Joint Analysis of Metabolite Markers of Fish Intake and Persistent Organic Pollutants in Relation to Type 2 Diabetes Risk in Swedish Adults

The Journal of Nutrition
Nutritional Epidemiology

Joint Analysis of Metabolite Markers of Fish Intake and Persistent Organic Pollutants in Relation to Type 2 Diabetes Risk in Swedish Adults

Lin Shi,[1] Carl Brunius,[1] Ingvar A Bergdahl,[2,3] Ingegerd Johansson,[4] Olov Rolandsson,[3] Carolina Donat Vargas,[5] Hannu Kiviranta,[6] Kati Hanhineva,[7,8] Agneta Åkesson,[5] and Rikard Landberg[1,3]

[1]Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden; [2]Department of Biobank Research, [3]Department of Public Health and Clinical Medicine, and [4]Department of Odontology, Umeå University, Umeå, Sweden; [5]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden; [6]Environmental Health Unit, National Institute for Health and Welfare, Kuopio, Finland; [7]LC-MS Metabolomics Center, Kuopio, Finland; and [8]Department of Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

ABSTRACT

Background: There is conflicting evidence regarding the association between fish intake and type 2 diabetes (T2D) incidence, possibly owing to measurement errors in self-reported intake and coexposure to persistent organic pollutants (POPs) present in fish.

Objective: The aim of this study was to identify plasma metabolites associated with fish intake and to assess their association with T2D risk, independently of POPs, in Swedish adults.

Methods: In a case-control study nested in the Swedish Västerbotten Intervention Programme, fasting plasma samples from 421 matched T2D case-control pairs of men and women aged 30–60 y at baseline and 10-y follow-up samples from a subset of 149 pairs were analyzed using untargeted metabolomics. Moreover, 16 plasma POPs were analyzed for the 149 pairs who had repeated samples available. Fish-related plasma metabolites were identified using multivariate modelling and partial correlation analysis. Reproducibility of metabolites and metabolite patterns, derived via principal component analysis (PCA), was assessed by intraclass correlation. A unique component of metabolites unrelated to POPs was dissected by integrating metabolites and POPs using 2-way orthogonal partial least squares regression. ORs of T2D were estimated using conditional logistic regression.

Results: We identified 31 metabolites associated with fish intake that had poor to good reproducibility. A PCA-derived metabolite pattern strongly correlated with fish intake ($\rho = 0.37$, $P < 0.001$) but showed no association with T2D risk. Integrating fish-related metabolites and POPs led to a unique metabolite component independent of POPs, which tended to be inversely associated with T2D risk (OR: 0.75; 95% CI: 0.54, 1.02, $P = 0.07$). This component mainly consisted of metabolites reflecting fatty fish intake.

Conclusions: Our results suggest that fatty fish intake may be beneficial for T2D prevention, after removing the counteractive effects of coexposure to POPs in Swedish adults. Integrating metabolite markers and POP exposures appears a promising approach to advance the understanding of associations between fish intake and T2D incidence. *J Nutr* 2019;149:1413–1423.

Keywords: metabolomics, type 2 diabetes, fish biomarkers, persistent organic pollutants, nested case-control study, O2PLS modeling

Introduction

The prevalence of type 2 diabetes (T2D) is increasing globally (1). Diet has been suggested as an important modifiable risk factor in T2D (2, 3). Among dietary exposures relevant for T2D prevention, fish has demonstrated beneficial effects on several risk factors in T2D, e.g., lipid profile, blood pressure, and inflammation, in randomized controlled trials. However, there is conflicting evidence regarding the association between fish intake and risk of developing T2D in observational studies (4–11).

There are several possible reasons for the inconsistent findings. Measurement errors in self-reported fish intake, e.g.,

Manuscript received January 8, 2019. Initial review completed February 13, 2019. Revision accepted March 20, 2019.
First published online June 18, 2019; doi: https://doi.org/10.1093/jn/nxz068.

via an FFQ, are difficult to avoid and may contribute to the inconsistency (12, 13). Moreover, biomarkers of fish intake, such as DHA (22:6n–3) and EPA (20:5n–3), show inconsistent results in relation to T2D risk (14, 15). Although many health effects have been attributed to these fatty acids, they may not entirely explain the role of fish intake in T2D development (16, 17). Wide-scale analysis of metabolites associated with fish consumption might better reflect fish intake and improve understanding of the role of fish intake in disease-related metabolic processes.

Untargeted metabolomics, defined as comprehensive analysis of metabolites in biological samples, has become a key technology for identifying biomarkers that reflect dietary exposure (18). Using this technique, recent dietary interventions have found changes in several metabolites in response to a diet rich in fish, e.g., 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid (CMPF), lipids containing PUFAs, and carnitine (19–21). However, to our knowledge, very few studies have investigated whether these metabolites can reflect habitual fish intake and whether they are associated with risk of developing T2D in a free-living population.

Importantly, fish intake is also a main contributor to environmental persistent organic pollutant (POP) exposure in the general population (22, 23). Such pollutants have been associated with increased T2D risk (24). Consequently, POP exposure has been indicated to modify the association between fish intake and T2D risk in observational studies, a possible explanation for the inconsistent findings on the association between fish and T2D risk (5, 7). To improve the understanding of health effects of fish per se, it is of great importance to deduce the effect of fish independently of that of POP exposures in T2D development.

The aim of the present study was to identify plasma metabolites associated with fish intake that could be used as potential biomarkers of fish intake in a Swedish free-living population. Moreover, for the first time, to our knowledge, we conducted joint analysis of identified fish-related metabolite markers and plasma POPs, in order to investigate associations between fish intake and T2D risk, independently of POP exposures.

The establishment of this study and the metabolomics analysis were supported by the Swedish University of Agricultural Sciences through a young investigators' quality grant (to RL). The salaries of RL, LS, and CB were covered by grants from the Chalmers Foundation, Swedish Research Council, Dr Håkanssons Foundation, and Kristina Stenborg Foundation. A Swedish Research Council for Health, Working Life and Welfare (FORTE) grant (to RL) supported the setting up of the case-control study. The LC-MS metabolomics unit at University of Eastern Finland is supported by Biocenter Finland and KH holds an Academy of Finland Research fellowship. We also acknowledge Swedish Research Council (VR) grant 2017-00822 (to RL).

Author disclosures: LS, CB, IAB, IJ, OR, CDV, HK, KH, AÅ, and RL, no conflicts of interest.

None of the funding bodies had any role in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

Supplemental Tables 1–6 and Supplemental Figures 1–6 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/jn/.

Address correspondence to LS (e-mail: shlin@chalmers.se).

Abbreviations used: AROC, area under the receiver operating characteristic curve; CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid; ESI, electrospray ionization; HILIC, hydrophilic interaction LC; ICC, intraclass correlation; lysoPE, lysophosphatidylethanolamine; MSI, Metabolomics Standard Initiative; O2PLS, 2-way orthogonal partial least squares regression; PCA, principal component analysis; PLS, partial least squares; POP, persistent organic pollutant; T2D, type 2 diabetes; VIP, Västerbotten Intervention Programme.

Methods

Study population

A nested case-control study was set up within the Västerbotten Intervention Programme (VIP) (25) (**Figure 1**A). Detailed information on the VIP and selection of participants is reported elsewhere (25, 26). In brief, inhabitants in Västerbotten County aged 40, 50, and 60 y (and until 1996 also 30 y) were invited to participate in a health examination with blood sampling. Participants were asked to complete a questionnaire capturing information on diet, socioeconomic conditions, self-rated health, personal health history, family history of diabetes, tobacco use, and physical activity. The VIP was initiated in 1985 and participants included in the present study were recruited from 1991 to 2005 (baseline). T2D cases were diagnosed according to the Diabetes Register in Northern Sweden (DiabNorth) and were confirmed by diabetes specialists (27). In total, 421 participants (male = 196, female = 225) at baseline (median 7 y before diagnosis) among eligible diabetes cases who had an unthawed fasting plasma sample in the biobank were included in the study. Each case was individually matched to 1 nondiabetic VIP participant (healthy controls). Matching criteria included gender, age (±2 y), sample date (±90 d), and version of questionnaire at baseline examination. Among the 421 case-control pairs, 149 pairs had plasma POPs measured at baseline (see the section "Assessment of plasma POP concentrations" for details) and a follow-up sample drawn 10 y after baseline. In the present study, the samples at baseline were used to identify metabolite markers of fish intake (n = 421 pairs) and to investigate associations between fish markers and T2D (n = 149 pairs that had both baseline metabolomics and POP measurements available). The follow-up samples from healthy controls (n = 149) were included in order to estimate the long-term reproducibility of putative dietary biomarkers (see the section "Statistical analysis" for details). The entire study protocol was approved by the Regional Ethics Committee in Uppsala, Sweden (registration no. 2014/011). In the study, we explored plasma metabolites reflecting fish intake and their association with incident T2D. We also took plasma POP exposures into account in the analysis. The outcomes of interest did not change during the analysis.

Assessment of fish intake

Fish intake was assessed by 2 modified and validated versions of the Northern Sweden FFQ: 1 with 84 food items and 1 with 64 items (28, 29). Both questionnaires included the same 2 questions on fish intake. Total fish intake (grams per day) for each participant was estimated by adding together consumption of fatty fish (including herring, salmon, and whitefish) and lean fish (including perch, bass, and cod).

Assessment of plasma POP concentrations

Plasma concentrations of 16 POPs were measured for the 149 case-control pairs that had both baseline and 10-y follow-up samples available. The analysis covered several of the most widely dispersed POPs, including 10 polychlorinated biphenyl congeners (74, 99, 118, 153, 138, 156, 187, 183, 170, and 180), the pesticide dichlorodiphenyltrichloroethane and its metabolite dichlorodiphenyldichloroethylene, hexachlorobenzene, β-hexachlorocyclohexane, oxychlordane, and *trans*-nonachlordane. All analyses were performed at the National Institute for Health and Welfare in Finland, using GC-triple quadrupole MS (30). Plasma environmental POPs (nanograms per gram total lipids) were standardized for total lipids (i.e., plasma POPs divided by total lipids = 2.27 × total cholesterol + TGs + 0.623) when used for further statistical analysis (31, 32). Detailed analysis of blood lipids including serum cholesterol and serum TGs has been reported elsewhere (25).

Untargeted metabolomics analysis

A detailed description of the untargeted LC-MS metabolomics analysis is provided elsewhere (33). In brief, de-proteinized fasting heparin plasma samples were analyzed by an HPLC-quadrupole time-of-flight mass spectrometer system (Agilent Technologies) that consisted of a 1290 LC system, a Jetstream electrospray ionization (ESI) source, and

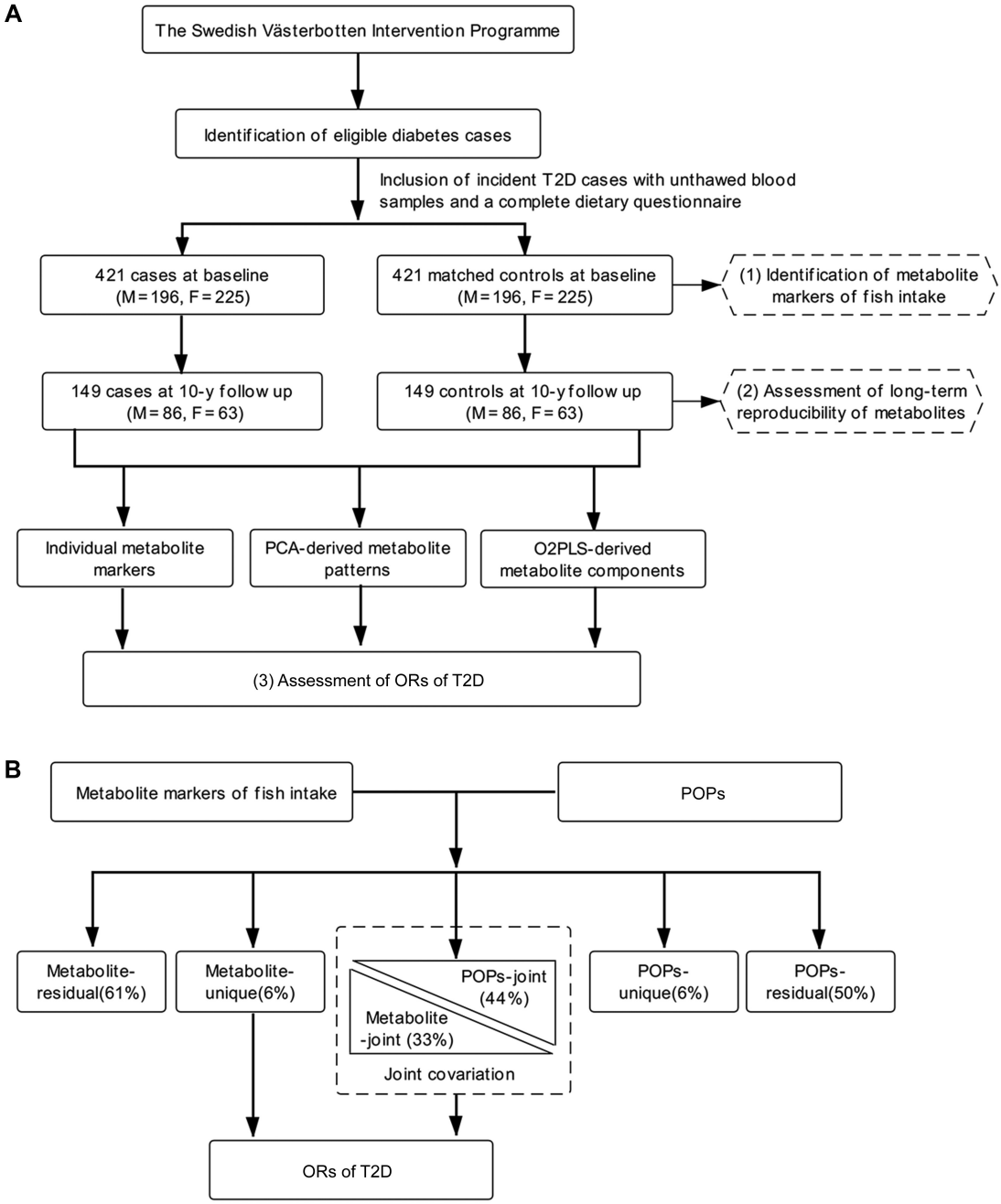

FIGURE 1 Flowchart of participant selection from the Swedish Västerbotten Intervention Programme (A) and overview of the O2PLS model structures obtained for integration of the metabolite markers of fish intake and plasma concentrations of POPs (B). The case-control study nested in the Swedish Västerbotten Intervention Programme consisted of 421 matched T2D case-control pairs of men and women aged 30–60 y at baseline and a subpopulation for which repeated samples collected after 10-y follow-up were available (149 pairs). Three major analyses were performed: *1*) metabolite markers of fish intake were identified using random forest modeling among healthy controls ($n = 421$), followed by partial correlation analysis adjusted for several lifestyle-related factors; *2*) the long-term reproducibility of metabolite markers of fish intake was assessed by intraclass correlation from repeated samples 10 y apart among healthy controls ($n = 149$); and *3*) ORs of T2D were estimated using conditional logistic regression on case-control pairs ($n = 149$) for individual metabolite markers of fish intake, metabolite patterns derived from PCA, and metabolite components derived from O2PLS. O2PLS, 2-way orthogonal partial least squares regression; PCA, principal component analysis; POP, persistent organic pollutant; T2D, type 2 diabetes.

a 6540 UHD accurate-mass quadrupole time-of-flight spectrometer. Plasma samples were analyzed in 8 batches and a constrained randomization was applied to keep sample pairs and follow-up samples from the same individual within the same batch, but with randomized order within batches. Instrumental analyses were performed with ~250 injections per batch, including study samples and quality-control samples that constituted ~16% of study samples to monitor the stability and functionality of the system. The samples were analyzed using reverse phase chromatography and hydrophilic interaction LC (HILIC) in positive (ESI+) and negative (ESI−) polarity. Instrument data were acquired using MassHunter Acquisition B.04.00 software (Agilent Technologies). Raw data acquired in each analytical batch were converted to mzXML format and deconvolution was performed with the open-source R package "XCMS" (34). Two types of quality-control samples were used to monitor the stability of the instrumental analysis and the R package "batchCorr" (33) was applied to correct

for within- and between-batch measurement error. In total, 19,180 and 10,060 features with CV <0.3 were considered to qualify for reverse phase chromatography and HILIC, respectively, after a stringent normalization procedure. Missing values were replaced with values randomly selected from a normal distribution between 0 and the lowest measured peak intensity within each feature. Throughout this article, the term "feature" refers to a mass spectral peak, i.e., a molecular entity with a unique *m/z* and retention time as measured by an LC-MS instrument. The term "metabolite" refers to a compound, with or without successful identification (see the section "Metabolite identification").

Statistical analysis

Differences between cases and controls in the baseline characteristics of participants were compared by 2-tailed Student's *t* test (or Wilcoxon's Signed Rank test if data were nonnormally distributed) for continuous variables and chi-square test for categorical variables.

Identification of individual metabolite markers of fish intake

Plasma metabolites associated with fish intake were identified using a multivariate analysis pipeline adapted for large-scale metabolomics data to minimize false positive associations, followed by partial Spearman rank correlation analysis adjusted for covariates. In brief, a sparse partial least squares (PLS) regression [R package "mixOmics" (35)] was applied on qualified metabolite features obtained from reverse phase chromatography and HILIC as a prefilter to remove the majority of uninformative features per chromatographic mode. After stringent filtering, the subset of data was then modeled with a PLS algorithm incorporated into a repeated double cross-validation framework with unbiased variable selection [R package "MUVR" (36)], in order to determine the features that predicted estimated fish intake from the FFQ. Fish intake measured by FFQ among healthy controls at baseline (n = 421) was energy-adjusted using the density method (26, 37) and was used as the dependent variable in the PLS modeling. Model performance was confirmed by permutation analysis (38). The model based on HILIC features showed poor performance (permutation test P = 0.25) and HILIC features were thus not considered further in the analysis. In total, 155 features from reverse phase chromatography were optimally selected to best predict fish intake.

For each feature selected by multivariate modeling, the direct association with fish intake was assessed using partial Spearman rank correlation analysis among all participants included and adjusting for case-control status and the following variables collected at baseline: age, gender, BMI (in kg/m^2), smoking status (smoker, former smoker, occasional smoker, nonsmoker), education (elementary school, vocational school, secondary school, university education/college), physical activity (inactive, moderately inactive, moderately active, active), and food items that were highly correlated with fish intake in the study population, i.e., whole grains and alcohol. In total, 38 features were independently associated with fish intake (Bonferroni-adjusted P < 0.05) and these features were then subjected to identification and further analyses (see the section "Identification of metabolite patterns of fish intake" below, **Supplemental Table 1**).

Identification of metabolite patterns of fish intake

We performed principal component analysis (PCA) [R package "psych" (39)] on 31 fish-related metabolites identified from the aforementioned 38 features to deduce metabolite patterns that reflected total fish intake. We considered the first 2 components that had an eigenvalue >2 and met the criteria of Very Simple Structure (40).

Joint analysis of metabolite markers of fish intake and plasma POPs

We found that fish-related metabolites identified were significantly correlated with several species of POPs (**Supplemental Figure 1**). To allow investigation of the independent role of fish-related metabolite biomarkers in T2D risk, we decomposed the unique metabolite component reflecting fish intake that was unrelated to plasma POPs from the shared component of fish metabolite markers and POP covariance (Figure 1B), using 2-way orthogonal partial least squares analysis [O2PLS, R package "OmicsPLS" (41)]. O2PLS is a generalization of orthogonal projections to latent structures, which is designed for integrating 2 high-dimensional data sets (42). It uses multiple linear regression to decompose systematic variability in 2 data sets, e.g., metabolites and POPs, into 3 partitions (Figure 1B): a joint covariation that represents the variability shared between metabolites and POPs (i.e., a joint component); an orthogonal variation that represents the unique systematic variation in each of the data matrixes (i.e., a unique component); and a residual partition that captures the unsystematic variation in the 2 data sets. Importantly, the joint component of metabolites described the variability in the plasma metabolome data that were useful for predicting plasma POPs. This component was thus highly correlated with POPs. The unique metabolite component described the systematic variability in the metabolite data that were not correlated with POPs. The data were autoscaled for O2PLS and the number of components was determined by 7-fold cross-validation. Permutation analysis showed that the resulting model had better performance than models with a permuted data matrix (n = 10,000) (**Supplemental Figure 2**).

Assessment of the discriminative ability of metabolite markers of fish intake and their long-term reproducibility

Metabolite markers of fish intake were evaluated for their ability to discriminate high self-reported intake (top tertile) from low intake (bottom tertile) among healthy controls, using the Biomarker Analysis Module of MetaboAnalyst 3.0 (43). Predictors with an area under the receiver operator characteristic curve (AROC) value >0.7 were considered good for classification of low compared with high intakes (44, 45).

The long-term reproducibility of identified metabolites was estimated by intraclass correlation (ICC) between the 2 sampling occasions over the 10-y period among a subset of the healthy controls (n = 149). To account for potential intraindividual systematic changes in fish intake over the years, ICC was calculated based on rank-transformed data (26).

Assessment of associations between metabolite markers of fish intake and T2D risk

Individual metabolite markers of fish intake (i), PCA-derived metabolite patterns (ii), and an O2PLS-derived unique metabolite component of fish intake (iii) were used as predictors of the likelihood of developing T2D, using conditional logistic regression [R package "survival" (46)]. ORs and 95% CIs were calculated using tertiles of exposures and per-SD increments among the subset of case-control pairs for which plasma concentrations of POPs were analyzed (n = 149 pairs). Three models were constructed for each of the exposures (i–iii): *1*) a crude model without adjustment; *2*) a multivariable model adjusted for lifestyle-related factors: BMI, smoking status, education, physical activity, and daily energy intake (kilocalories per day); and *3*) a multivariable model further adjusted for overall POP exposure. To address collinearity potentially attributable to the high correlations between different POPs, the overall and concerted multipollutant exposure was estimated from PCA, with 2 orthogonal principal components together accounting for 70% of the total variance in plasma concentrations of POPs. In a complementary analysis, we also investigated the OR of T2D risk for FFQ-derived fish intake.

In addition, we applied the bootstrap resampling technique to obtain the distribution of ORs of T2D risk for fish intake before and after adjusting for POPs, aiming to examine the robustness and stability of OR estimates in the subpopulation consisting of a limited number of participants (n = 149 case-control pairs with both metabolomics data and POPs available). ORs of T2D risk for fish intake were computed among 200 resampled case-control pairs with replacement and the analysis was repeated 10,000 times.

All statistical analyses were conducted in R version 3.5.3 (R Foundation), except for ICC, for which the SAS macro "%icc9" (SAS

version 9.4, SAS Institute; https://www.hsph.harvard.edu/donna-spiegelman/software/icc9/) was used.

Metabolite identification

Each of the metabolite features related to fish intake was identified based on accurate mass and tandem MS fragmentation matched against online databases or the literature (26, 47–50). The confidence level of annotation was categorized according to the Metabolomics Standard Initiative (MSI) (51): the use of literature and databases results in level 2 identifications (MSI 2). Annotated classes (MSI 3) were presented as "putative chemical class mass @retention time," whereas unknown compounds (MSI 4) were presented as "analytical mode m/z @ retention time" only (Supplemental Table 1).

Results

Participant characteristics

The baseline characteristics of the study population are presented in **Table 1**. No difference in self-reported fish intake was observed between T2D cases and matched controls. In general, cases had higher plasma concentrations of POPs than controls in the subset of the population for which POPs were analyzed.

Metabolite markers of fish intake

We identified 31 metabolites associated with fish intake, i.e., potential fish markers, independent of case/control status, age, gender, BMI, and lifestyle-related factors (**Table 2**). The 31 fish-metabolite associations were also confirmed among the subset of 149 case-control pairs for which plasma concentrations of POPs were analyzed (**Supplemental Table 2**). Identified metabolites included EPA, DHA, several species of phosphatidylcholines and phosphatidylethanolamines (containing C18:2, C20:4, C20:5, C22:5, and C22:6), 2 furan fatty acids (CMPF and 3-carboxy-4-methyl-5-pentyl-2-furanpropionic acid), carnitine, and several unannotated metabolites. Most of the individual metabolites had poor predictive accuracy in discriminating self-reported high intake from low intake of fish (AROC $\leq$ 0.7, **Supplemental Table 3**). Among the 31 metabolites identified, 24 showed good reproducibility among healthy controls (0.4 $\leq$ ICC $\leq$ 0.56) sampled 10 y apart (Table 2).

We found that 2 components derived from PCA accounted for 62% of the total variance in the fish-related metabolite markers identified. Metabolites that correlated positively with fish intake had a high loading in the fish-PCA-component 1, which also showed higher correlation with intake of fish (ρ = 0.37) and somewhat better predictive ability for fish intake (AROC = 0.79) than individual metabolites (0.57 $\leq$ AROC $\leq$ 0.73), and had good reproducibility (ICC = 0.51) (Table 2). In contrast, the fish-PCA-component 2 showed a weak correlation with fish intake.

Metabolite markers of fish intake and T2D risk

Among the 31 fish-related metabolite markers identified, 4 were positively associated with T2D risk, whereas 2 were inversely associated, when investigated individually. The results were not substantially influenced after adjusting for lifestyle factors (**Supplemental Figure 3**, **Supplemental Table 4**). No significant association between the metabolite patterns and T2D risk was observed (fish-PCA-component 1: OR: 1.12; 95% CI: 0.85, 1.50, P = 0.45; fish-PCA-component 2: OR: 0.92; 95% CI: 0.68, 1.20, P = 0.59). Adjusting plasma POPs in the risk models attenuated ORs of T2D for individual fish-related metabolites (Supplemental Figure 3): only 1 metabolite remained positively associated with T2D, whereas 4 showed inverse associations. However, no significant association between the fish-PCA-component 1 that better reflected fish intake and T2D risk was observed after adjustment for plasma POPs (OR: 0.86; 95% CI: 0.66, 1.17, P = 0.37) (**Supplemental Figure 4**).

O2PLS-derived metabolite components reflecting fish intake and T2D risk

Using O2PLS, we decomposed 1 unique metabolite component that was unrelated to plasma POPs and 1 joint metabolite component that represented the variability shared between plasma metabolite markers of fish intake and POPs (Figure 1B). Of the 10 metabolites that contributed to the unique metabolite component, 8 were phosphatidylcholines and phosphatidylethanolamines (**Figure 2**A). Moreover, this unique component was in particular correlated with fatty fish intake and, as expected, uncorrelated with plasma POPs (Figure 2B). Unlike the unique component, the joint metabolite component mainly consisted of carnitine, EPA, DHA, CMPF, and unknown metabolites that were all highly correlated with EPA and DHA (**Supplemental Table 5**).

Interestingly, the unique metabolite component showed a clear trend to be inversely associated with T2D risk in the small subsample set of n = 149 cases and controls, but did not reach statistical significance (OR: 0.78; 95% CI: 0.59, 1.06, P = 0.09) (Figure 2C). As expected, as an inherent result of the O2PLS method (i.e., the deconvoluted unique metabolite component was unrelated to POPs), adjusting for plasma POPs in risk models did not substantially affect this association (OR: 0.75; 95% CI: 0.54, 1.02, P = 0.07). In contrast, we observed that the association between the joint component derived from O2PLS, i.e., the component that was affected by POPs, and T2D risk was greatly attenuated after adjusting for plasma POPs (Figure 2C): the OR of T2D risk was reduced from 1.26 to 0.93. In the bootstrapping analysis that was conducted to evaluate the accuracy of the OR estimates, we found no marked difference between 10,000 bootstrapped OR estimates of T2D risk for metabolite components and the estimates in the original sample (**Supplemental Table 6**).

Discussion

Using untargeted metabolomics, we identified 31 plasma metabolites that were associated with fish intake in Swedish men and women and assessed their long-term reproducibility over 10 y. We found no apparent association between fish intake and T2D risk. Importantly, for the first time, to our knowledge, we performed a supervised multivariate modeling (O2PLS) on fish-related metabolites and plasma POPs. This approach allowed us to identify the unique metabolite component of fish intake unrelated to POPs, which mainly consisted of metabolites reflecting fatty fish intake and tended to be inversely associated with T2D risk (OR: 0.75; 95% CI: 0.54, 1.02, P = 0.07). The overall findings of our study highlight the potential of using plasma metabolites reflecting habitual fish intake in Swedish adults as a complement to traditional dietary assessment, and also support a benefit of fatty fish intake in T2D prevention after removal of the counteractive effects of POP exposures.

TABLE 1 Baseline characteristics of the Swedish adults included in the nested case-control study of incident type 2 diabetes in the Swedish Västerbotten Intervention Programme[1]

Characteristics	Cases ($n = 421$)	Controls ($n = 421$)	P
Men,[2] %	46.5	46.5	
Age,[2] y	50.4 ± 7.91	50.4 ± 7.90	
BMI, kg/m^2	29.7 ± 4.73	25.5 ± 3.81	<0.001
Plasma glucose, mmol/L	6.12 ± 0.91	5.45 ± 1.15	<0.001
Serum TGs, mmol/L	2.01 ± 1.26	1.34 ± 0.65	<0.001
Serum total cholesterol, mmol/L	5.90 ± 1.26	5.70 ± 1.13	0.005
Systolic blood pressure, mm Hg	139 ± 17.8	128 ± 16.7	<0.001
Diastolic blood pressure, mm Hg	85.1 ± 10.4	79.4 ± 9.70	<0.001
Total caloric intake, kcal/d	1742 ± 618	1745 ± 623	0.6
Smoking status			0.03
Current smoker, %	21.7	18.2	
Former smoker, %	38.6	33.3	
Occasional smoker, %	0.70	3.82	
Nonsmoker, %	39.0	44.2	
Physical activity index			0.08
Inactive, %	17.2	16.3	
Moderately inactive, %	36.4	35.0	
Moderately active, %	28.8	28.1	
Active, %	17.7	20.1	
Education			0.03
Elementary school	33.3	28.6	
Vocational (training) school	28.8	26.2	
Secondary school	21.5	21.1	
University education/college	16.3	23.6	
Fish measured by FFQ, g · d^{-1} · 1000 kcal^{-1}			
Fatty fish	3.02 (0.14, 7.3)	3.31 (0.13, 8.5)	0.09
Lean fish	6.21 (0.32, 13.1)	6.32 (0.12, 11.9)	0.9
Plasma environmental POPs, ng/g total lipids[3]			
HCB	35.4 ± 14.2	30.6 ± 11.6	0.002
β-HCH	21.5 ± 15.3	15.7 ± 9.60	<0.0001
Oxychlordane	6.71 ± 3.91	5.22 ± 2.91	0.001
Trans-nonachlordane	12.2 ± 6.61	10.1 ± 5.43	0.002
DDT	8.21 ± 5.51	6.22 ± 5.02	0.001
DDE	294 ± 187	248 ± 174	0.04
PCB74	7.21 ± 3.42	6.74 ± 3.92	0.02
PCB99	10.8 ± 4.90	9.81 ± 5.90	0.01
PCB118	22.4 ± 11.4	20.6 ± 16.5	0.02
PCB153	145 ± 54.3	144 ± 62.1	0.4
PCB138	98.3 ± 38.1	92.4 ± 42.4	0.05
PCB156	16.1 ± 6.40	16.0 ± 6.71	0.9
PCB187	30.6 ± 13.8	30.1 ± 15.2	0.5
PCB183	13.4 ± 6.42	12.3 ± 6.91	0.06
PCB180	99.6 ± 37.5	100 ± 40.7	0.3
PCB170	49.9 ± 18.7	50.8 ± 19.7	0.9

[1]Values are means ± SDs, medians (10th, 90th percentiles), or percentages. DDE, dichlorodiphenyldichloroethylene; DDT, dichlorodiphenyltrichloroethane; HCB, hexachlorobenzene; HCH, hexachlorocyclohexane; PCB, polychlorinated biphenyl; POP, persistent organic pollutant.
[2]Matching factors in the nested case-control study.
[3]$n = 149$ for cases and controls.

Metabolite markers of fish intake

We replicated EPA, DHA, and CMPF, which have been reported as potential biomarkers of fish intake in previous observational studies (44, 52, 53). Of note, metabolites found to change in response to diets rich in fish in previous randomized controlled trials, including 3-carboxy-4-methyl-5-pentyl-2-furanpropionic acid, phosphatidylcholines, and carnitine (20, 49), were also confirmed in our observational study. Although the magnitudes of FFQ–metabolite correlations were weaker than those of correlations reported in intervention studies (19, 20, 49), metabolite markers identified in observational studies tend to be more sensitive and robust and might be used to reflect habitual fish intake.

Moreover, 19 metabolites showed similar reproducibility over 10 y ($0.43 \leq \text{ICC} \leq 0.56$) to fish intake assessed by FFQ (ICC = 0.43), reinforcing the potential for using metabolite markers in studies investigating the relation between fish and diet-related diseases (54, 55). We acknowledge that ICC estimates for metabolite markers might be lower than the true ICCs owing to a long period between repeated sampling, i.e., 10 y, which may be inferred from systematic changes in metabolite concentrations between two occasions. We found

TABLE 2 Metabolites that were significantly associated with fish intake in Swedish adults in the nested case-control study of incident type 2 diabetes in the Swedish Västerbotten Intervention Programme and in studies reported in the literature[1]

Metabolite	ρ[2]	P[2]	ICC[3]	References[4]
RP303.6468@9.99	0.22	2.73×10^{-10}	0.49 (0.37, 0.61)	
RP490.3516@9.00	0.21	1.55×10^{-09}	0.54 (0.43, 0.65)	
PC (22:6/16:0)	0.19	2.30×10^{-08}	0.53 (0.41, 0.64)	(49, 52)
RP326.0461@7.52	0.19	3.53×10^{-08}	0.40 (0.27, 0.53)	
CMPF	0.19	3.66×10^{-08}	0.45 (0.33, 0.58)	(20, 44, 53)
PC (18:2/20:4)	0.19	3.69×10^{-08}	0.56 (0.45, 0.66)	(49)[5]
RP255.1228@7.51	0.18	9.90×10^{-08}	0.45 (0.32, 0.57)	
RP223.0968@7.52	0.18	1.23×10^{-07}	0.45 (0.33, 0.58)	
DHA	0.18	2.52×10^{-07}	0.51 (0.40, 0.63)	(44, 53)
RP237.1121@7.52	0.17	3.84×10^{-07}	0.39 (0.26, 0.53)	
lysoPC (22:6)	0.17	4.04×10^{-07}	0.47 (0.34, 0.59)	(20, 49, 52)
Carnitine (13:0)	0.17	7.56×10^{-07}	0.43 (0.30, 0.56)	(49)[5]
PC850.5620@12.37	0.17	1.22×10^{-06}	0.39 (0.27, 0.53)	
RP320.0805@7.52	0.17	1.27×10^{-06}	0.44 (0.31, 0.57)	
RP211.1325@8.56	0.16	1.86×10^{-06}	0.21 (0.10, 0.40)	
PC (22:6/18:3)	0.16	4.55×10^{-06}	0.38 (0.26, 0.53)	
7-oxo-11E,13-tetradecadienoic acid	0.15	8.95×10^{-06}	0.35 (0.22, 0.50)	
RP684.5546@13.42	0.15	9.28×10^{-06}	0.55 (0.43, 0.65)	
PE (22:6/P-18:0)	0.15	2.18×10^{-05}	0.47 (0.35, 0.59)	
PC (22:6/18:0)	0.14	2.45×10^{-05}	0.52 (0.41, 0.64)	(49)[5]
RN738.3984@10.54	0.14	2.75×10^{-05}	0.35 (0.23, 0.50)	
RP383.1670@10.53	0.14	2.94×10^{-05}	0.50 (0.38, 0.61)	
Carnitine391.27@8.1	0.14	3.70×10^{-05}	0.37 (0.25, 0.52)	(49)[5]
RN443.1569@7.87	0.14	4.30×10^{-05}	0.38 (0.25, 0.52)	
RP854.5664@11.92	0.13	1.11×10^{-04}	0.43 (0.31, 0.56)	
3-carboxy-4-methyl-5-pentyl-2-furanpropionic acid	0.12	4.62×10^{-04}	0.44 (0.32, 0.57)	(20, 49)[5]
EPA	0.12	1.63×10^{-03}	0.43 (0.31, 0.56)	(44, 53)
RN995.6119@10.01	−0.14	5.92×10^{-05}	0.47 (0.35, 0.59)	
lysoPE (18:2)	−0.14	3.19×10^{-05}	0.50 (0.38, 0.61)	(52)
RN802.5549@12.00	−0.15	1.14×10^{-05}	0.12 (0.03, 0.38)	
lysoPE (20:3)	−0.17	3.94×10^{-07}	0.42 (0.29, 0.55)	
PCA of fish-related metabolites				
Fish-PCA-component 1	0.37	4.57×10^{-12}	0.51 (0.39, 0.61)	
Fish-PCA-component 2	0.04	0.30	0.30 (0.21, 0.42)	

[1] Identified metabolites associated with fish intake measured by FFQ. CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid; ICC, intraclass correlation; lysoPC, lysophosphatidylcholine; lysoPE, lysophosphatidylethanolamine; PC, phosphatidylcholine; PCA, principal component analysis; PE, phosphatidylethanolamine; RN, reverse phase chromatography negative ionization mode; RP, reverse phase chromatography positive ionization mode.
[2] Partial Spearman rank correlation coefficients between metabolites and fish intake, controlling for age, gender, case/control status, BMI, smoking status, education, and physical activity index ($n = 421$ case-control pairs).
[3] ICC coefficient representing 10-y long-term reproducibility of metabolites among a subset of healthy controls ($n = 149$) with repeated samples available. ICC $\geq$ 0.4 denotes good reproducibility.
[4] Previous findings regarding fish–metabolite associations reported in the literature. For EPA, DHA, and CMPF, the list of articles is not exhaustive. Reviews are not considered. For lipids, reference is made only to publications that report fatty acid constituents.
[5] The metabolite was only found to change in response to fish intake in controlled intervention studies.

that ICCs for EPA (0.43) and DHA (0.51), well-established biomarkers, were lower than in a previous study where a short-time reproducibility was assessed (within 6 mo, ICC = 0.67) (56). However, no study that we know of has investigated ICCs of other identified fish-related metabolites, e.g., CMPF, and several phosphatidylcholines containing marine-derived PUFAs, making comparisons impossible.

Furthermore, we observed that fish intake was somewhat better reflected when using metabolite patterns derived by PCA, rather than single metabolites, as exposure measurements. Combining metabolites in a PCA yielded slightly better discrimination between high and low fish intakes than did single markers. This is in line with other studies in which a panel of metabolites was found to better predict intake of coffee (57) and sugar-sweetened beverages (45) than single metabolites. However, the performance of panels needs to be carefully evaluated across different populations.

Associations between metabolite markers of fish intake and T2D risk

Fish is an abundant source of omega-3 PUFAs, vitamins, and minerals, and is also considered as a central food item for healthy dietary patterns that may prevent development of T2D (7–10, 58, 59). It is also noteworthy that fish, in particular fatty fish (e.g., Baltic herring, salmon), are exposed to POP contamination (60) and coexposure to POPs present in fish may affect the association between fish intake and T2D risk in populations (5, 7). We found strong correlations between POPs in plasma and fatty fish intake in the current population (Figure 2B). This is plausible, because these measured POPs often

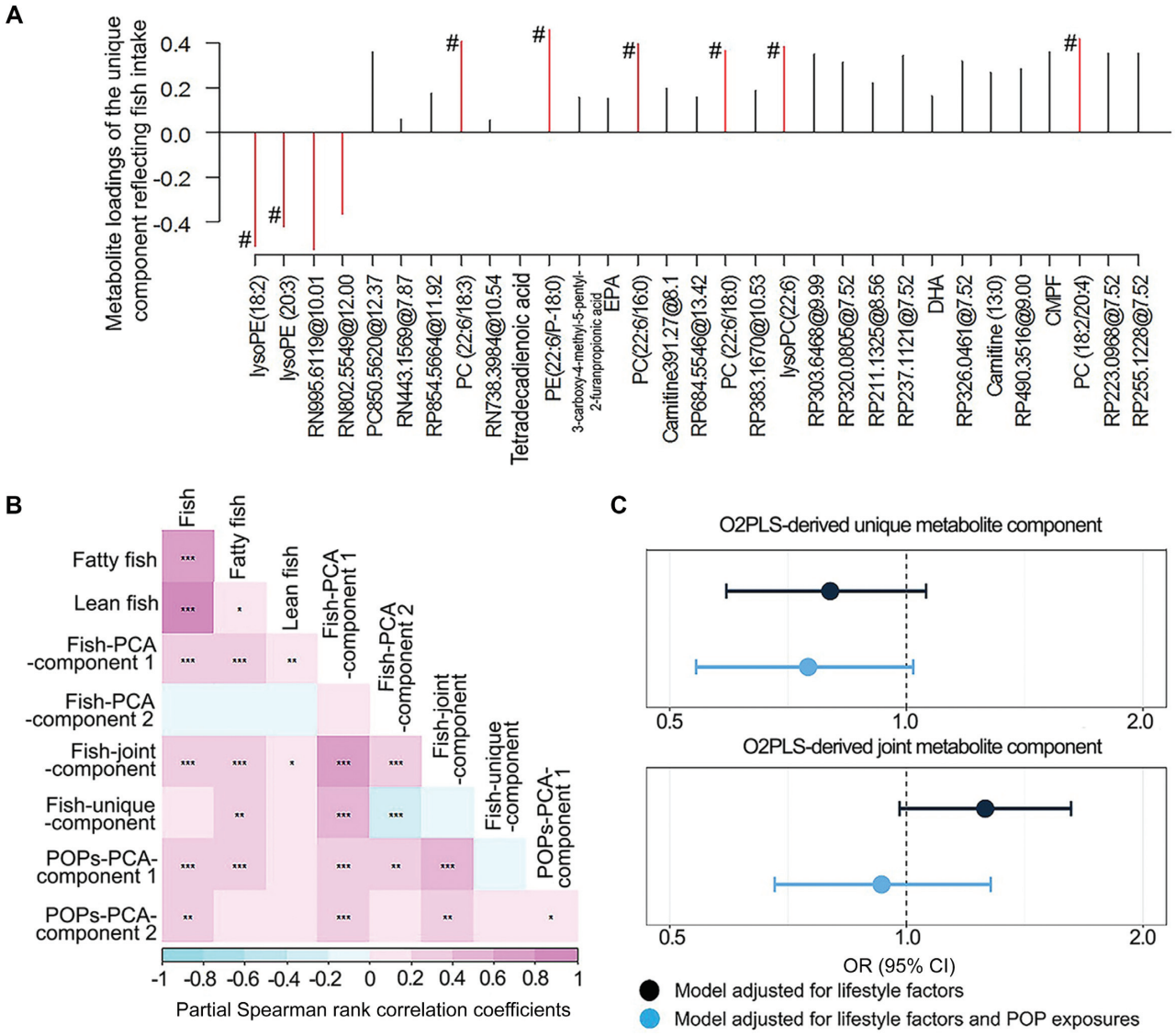

FIGURE 2 Associations between metabolite markers of fish intake, plasma concentrations of POPs, and ORs of type 2 diabetes in Swedish adults. (A) Metabolite loadings of the unique metabolite component derived from O2PLS modeling. Loadings of the top 10 variables contributing to the components are presented in red. (B) Partial Spearman rank correlation coefficients between metabolite patterns of fish markers (i.e., Fish-PCA-component 1, Fish-PCA-component 2, Fish-joint component, and Fish-unique component of metabolites), FFQ-assessed fish intake, and PCA-derived principal components of pollutants that reflected overall and concerted multipollutant exposures (POPs-PCA-component 1 and POPs-PCA-component 2), controlled for age, gender, case/control status, BMI, smoking status, education, and physical activity index ($n = 421$ case-control pairs). (C) Associations between variations in metabolites determined by O2PLS and risk of developing type 2 diabetes. OR per SD and 95% CIs were obtained from the conditional logistic regression model adjusted for lifestyle factors (presented in black) and the model further adjusted for POP exposure reflected by principal components of POPs (presented in blue). #Annotated PCs and PEs that had high loadings on the unique metabolite component of fish intake. *,**,***Significant correlation: $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. *P* values are false discovery rate–adjusted. CMPF, 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid; lysoPC, lysophosphatidylcholine; lysoPE, lysophosphatidylethanolamine; O2PLS, 2-way orthogonal partial least squares regression; PC, phosphatidylcholine; PCA, principal component analysis; PE, phosphatidylethanolamine; POP, persistent organic pollutant; RN, reverse phase chromatography negative ionization mode; RP, reverse phase chromatography positive ionization mode.

accumulate in fat deposits of fish (60). In our study, we found no conclusive evidence to support any association between fish intake, either measured by metabolomics (i.e., individual metabolites or PCA-derived metabolite patterns) or by FFQ-derived intake (**Supplemental Figure 5**), and T2D risk. By using a conventional multivariate logistic regression, adjustment for POPs in the risk models led to overall attenuation of the relation between identified metabolite markers of fish intake including well-known biomarkers of fatty fish intake, i.e., DHA, EPA, and CMPF, and T2D risk. This indicated that POPs adversely affected the association between fish intake (in particular, fatty fish intake) and T2D risk in the population.

In response to this and in order to uncover the independent association between fish intake and T2D risk, we employed O2PLS integrative modelling on fish-related metabolites and plasma POPs to dissect the unique metabolite component that reflected the variability in metabolites independently of POP exposure. Interestingly, we found that the unique metabolite component tended to be inversely associated with T2D risk (OR: 0.75; 95% CI: 0.54, 1.02, $P = 0.07$). The lack of

statistical significance at an α level of 0.05 may be attributed to the limited number of participants with both metabolomics data and plasma POP data available ($n = 149$ case-control pairs) (61). However, we found no marked difference between 10,000 bootstrapped OR estimates of T2D risk for fish intake and the estimates in the original sample, suggesting that our estimates reflect a robust finding. Moreover, the deconvoluted unique metabolite component of fish intake was mainly based on higher concentrations of phosphatidylcholines containing PUFAs found in marine foods and lower concentrations of lysophosphatidylethanolamines (lysoPEs) containing C20:3 and C18:2 (Figure 2A). Although the mechanisms remain unclear, these phospholipids have previously been associated with risk factors of T2D, such as fasting plasma insulin, LDL cholesterol, and inflammatory markers (49, 62, 63). Of note, the unique component derived from metabolite markers of fish intake was in particular correlated with fatty fish intake, not with lean fish intake. This finding indicates that the observed independent benefit of fish intake reflected by this unique component on T2D risk might be mainly driven by metabolites reflecting fatty fish intake. In order to investigate that, a similar O2PLS model was constructed based on the optimally identified set of metabolites that best predicted fatty fish intake and the results were very similar to those for total fish intake (**Supplemental Figure 6**). Moreover, metabolites that had high loadings in the fatty fish unique metabolite component included the lipids that contributed to the unique component of total fish intake, e.g., lysoPE (18:2), lysoPE (20:3), phosphatidylcholine (22:6/18:3), phosphatidylcholine (22:6/18:0), and lysophosphatidylcholine (22:6) (data not shown). Taken together, our results support a beneficial effect of fatty fish intake on T2D prevention, possibly through regulating phospholipids, and also indicate that this benefit might be masked by coexposure to POPs.

Our data did not support an association between lean fish intake and risk of T2D. However, we cannot exclude the possibility that other factors may confound potential associations between lean fish and T2D, such as cooking methods or other unmeasured contaminants that accumulate predominantly in lean fish (5, 10, 16, 64).

Strengths and limitations

Our study has several strengths. First, to our knowledge, it is the first study to investigate potential metabolite markers of fish intake and at the same time to assess their associations with risk of developing T2D in a free-living population. Second, we evaluated the sensitivity and specificity (as indicated by AROC), as well as the long-term reproducibility (as indicated by ICC), of identified metabolite markers, the results of which suggested an advantage of using metabolite patterns reflecting fish intake as a complement to traditional dietary assessment. In fact, a metabolite pattern showed a long-term reproducibility of the same magnitude as that of FFQ-derived fish intakes. Third, using O2PLS integrative modelling, we disentangled a metabolite component that was independent of plasma POP concentrations, which enabled us to unmask the independent benefits of fish intake on T2D risk. To the best of our knowledge, this has never been done previously. Our study thus provides important complementary information on the role of fish intake in T2D risk that cannot be obtained in studies that use fish intake assessed by self-reported measurements.

Our study also has several limitations. First, correlations between dietary intake and plasma metabolites might be lower than the true associations, owing to measurement errors of an FFQ administered at a single time point, variations in single measurements of metabolites reflecting day-to-day variation in the intake of fish, and the short half-lives of food-related metabolites in plasma. Second, we acknowledge that our study design makes it impossible to differentiate metabolites reflecting fish intake per se from associated lifestyles. However, our study has replicated metabolites reported previously as well-known biomarkers of fish intake in cohort studies and also identified metabolites shown to be affected in response to fish intake in dietary interventions, suggesting that identified metabolites could be used as potential biomarkers reflecting habitual fish intake. Third, we cannot rule out the risk that the observed inverse association between the unique metabolite component reflecting fatty fish intake and T2D risk might be subject to residual confounding inherent to a healthy lifestyle, despite our comprehensive efforts to adjust for plausible confounding lifestyle factors in multivariable models, including physical activity, smoking, and daily energy intake. It is also noteworthy that food items characterizing a Healthy Nordic Index (58), i.e., cabbage, root vegetables, and whole grains, did not correlate with the unique metabolite component ($|\rho| < 0.01$) and had no influence on the OR of T2D risk for this unique metabolite component (data not shown). Moreover, the data set does not allow identification of biomarkers for fatty fish intake per se, owing to apparent coconsumption of fatty fish and lean fish in the current population or misclassification inherent to the FFQ, e.g., difficulties for consumers in classifying fishes. However, the unique metabolite component consisted of metabolite markers that correlated with mainly fatty fish intake. Lastly, we did not manage to annotate all metabolites associated with fish intake. However, unknown compounds did not affect the major finding in this study concerning the beneficial role of fatty fish intake on T2D prevention when not counteracted by coexposure to POPs.

Conclusions

In conclusion, we found several fasting plasma metabolites that appeared to reflect habitual fish intake in men and women in Västerbotten County, Sweden. No apparent association between fish intake, either measured by metabolomics or by FFQ, and T2D risk was found. However, the joint analysis of fish-related metabolites and plasma POPs using supervised multivariate modeling (O2PLS) allowed us to deconvolute the metabolite component uniquely related to in particular fatty fish intake, mainly comprised of phospholipids, that showed a clear tendency to be inversely associated with T2D risk independently of POP exposures ($P = 0.07$). The finding supports beneficial effects of fatty fish on T2D prevention, but replication of study results in other cohorts with a larger sample size is necessary. Our results highlight the potential for using fasting plasma metabolites for objective assessments of habitual fish intake, along with measurement of pollutant exposures for improved understanding of associations between fish intake and risk of developing T2D in populations.

Acknowledgments

The authors' responsibilities were as follows—LS: participated in the sample analyses, performed the data processing and statistical analyses, interpreted the data, and wrote the manuscript; CB, IAB, and RL: conceived and designed the study; CB: supervised the statistical analyses and interpretation, and revised the manuscript; IAB, IJ, KH, OR, HK, CDV, and AÅ: interpreted the data and provided critical intellectual input; IAB, CDV, HK, and AÅ: contributed data on persistent organic pollutants; KH: participated in the untargeted LC-MS analysis and metabolite identification; RL: supervised the data

analyses and interpretation, revised the manuscript, and had the overall responsibility for the project; and all authors: read and approved the final manuscript.

References

1. World Health Organization. Global Report on Diabetes. Geneva: WHO; 2016.
2. Alhazmi A, Stojanovski E, McEvoy M, Garg ML. The association between dietary patterns and type 2 diabetes: a systematic review and meta-analysis of cohort studies. J Hum Nutr Diet 2014;27: 251–60.
3. Franks PW, McCarthy MI. Exposing the exposures responsible for type 2 diabetes and obesity. Science 2016;354:69–73.
4. Zhang M, Picard-Deland E, Marette A. Fish and marine omega-3 polyunsatured fatty acid consumption and incidence of type 2 diabetes: a systematic review and meta-analysis. Int J Endocrinol 2013;2013:501015.
5. Wallin A, Di Giuseppe D, Orsini N, Åkesson A, Forouhi NG, Wolk A. Fish consumption and frying of fish in relation to type 2 diabetes incidence: a prospective cohort study of Swedish men. Eur J Nutr 2017;56:843–52.
6. Chen C, Yang Y, Yu X, Hu S, Shao S. Association between omega-3 fatty acids consumption and the risk of type 2 diabetes: a meta-analysis of cohort studies. J Diabetes Investig 2017;8:480–8.
7. Marushka L, Batal M, David W, Schwartz H, Ing A, Fediuk K, Sharp D, Black A, Tikhonov C, Chan HM. Association between fish consumption, dietary omega-3 fatty acids and persistent organic pollutants intake, and type 2 diabetes in 18 First Nations in Ontario, Canada. Environ Res 2017;156:725–37.
8. Anil E. The impact of EPA and DHA on blood lipids and lipoprotein metabolism: influence of apoE genotype. Proc Nutr Soc 2007;66: 60–8.
9. Jacob J. Fish consumption and omega-3-fatty acids in prevention of diet-related noncommunicable diseases. J Soc Heal Diabetes 2016;4: 115–20.
10. Zhuang P, Wang W, Wang J, Zhang Y, Jiao J. Current level of fish consumption is associated with mortality in Chinese but not US adults: new findings from two nationwide cohort studies with 14 and 9.8 years of follow-up. Mol Nutr Food Res 2018;62:1700898.
11. Tørris C, Småstuen MC, Molin M. Nutrients in fish and possible associations with cardiovascular disease risk factors in metabolic syndrome. Nutrients 2018;10:952.
12. Bennett DA, Landry D, Little J, Minelli C. Systematic review of statistical approaches to quantify, or correct for, measurement error in a continuous exposure in nutritional epidemiology. BMC Med Res Methodol 2017;17:146.
13. Paeratakul S, Popkin BM, Kohlmeier L, Hertz-Picciotto I, Guo X, Edwards LJ. Measurement error in dietary data: implications for the epidemiologic study of the diet–disease relationship. Eur J Clin Nutr 1998;52:722–7.
14. Forouhi NG, Imamura F, Sharp SJ, Koulman A, Schulze MB, Zheng J, Ye Z, Sluijs I, Guevara M, Huerta JM, et al. Association of plasma phospholipid n-3 and n-6 polyunsaturated fatty acids with type 2 diabetes: the EPIC-InterAct Case-Cohort Study. PLoS Med 2016;13:1002094.
15. Virtanen JK, Mursu J, Voutilainen S, Uusitupa M, Tuomainen TP. Serum omega-3 polyunsaturated fatty acids and risk of incident type 2 diabetes in men: the Kuopio Ischemic Heart Disease Risk Factor study. Diabetes Care 2014;37:189–96.
16. van Woudenbergh GJ, van Ballegooijen AJ, Kuijsten A, Sijbrands EJG, van Rooij FJA, Geleijnse JM, Hofman A, Witteman J, Feskens EJM. Eating fish and risk of type 2 diabetes. Diabetes Care 2009;32: 2021–6.
17. The ASCEND Study Collaborative Group. Effects of n−3 fatty acid supplements in diabetes mellitus. N Engl J Med 2018;379:1540–50.
18. Rappaport SM, Barupal DK, Wishart D, Vineis P, Scalbert A. The blood exposome and its role in discovering causes of disease. Environ Health Perspect 2014;122(8):769–74.
19. Cheung W, Keski-Rahkonen P, Assi N, Ferrari P, Freisling H, Rinaldi S, Slimani N, Zamora-Ros R, Rundle M, Frost G, et al. A metabolomic study of biomarkers of meat and fish intake. Am J Clin Nutr 2017;105:600–8.
20. Hanhineva K, Lankinen MA, Pedret A, Schwab U, Kolehmainen M, Paananen J, De Mello V, Sola R, Lehtonen M, Poutanen K, 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.
21. Guasch-Ferré M, Bhupathiraju SN, Hu FB. Use of metabolomics in improving assessment of dietary intake. Clin Chem 2018;64:82–98.
22. MacKenzie BR, Almesjö L, Hansson S. Fish, fishing, and pollutant reduction in the Baltic Sea. Environ Sci Technol 2004;38:1970–6.
23. Järv L, Kiviranta H, Koponen J, Rantakokko P, Ruokojärvi P, Radin M, Raid T, Roots O, Simm M. Persistent organic pollutants in selected fishes of the Gulf of Finland. J Mar Syst 2017;171:129–33.
24. Ngwa EN, Kengne A-P, Tiedeu-Atogho B, Mofo-Mato E-P, Sobngwi E. Persistent organic pollutants as risk factors for type 2 diabetes. Diabetol Metab Syndr 2015;7:41.
25. Norberg M, Wall S, Boman K, Weinehall L. The Västerbotten Intervention Programme: background, design and implications. Glob Health Action 2010;3:4643.
26. Shi L, Brunius C, Johansson I, Bergdahl IA, Lindahl B, Hanhineva K, Landberg R. Plasma metabolites associated with healthy Nordic dietary indexes and risk of type 2 diabetes—a nested case-control study in a Swedish population. Am J Clin Nutr 2018;108:564–75.
27. Rolandsson O, Norberg M, Nyström L, Söderberg S, Svensson M, Lindahl B, Weinehall L. How to diagnose and classify diabetes in primary health care: lessons learned from the Diabetes Register in Northern Sweden (DiabNorth). Scand J Prim Health Care 2012;30: 81–7.
28. Johansson I, Hallmans G, Wikman A, Biessy C, Riboli E, Kaaks R. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. Public Health Nutr 2002;5:487–96.
29. Johansson I, Nilsson LM, Stegmayr B, Boman K, Hallmans G, Winkvist A. Associations among 25-year trends in diet, cholesterol and BMI from 140,000 observations in men and women in northern Sweden. Nutr J 2012;11:40.
30. Koponen J, Rantakokko P, Airaksinen R, Kiviranta H. Determination of selected perfluorinated alkyl acids and persistent organic pollutants from a small volume human serum sample relevant for epidemiological studies. J Chromatogr A 2013;1309:48–55.
31. Donat-Vargas C, Åkesson A, Tornevi A, Wennberg M, Sommar J, Kiviranta H, Rantakokko P, Bergdahl IA. Persistent organochlorine pollutants in plasma, blood pressure, and hypertension in a longitudinal study. Hypertension 2018;71:1258–68.
32. Phillips DL, Pirkle JL, Burse VW, Bernert JT, Henderson LO, Needham LL. Chlorinated hydrocarbon levels in human serum: effects of fasting and feeding. Arch Environ Contam Toxicol 1989;18:495–500.
33. Brunius C, Shi L, Landberg R. Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. Metabolomics 2016;12:173.
34. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 2006;78:779–87.
35. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput Biol 2017;13:e1005752.
36. Shi L, Westerhuis JA, Rosén J, Landberg R, Brunius C. Variable selection and validation in multivariate modelling. Bioinformatics 2019;35(6):972–80.
37. Playdon MC, Moore SC, Derkach A, Reedy J, Subar AF, Sampson JN, Albanes D, Gu F, Kontto J, Lassale C, et al. Identifying biomarkers of dietary patterns by using metabolomics. Am J Clin Nutr 2017;105: 450–65.
38. Lindgren F, Hansen B, Karcher W. Model validation by permutation tests. J Chemom 1996;10:521–32.
39. Revelle W. psych: Procedures for Personality and Psychological Research. Version 1.8.12. Evanston (IL): Northwest University; 2018.
40. Revelle W, Rocklin T. Very Simple Structure: an alternative procedure for estimating the optimal number of interpretable factors. Multivariate Behav Res 1979;14:403–14.

41. el Bouhaddani S, Houwing-Duistermaat J, Salo P, Perola M, Jongbloed G, Uh H-W. Evaluation of O2PLS in omics data integration. BMC Bioinformatics 2016;17:117–202.
42. Trygg J, Wold S. O2-PLS, a two-block (X–Y) latent variable regression (LVR) method with an integral OSC filter. J Chemom 2003;17:53–64.
43. Xia J, Sinelnikov IV, Han B, Wishart DS. MetaboAnalyst 3.0—making metabolomics more meaningful. Nucleic Acids Res 2015;43: W251–7.
44. Wang Y, Gapstur SM, Carter BD, Hartman TJ, Stevens VL, Gaudet MM, McCullough ML. Untargeted metabolomics identifies novel potential biomarkers of habitual food intake in a cross-sectional study of postmenopausal women. J Nutr 2018;148:932–43.
45. Gibbons H, McNulty BA, Nugent AP, Walton J, Flynn A, Gibney MJ, Brennan L. A metabolomics approach to the identification of biomarkers of sugar-sweetened beverage intake. Am J Clin Nutr 2015;101:471–7.
46. Therneau T (2015). A Package for Survival Analysis in S. version 2.38. [Internet]. Available from: https://CRAN.R-project.org/package=survival [cited Mar 2019].
47. Shi L, Brunius C, Lehtonen M, Auriola S, Bergdahl IA, Rolandsson O. Plasma metabolites associated with type 2 diabetes in a Swedish population: a case–control study nested in a prospective cohort. Diabetologia 2018;61(4):849–61.
48. Hanhineva K, Brunius C, Andersson A, Marklund M, Juvonen R, Keski-Rahkonen P, Auriola S, Landberg R. Discovery of urinary biomarkers of whole grain rye intake in free-living subjects using nontargeted LC-MS metabolite profiling. Mol Nutr Food Res 2015;59: 2315–25.
49. Tovar J, de Mello VD, Nilsson A, Johansson M, Paananen J, Lehtonen M, Hanhineva K, Björck I. Reduction in cardiometabolic risk factors by a multifunctional diet is mediated via several branches of metabolism as evidenced by nontargeted metabolite profiling approach. Mol Nutr Food Res 2017;61:1600552.
50. de Mello VD, Paananen J, Lindström J, Lankinen MA, Shi L, Kuusisto J, Pihlajamäki J, Auriola S, Lehtonen M, Rolandsson O, et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the Finnish Diabetes Prevention Study. Sci Rep 2017;7:46337.
51. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL, et al. Proposed minimum reporting standards for chemical analysis. Metabolomics 2007;3: 211–21.
52. Lu Y, Zou L, Su J, Tai ES, Whitton C, van Dam RM, Ong CN. Meat and seafood consumption in relation to plasma metabolic profiles in a Chinese population: a combined untargeted and targeted metabolomics study. Nutrients 2017;9:683.
53. Pallister T, Jennings A, Mohney RP, Yarand D, Mangino M, Cassidy A, MacGregor A, Spector TD, Menni C. Characterizing blood metabolomics profiles associated with self-reported food intakes in female twins. PLoS One 2016;11:0158568.
54. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutr J 2012;11:109.
55. Brennan L, Hu FB. Metabolomics-based dietary biomarkers in nutritional epidemiology—current status and future opportunities. Mol Nutr Food Res 2019;63(1):e1701064.
56. Albani V, Celis-Morales C, O'Donovan CB, Walsh MC, Woolhead C, Forster H, Fallaize R, Macready AL, Mathers JC, Adamson AJ, et al. Within-person reproducibility and sensitivity to dietary change of C15:0 and C17:0 levels in dried blood spots: data from the European Food4Me Study. Mol Nutr Food Res 2017;61(10): 1700142.
57. Rothwell JA, Fillâtre Y, Martin J-F, Lyan B, Pujos-Guillot E, Fezeu L, Hercberg S, Comte B, Galan P, Touvier M, et al. New biomarkers of coffee consumption identified by the non-targeted metabolomic profiling of cohort study subjects. PLoS One 2014;9: e93474.
58. Olsen A, Egeberg R, Halkjaer J, Christensen J, Overvad K, Tjonneland A. Healthy aspects of the Nordic diet are related to lower total mortality. J Nutr 2011;141:639–44.
59. Kanerva N, Kaartinen NE, Schwab U, Lahti-Koski M, Männistö S. The Baltic Sea Diet Score: a tool for assessing healthy eating in Nordic countries. Public Health Nutr 2014;17: 1697–705.
60. Lee DH, Porta M, Jacobs DR, Vandenberg LN. Chlorinated persistent organic pollutants, obesity, and type 2 diabetes. Endocr Rev 2014;35:557–601.
61. Nemes S, Jonasson JM, Genell A, Steineck G. Bias in odds ratios by logistic regression modelling and sample size. BMC Med Res Methodol 2009;9:56.
62. Suvitaival T, Bondia-Pons I, Yetukuri L, Pöhö P, Nolan JJ, Hyötyläinen T, Kuusisto J, Orešič M. Lipidome as a predictive tool in progression to type 2 diabetes in Finnish men. Metabolism 2018;78: 1–12.
63. Lankinen M, Schwab U, Erkkilä A, Seppänen-Laakso T, Hannila ML, Mussalo H, Lehto S, Uusitupa M, Gylling H, Orešič M. Fatty fish intake decreases lipids related to inflammation and insulin signaling—a lipidomics approach. PLoS One 2009;4:0005258.
64. Llull RM, Garí M, Canals M, Rey-Maquieira T, Grimalt JO. Mercury concentrations in lean fish from the western Mediterranean Sea: dietary exposure and risk assessment in the population of the Balearic Islands. Environ Res 2017;158:16–23.