fish, and bilberries in a randomized controlled trial. J Nutr 2015;145:7–17.
[33] Savolainen O, Lind MV, Bergström G, Fagerberg B, Sandberg A-S, Ross A. Biomarkers of food intake and nutrient status are associated with glucose tolerance status and development of type 2 diabetes in older Swedish women. Am J Clin Nutr 2017;106:ajcn152850.
[34] McCullough ML, Maliniak ML, Stevens VL, Carter BD, Hodge RA, Wang Y. Metabolomic markers of healthy dietary patterns in US postmenopausal women. Am J Clin Nutr 2019;109:1439–51.
[35] Xu L, Sinclair AJ, Faiza M, Li D, Han X, Yin H, et al. Furan fatty acids – beneficial or harmful to health? Prog Lipid Res 2017;68:119–37.
[36] Bordiga M, Locatelli M, Travaglia F, Arlorio M, Reyneri A, Blandino M, et al. Alkylresorcinol content in whole grains and pearled fractions of wheat and barley. J Cereal Sci 2016;70:38–46.
[37] Arab-Tehrany E, Jacquot M, Gaiani C, Imran M, Desobry S, Linder M. Beneficial effects and oxidative stability of omega-3 long-chain polyunsaturated fatty acids. Trends Food Sci Technol 2012;25:24–33.
[38] Chung H, Nettleton JA, Lemaitre RN, Barr RG, Tsai MY, Tracy RP, et al. Frequency and type of seafood consumed influence plasma (n-3) fatty acid concentrations. J Nutr 2008;138:2422–7.

[39] Moujahed N, Guizani SEO. Seasonal variation of chemical and fatty acids composition in Atlantic Mackerel from the Tunisian Northern-East Coast. J Food Process Technol 2015;6:1–4.
[40] Cuparencu C, Praticó G, Hemeryck LY, Sri Harsha PSC, Noerman S, Rombouts C, et al. Biomarkers of meat and seafood intake: an extensive literature review. Genes Nutr 2019;14:35.
[41] Lankinen M, Schwab U, Kolehmainen M, Paananen J, Nygren H, Seppänen-Laakso T, et al. A healthy Nordic diet alters the plasma lipidomic profile in adults with features of metabolic syndrome in a multicenter randomized dietary intervention. J Nutr 2015;146:662–72.
[42] Colas R, Sassolas A, Guichardant M, Cugnet-Anceau C, Moret M, Moulin P, et al. LDL from obese patients with the metabolic syndrome show increased lipid peroxidation and activate platelets. Diabetologia 2011;54:2931–40.
[43] Ulven SM, Leder L, Elind E, Ottestad I, Christensen JJ, Telle-Hansen VH, et al. Exchanging a few commercial, regularly consumed food items with improved fat quality reduces total cholesterol and LDL-cholesterol: a double-blind, randomised controlled trial. Br J Nutr 2016;116:1383–93.
[44] Hernán MA, Robins JM. Per-protocol analyses of pragmatic trials. N Engl J Med 2017;377:1391–8.
[45] Renzella J, Townsend N, Jewell J, João B, Roberts N, Rayner M, et al. What national and subnational interventions and policies based on Mediterranean and Nordic diets are recommended or implemented in the WHO European Region, and is there evidence of effectiveness in reducing noncommunicable diseases? (Health Evidence Network). Copenhagen: World Health Organization; 2018.