scientific reports

Check for updates

OPEN

Plasma lipidomic profiles of kidney, breast and prostate cancer patients differ from healthy controls

Denise Wolrab[1], Robert Jirásko[1], Ondřej Peterka[1], Jakub Idkowiak[1], Michaela Chocholoušková[1], Zuzana Vaňková[1], Karel Hořejší[1], Ivana Brabcová[1], David Vrána[2,3], Hana Študentová[2], Bohuslav Melichar[2] & Michal Holčapek[1✉]

Early detection of cancer is one of the unmet needs in clinical medicine. Peripheral blood analysis is a preferred method for efficient population screening, because blood collection is well embedded in clinical practice and minimally invasive for patients. Lipids are important biomolecules, and variations in lipid concentrations can reflect pathological disorders. Lipidomic profiling of human plasma by the coupling of ultrahigh-performance supercritical fluid chromatography and mass spectrometry is investigated with the aim to distinguish patients with breast, kidney, and prostate cancers from healthy controls. The mean sensitivity, specificity, and accuracy of the lipid profiling approach were 85%, 95%, and 92% for kidney cancer; 91%, 97%, and 94% for breast cancer; and 87%, 95%, and 92% for prostate cancer. No association of statistical models with tumor stage is observed. The statistically most significant lipid species for the differentiation of cancer types studied are CE 16:0, Cer 42:1, LPC 18:2, PC 36:2, PC 36:3, SM 32:1, and SM 41:1 These seven lipids represent a potential biomarker panel for kidney, breast, and prostate cancer screening, but a further verification step in a prospective study has to be performed to verify clinical utility.

Cancer incidence and mortality are increasing worldwide as a result of population aging and changing patterns of other risk factors[1]. Malignant disorders are categorized according to the site where tumor growth has started, regardless of subsequent metastatic spread to other parts of the body[2]. Prostate cancer is the second most frequently diagnosed cancer in men[3]. Breast cancer is the most frequently diagnosed cancer and one of the main causes of cancer-related death in women[4]. On the other hand, kidney cancer is the 9th most common cancer in men and the 14th most common cancer in females[5]. However, there is a marked geographical variation in the incidence rate with the highest incidence rates of kidney cancer for men in the Czech Republic among all European countries[6]. The first tests to detect prostate cancer include the levels of prostate specific antigen (PSA) in peripheral blood and the digital rectal examination (DRE). In case of abnormal DRE or elevated PSA levels, a transrectal ultrasound-guided biopsy is performed for verification[7]. Mammography represents the principal method for detecting breast cancer screening, but the diagnosis is commonly supported by other imaging methods including magnetic resonance imaging, positron emission tomography, computed tomography, or single-photon emission computed tomography[8]. Kidney cancer is often discovered by chance during examination with imaging methods for another purpose[9], such as ultrasound, computed tomography, or magnetic resonance imaging[10].

Generally, all imaging methods are subject to the limitation that very small tumors are not properly visualized, resulting in low sensitivity for early stage[11]. Staging examinations are performed after cancer diagnosis based on information regarding the location, spread, and extent of the tumor[12]. The diagnosis and treatment of patients in the early stage of cancer increase the chance for survival and cure compared to patients diagnosed at the late stage.

Cancer screening methods aim at the early detection of cancer for high-risk individuals[13]. Nowadays, more attention is devoted to the development of cancer screening methods based on the examination of peripheral blood, including liquid biopsy. The analysis of circulating cells, platelets, extracellular vesicles, mRNA, miRNA, proteins, cell-free DNA (cfDNA), and circulating tumor DNA (ctDNA) in blood are investigated as potential approaches in cancer screening[14]. Recently, the successful detection of eight cancer types based on the analysis of proteins and mutations in ctDNA in blood was reported[15]. Metabolomics also attracts research attention in

[1]Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 573, 532 10 Pardubice, Czech Republic. [2]Department of Oncology, Faculty of Medicine and Dentistry, Palacký University and University Hospital, I.P. Pavlova 6, 775 20 Olomouc, Czech Republic. [3]Present address: Comprehensive Cancer Center Nový Jičín, Hospital Nový Jičín, Nový Jičín, Czech Republic. ✉email: Michal.Holcapek@upce.cz
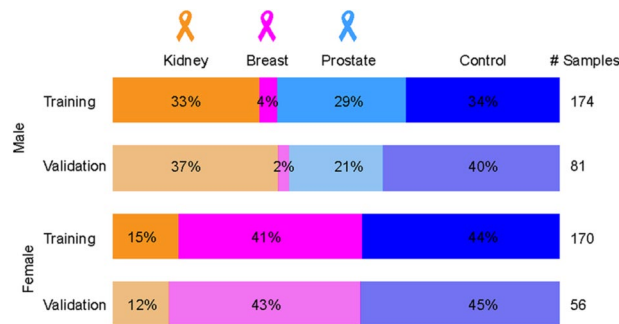
Figure 1. Overview of the sample set (n = 481) used for the study of 3 cancer types. Samples were divided into training (75%) and validation sets (25%). Plasma samples obtained from patients with kidney (n = 119), breast (n = 103) and prostate (n = 67) cancers and healthy controls (n = 192) were included in the study.

cancer screening[16] and other metabolic disorders[17]. Lipidomics can be considered as a part of metabolomics that deals with the comprehensive analysis of lipids, which are important biomolecules involved in many biological processes, such as signaling molecules or constituents of cell membranes, energy storage, and various other metabolic pathways[18]. Clinical lipidomics revealed that plasma lipid concentrations can be changed for various malignant diseases[19,20].

Here, our aim is to quantitatively determine the plasma lipidome of patients with kidney, breast, and prostate cancer and compare it with the lipidome of healthy controls. Lipidomic changes are investigated and visualized using various multivariate data analysis (MDA) tools, such as principal component analysis (PCA), orthogonal projection to latent structures discriminant analysis (OPLS-DA), and receiver operating characteristics (ROC). The ultrahigh-performance supercritical fluid chromatography—mass spectrometry (UHPSFC/MS) is used as a powerful high-throughput and sensitive method for quantitative lipidomic analysis based on the lipid class separation approach recommended for reliable quantitation together with the use of exogenous internal standards for individual lipid classes[21,22].

Results

Study design. Heparin plasma samples from 289 cancer patients and 192 volunteers without the history of previous malignant disease (further referred to as healthy controls) were obtained. Patients were diagnosed with breast, prostate, or kidney cancer based on standard medical procedures at the University Hospital in Olomouc.

The sample set was divided into training and validation sets, whereby about 25% of samples for each cancer type and healthy controls were assigned to the validation set. Finally, the training set included 135 healthy controls, 209 cancer patients (77 breast, 82 kidney, and 50 prostate cancers), and the validation set included 57 healthy controls and 80 cancer patients (26 breast, 37 kidney, and 17 prostate cancers). The overview of all samples together with the clinical information is summarized in Fig. 1 and Supplementary Tables S1 and S2. The average age of healthy volunteers was lower than that of cancer patients, and the average body mass index (BMI) was comparable for both sample groups. Cancer patients are classified according to the TNM system. The majority of samples are assigned as T1 stage, typically for breast cancer (59%) and kidney cancer (47%), while T2 stage is predominant for prostate cancer (69%).

Previous studies reported differences in plasma lipidome depending on gender[23–27]. As a consequence, the gender effect on the prediction performance using MDA for both genders and gender-separated models was evaluated (Supplementary Fig. S1) using OPLS-DA models. The accuracy was slightly higher for gender-separated models, in particular, for females. Therefore, the sample set was divided according to gender. Obviously, prostate cancer occurs only in men and the overwhelming majority of breast cancer patients are women, so the gender separation is an important issue only for kidney cancer, where this study has 73% of men and 27% of women samples.

Discovery phase measured by UHPSFC/MS. The order of samples was randomized separately for the extraction and UHPSFC/MS measurements to exclude any possible biases. The lipidomic analysis of human plasma in the discovery phase resulted in the quantitation of 138 lipids (Supplementary Table S3a) belonging to glycerolipids, glycerophospholipids, and sphingolipids.

Non-supervised PCA and supervised OPLS-DA were applied for all training set samples to visualize differences between sample groups (healthy controls and cancer patients) for the three cancer types studied (Fig. 2A–D). MDA allows the prediction of samples to belong to a particular sample group. The samples of the validation set were predicted by the corresponding OPLS-DA model built on the samples of the training set samples (Supplementary Tables S4 and S5). The specificity, sensitivity, and accuracy values for individual models were the following: kidney cancer males—91%, 73%, and 82%; kidney cancer females—88%, 71%, and 84%; breast cancer females—88%, 63%, and 76%; prostate cancer—90%, 82%, and 87%. Specificity and sensitivity values depending on the cancer stage and accuracy values depending on the cancer type are summarized in Fig. 3 for training and validation sets. The prediction performance is only slightly higher for the training set than for the validation set, which is an important confirmation that the statistical models do not collapse for the prediction
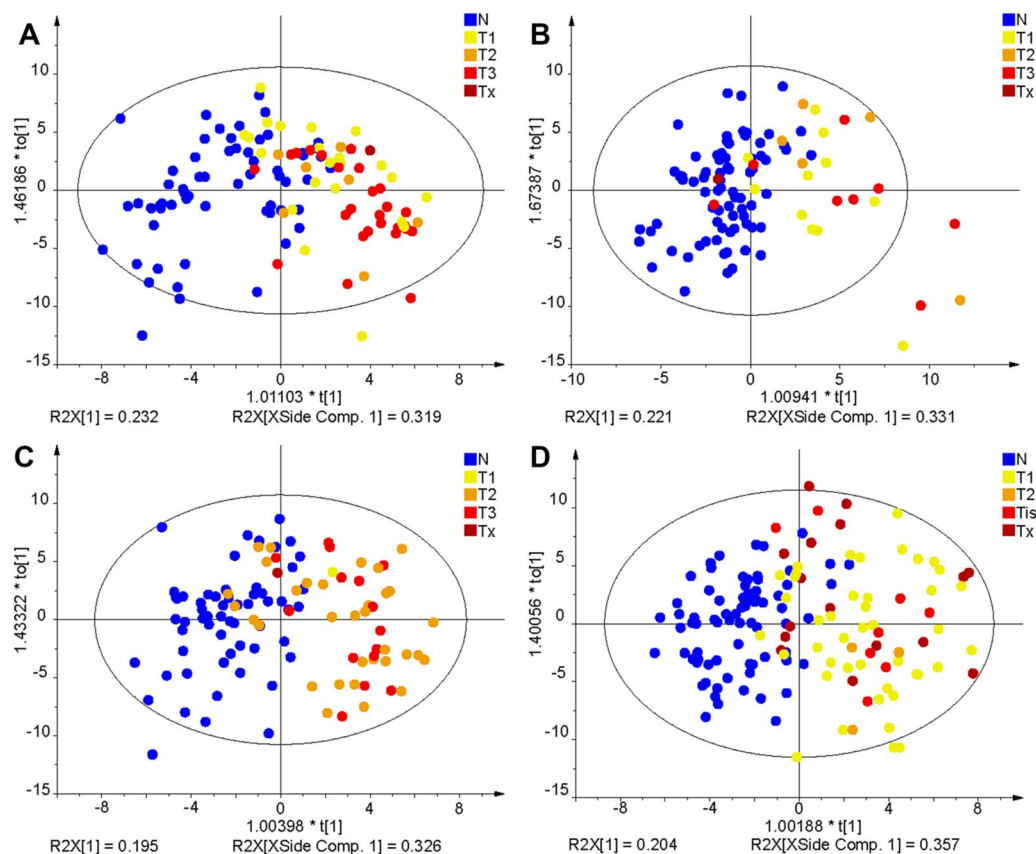
Figure 2. OPLS-DA models used for the differentiation of cancer and control plasma samples. The training set from the discovery phase was measured by UHPSFC/MS and then used to build OPLS-DA models: (**A**) kidney cancer versus control samples for males, (**B**) kidney cancer versus control samples for females, (**C**) prostate cancer versus control samples for males, and (**D**) breast cancer versus control samples for females. Annotation: blue—healthy controls (N), yellow—cancer stage T1 (T1), orange—cancer stage T2 (T2), light red—cancer stage T3 or Tis (T3/Tis), and dark red—unknown cancer stage (Tx).

of samples with unknown classification. ROC curves for the diagnostic ability to classify healthy control and cancer patient samples are illustrated in Fig. 3I–L for individual cancer types. The AUC values for the different types of cancer types ranged from 0.917 to 0.967 for the training set and from 0.868 to 0.953 for the validation set.

Qualification phase measured by UHPSFC/MS. UHPSFC/MS measurements were repeated after several months to verify the repeatability of results, and these repeated measurements are called as the qualification phase. The same sample extracts were measured using different sample measurement sequences to minimize the risk of hidden biases. In total, 138 lipids were also quantified in the qualification phase (Supplementary Table S6a), with 126 of 138 lipids (91%) quantified in both the discovery and qualification phase. The small difference is caused mainly by low abundant short fatty acyl glycerolipids close to the limit of quantitation. PCA score plots for UHPSFC/MS measurements were compared in Fig. 4A,B. Both data sets show the quality control (QC) sample (green) cluster in the PCA score plot, indicating satisfactory method stability during the measurement sequence. The partial group separation between cancer (red) and control (blue) groups is already observed in non-supervised PCA score plots, which confirms a high reproducibility of the lipidomic profiling, as illustrated by numbers of selected samples in Fig. 4A,B. The first and second data sets were compared by calculating the relative standard deviation (RSD) for each lipid in each sample (Supplementary Table S7). In total, 65% of all values have RSD < 20%, and the average of all RSD for each lipid and all samples is 19%. Figure 4C–E further illustrates the reproducibility of quantitative results for selected dysregulated lipids in the discovery and qualification phases, as the medians of box plots for the first and second measurements are comparable. Furthermore, the box plots also show that the selected lipid species are downregulated in all cancer types compared to the control group (Fig. 4C–E). MDA was also applied for repeated measurements using the training sample set for building models. Generally, the prediction performance was comparable for both phases (summary in Supplementary Tables S4 and S5). ROC curves are shown in Fig. 5A–D for all cancer types. AUC values ranged from 0.888 to 0.994. Furthermore, MDA models for the discovery phase were used to predict the sample set of the qualification phase (Supplementary Tables S4 and S8). Specificity, sensitivity, and accuracy values for individual models were the following: kidney cancer males—62%, 91%, and 76%; kidney cancer males—86%, 72%, and
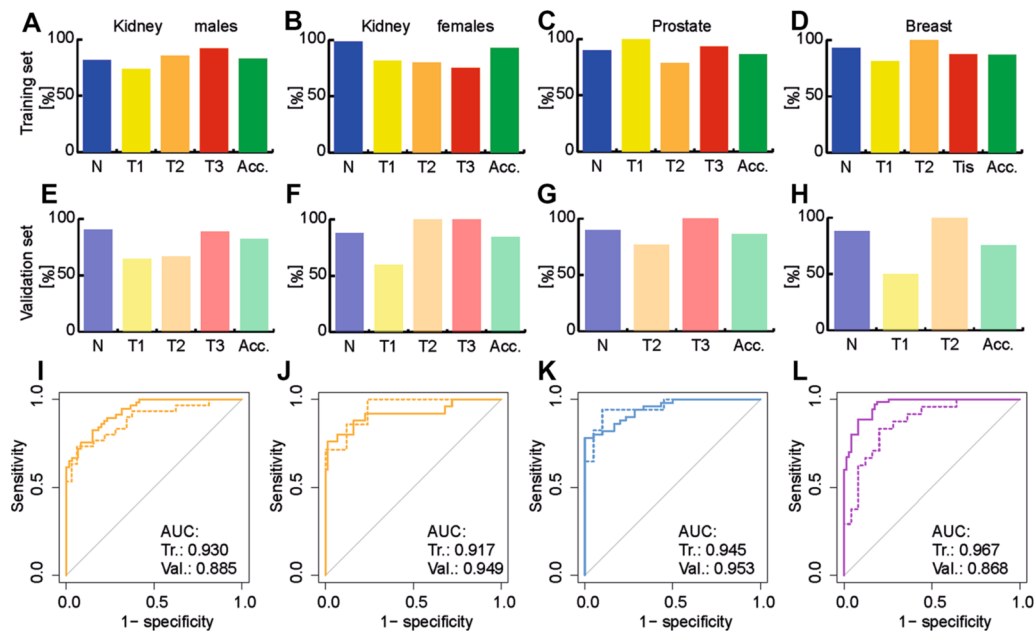
Figure 3. OPLS-DA models were used to predict the pathological state of human subjects. The training set was used to build OPLS-DA models. The percentage of specificity (blue), sensitivity (yellow, orange, red), and accuracy (green) for training and validation sets using UHPSFC/MS data from the discovery phase are presented. The sensitivity was determined for each stage of cancer (yellow—T1, orange—T2, and red—T3), excluding samples with unknown cancer stage. Training set: (**A**) kidney cancer versus healthy control for males, (**B**) kidney cancer versus healthy control for females, (**C**) prostate cancer versus healthy control for males, and (**D**) breast cancer versus healthy control for females. Validation set: (**E**) kidney cancer versus healthy control for males, (**F**) kidney cancer versus healthy control for females, (**G**) prostate cancer versus healthy control for males, and (**H**) breast cancer versus healthy control for females. ROC curves with the corresponding AUC values are presented, where the continuous lines represent the ROC curve for the training set, and dashed lines for the validation set. (**I**) kidney cancer versus healthy control for males, (**J**) kidney cancer versus healthy control for females, (**K**) prostate cancer versus healthy control for males, and (**L**) breast cancer versus healthy control for females.

83%; breast cancer females—94%, 57%, and 76%; prostate cancer—88%, 75%, and 82%. ROC curves are summarized for all samples in Fig. 5E–H for all types of cancer. AUC values ranged from 0.864 to 0.901.

Shotgun MS measurements. To further verify the conclusions of UHPSFC/MS measurements, all samples were also measured by shotgun MS as an independent alternative technique. 412 lipid species were quantified by shotgun MS for the categories of glycerolipid, glycerophospholipid, and sphingolipid (Supplementary Table S6c), which is about 3 times more quantified lipid species compared to UHPSFC/MS. The MDA prediction performance was comparable (Supplementary Tables S4 and S5). ROC curves for training and validation sets are summarized in Fig. 5I–L for all types of cancer. AUC values ranged from 0.841 to 1.00. For two of the most significantly regulated lipid species, the box plots are presented in Fig. 5M,N for all pathological states and data sets for both phases of UHPSFC/MS (1, 2) and shotgun MS (SG).

Influence of the number of lipid species on the prediction capability of statistical models. The influence of the number of quantified lipid species used for MDA on the accuracy to correctly classify samples was investigated. First, only lipid species common for both phases with UHPSFC/MS and shotgun MS were selected. In total, 91 common lipid species were used for MDA in each data set. The overall prediction performance was comparable to that obtained when all quantified lipid species were used for MDA, independent of the method and diagnosis (Supplementary Tables S4 and S9). The number of lipid species in MDA models was further reduced by considering additional statistical criteria, such as fold change (more than ± 20%), p-value (< 0.05), and VIP value (> 1). Supplementary Table S10 provide information how variables were reduced. The whole data set was divided into 5 data subsets (healthy control vs. kidney cancer samples for males and females, healthy control vs. breast cancer samples for males and females, and healthy control vs. prostate cancer samples for males) for UHPSFC/MS (1st and 2nd measurements) and shotgun MS, which resulted in 15 data subsets. In total, 29 lipid species were statistically significant after the Bonferroni correction for at least 10 from 15 data subsets considering all cancer types, methods, and measurements. The accuracy slightly decreased with decreasing number of lipids used for MDA independent of the investigated cancer type and method (Supplementary Tables S4 and S11). However, as the decrease of the prediction performance was not so pronounced, the effect of further reduction of lipids used for MDA on the prediction performance to correctly assign the sample type was
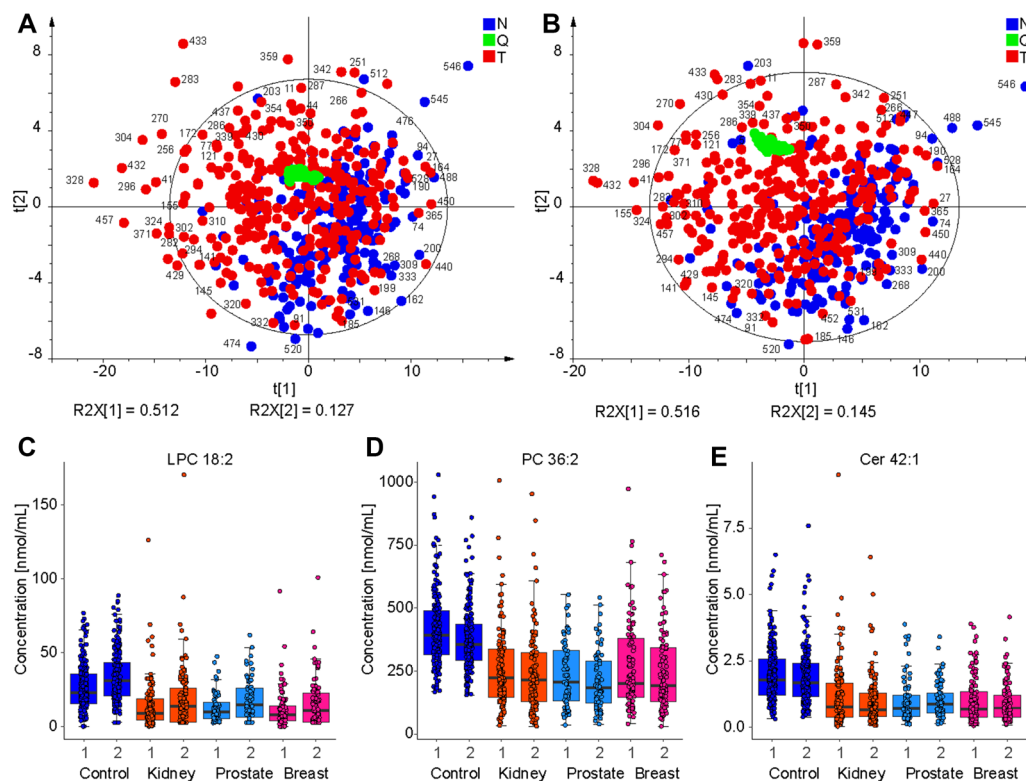
Figure 4. Comparison of UHPSFC/MS results in discovery and qualification phases. (**A**) PCA of all samples (validation and training sets) in the discovery phase. (**B**) PCA of all samples (validation and training sets) in the qualification phase (red—cancer (T), blue—control (N), and green—QC). Selected samples were annotated for comparison purposes. Boxplots comparing the concentrations of samples obtained from patients with different pathological states (blue: control, orange: kidney cancer, light blue: prostate cancer, and pink: breast cancer) for the discovery phase (1) and qualification phase (2) for (**C**) LPC 18:2, (**D**) PC 36:2, (**E**) Cer 42:1 using UHPSFC/ MS.

investigated. CE 16:0, Cer 42:1, LPC 18:2, PC 36:2, PC 36:3, SM 32:1, and SM 41:1 were significant according to the Bonferroni correction for at least 14 of the 15 data subsets considering all cancer types, methods and measurements. MDA was performed for these 7 lipid species and the prediction performance was evaluated (Supplementary Tables S4 and S12). The average of sensitivity, specificity, and accuracy values for the different number of lipids considering all methods and genders was calculated (Fig. 6). Generally, the sensitivity and consequently the accuracy decreased with decreasing number of lipids used for MDA. The specificity was not affected by the number of lipid species, independent of the cancer type (Fig. 6). No effect of the cancer stage on concentrations of the most significant lipid species was observed for all types of cancer (Supplementary Fig. S2).

Statistical evaluation of data. The different plasma lipidomic profiles depending on cancer type were investigated by evaluating statistically significant lipid species after Bonferroni correction, lipid species with a fold change of $\pm 20\%$, and VIP value > 1. The percentage of lipid species belonging to the lipid class fulfilling the defined criteria was calculated, as illustrated by the pie charts for different cancer types in Fig. 7A–D. Nonpolar lipid species, triacylglycerols and cholesterol esters, are of greater relevance in kidney cancer, while the influence of glycerophospholipids and sphingolipids appears to be dominant in breast and prostate cancer.

The most significant lipid species for all methods and data sets are downregulated in plasma samples of cancer patients, independent of the cancer type, as illustrated in Fig. 7E–G. MDA was used to investigate the differences between healthy control samples and different types of cancer types for all quantified lipid species, 91 lipid species for UHPSFC/MS and shotgun MS, 29, and finally, 7 most significant lipid species for all data sets. OPLS-DA models for samples obtained from healthy male controls and male patients suffering from prostate and kidney cancer as well as samples obtained from healthy female controls and female patients suffering from kidney and breast cancer are shown in Fig. 8A,B. The question was whether differentiation and prediction of cancer type and healthy control samples are possible using UHPSFC/MS. The specificity ranged from 67 to 97% with the average of 83%, sensitivity for kidney cancer from 49 to 74% with the average of 61%, sensitivity for prostate from 0 to 66% with the average of 43% and the accuracy from 57 to 74% with the average of 65% for the training and validation set and different numbers of lipids (138, 91, 29, and 7 lipid species) included to build the MDA models considering samples obtained from male donors (Fig. 8A, Supplementary Table S13). The specificity ranged from 80 to 96% with the mean of 90%, sensitivity for kidney cancer from 0 to 44% with
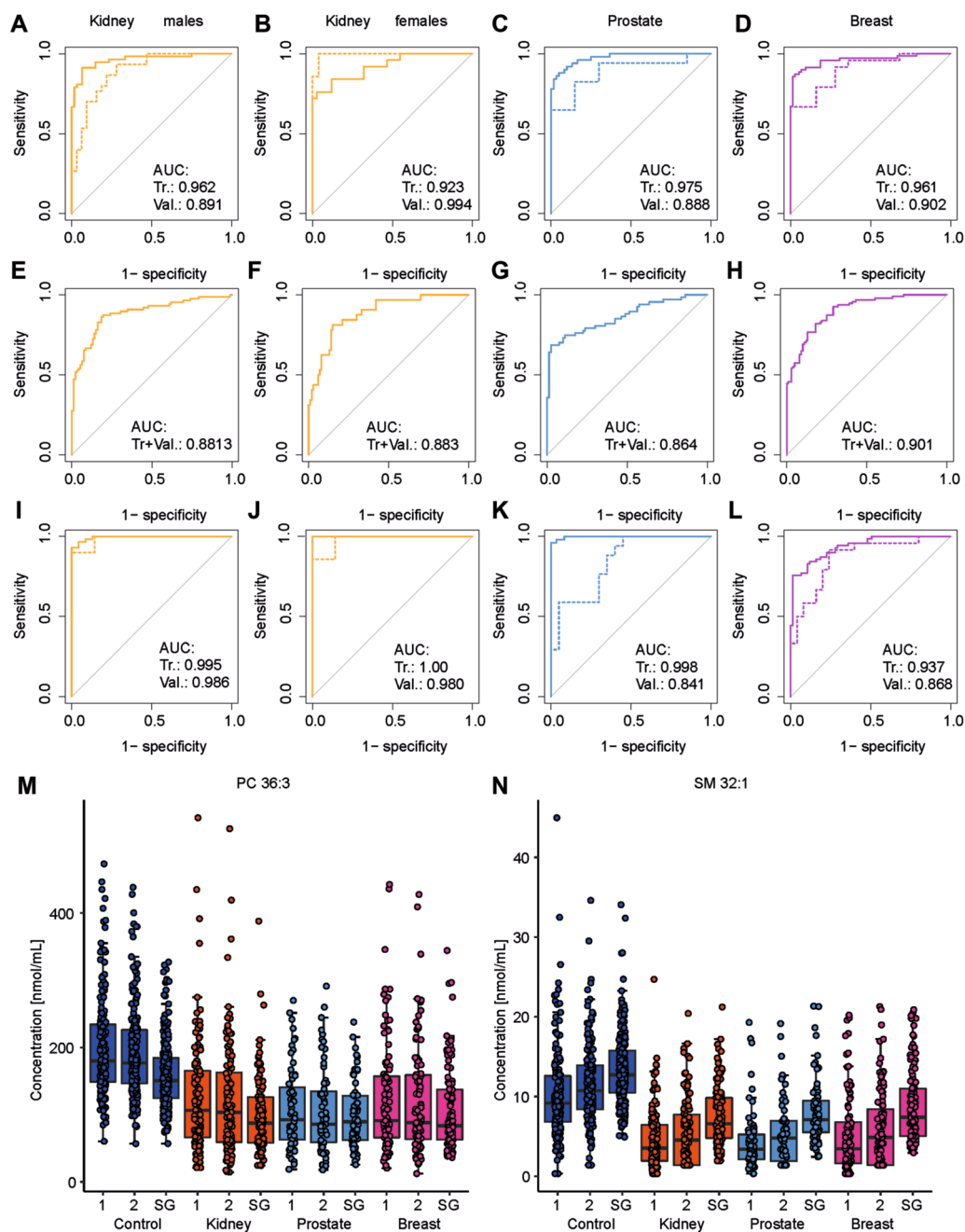
Figure 5. ROC curves with the corresponding AUC values are presented, where the continuous lines represent the ROC curve for the training set and dashed lines for the validation set using UHPSFC/MS. (**A**) kidney cancer versus healthy control for males, (**B**) kidney cancer versus healthy control for females, (**C**) prostate cancer versus healthy control for males, and (**D**) breast cancer versus healthy control for females using data from the qualification phase for MDA, and (**E**) kidney cancer versus healthy control for males, (**F**) kidney cancer versus healthy control for females, (**G**) prostate cancer versus healthy control for males, and (**H**) breast cancer versus healthy control for females predicting the data from the qualification phase using the discovery phase for MDA, (**I**) kidney cancer versus healthy control for males, (**J**) kidney cancer versus healthy control for females, (**K**) prostate cancer versus healthy control for males, and (**L**) breast cancer versus healthy control for females using shotgun MS data for MDA. Box plots comparing concentrations of samples obtained from patients with different pathological states (blue: healthy control, orange: kidney cancer, light blue: prostate cancer, and pink: breast cancer) for discovery (1) and qualification (2) phases using UHPSFC/MS and shotgun MS (SG): (**M**) PC 36:3, and (**N**) SM 32:1.
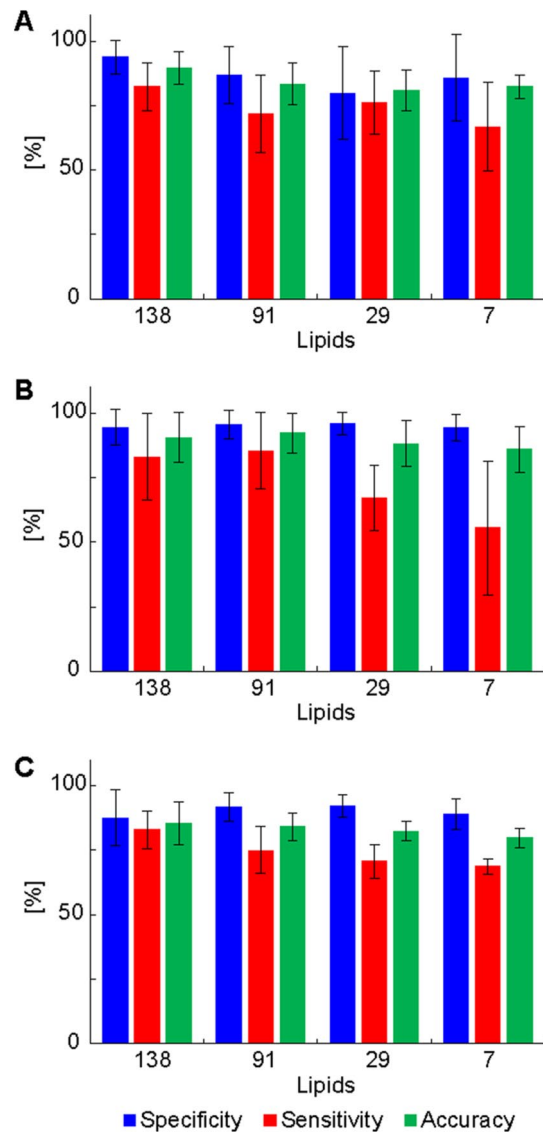
Figure 6. Influence of decreased number of lipid species used for MDA using UHPSFC/MS data in the discovery phase. (138 lipids: no exclusion, 91 lipids: common lipids from the discovery and qualification phase using UHPSFC/MS and shotgun MS, 29: only lipid species included, which are significant according to the Bonferroni correction for at least 10 from 15 models. SM 38:1 and SM 42:1 were excluded, because these variables were only significant for UHPSFC/MS data sets (10/10) and therefore a method bias cannot be excluded. 7: only lipid species included, which are significant according to the Bonferroni correction for at least 14 from 15 models. The average of the specificity (blue), sensitivity (red), and accuracy (green) for validation and training set as well as for both genders were calculated. (**A**) kidney cancer, (**B**) breast cancer, and (**C**) prostate cancer.

the mean of 24%, sensitivity for breast cancer from 60 to 83% with the mean of 73% and the accuracy from 66 to 78% with the mean of 74% for the training and validation set and different numbers of lipids (138, 91, and 29 lipid species) included to build the MDA models considering samples obtained from female donors (Fig. 8B). It was not possible to perform the MDA model using 7 lipid species as variables due to the insufficient number of components for samples obtained from female donors.

The differentiation of the cancer type was also investigated by performing OPLS-DA models that classify kidney cancer versus prostate cancer for males (Fig. 8C) and kidney cancer versus breast cancer for females (Fig. 8D and Supplementary Table S14). OPLS-DA models were evaluated using 138 and 91 lipids as variables for males and 138, 91, and 29 for females using UHPSFC/MS data, since for the lower number of lipids a lack of components was observed. The sensitivity for prostate cancer was 71–88% with the mean of 81% and for kidney cancer 57–82% with the mean of 72% for the training and validation set and both UHPSFC/MS data sets considering male samples. The sensitivity for breast cancer ranged between 94 and 100% with the mean of 98%,
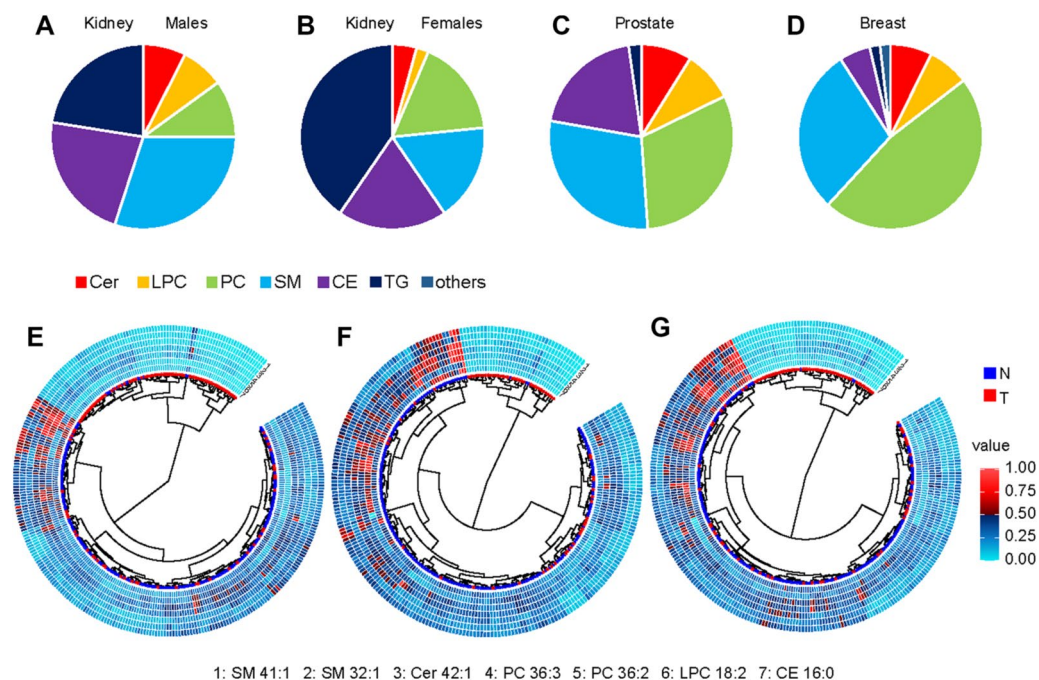
Figure 7. Distribution of lipid class percentage for individually lipid classes, which are significant according to the Bonferroni correction for the discovery phase: (**A**) kidney cancer males, (**B**) kidney cancer females, (**C**) prostate cancer males, and (**D**) breast cancer females. Dendrograms for 7 lipid species quantified in the discovery phase using UHPSFC/MS for both genders and training and validation sets for: (**E**) kidney cancer, (**F**) prostate cancer, and (**G**) breast cancer.

the sensitivity for kidney cancer ranged between 14 and 86% with the mean of 57%, and the accuracy ranged between 80 and 97% with the mean of 88% for the training and validation set and both UHPSFC/MS data sets considering female samples.

Discussion

Cancer screening as part of a regular health examination can allow early cancer detection and timely treatment, resulting in improved clinical outcomes. Circulating biomarker measurement as a minimally invasive and routinely used method seems to be one of the most attractive and convenient method for screening of high-risk individuals. Current approaches focus on the analysis of genetic mutations, ctDNA, or proteins for early cancer diagnosis in plasma or serum. To date, the clinical utility of lipidomic analysis for this purpose has not been clearly demonstrated.

In the present study, we performed quantitative lipidomics of human plasma samples collected from healthy controls and cancer patients by UHPSFC/MS. We paid special emphasis to the accurate molar quantitation of lipid species allowing the future interlaboratory comparison of the results if the same measurement protocol is applied.

MDA revealed the applicability of lipidomics as a diagnostic tool for all three cancer types studied. The performance of classification models in cancer prediction was characterized by high sensitivity, specificity, and accuracy. Nonpolar lipids, such as cholesteryl esters and triacylglycerols, are more important for kidney cancer, while differences in sphingolipid and glycerophospholipid profiles are more pronounced in breast and prostate cancers. The reduction in the number of quantified lipid species used for MDA showed only a slight loss in sensitivity, specificity, and accuracy. Nevertheless, the decrease in method complexity compared to the overall lipidomic profiling could facilitate potential clinical use.

Our results are consistent with previous reports on plasma or serum alterations in patients with different types of cancer[19], including breast cancer[28–30], pancreatic cancer[23,31], kidney cancer[32], lung cancer[33,34], and prostate cancer[35,36]. We observed downregulation of multiple plasma lipid species in patients compared to healthy volunteers. Previous reports also showed the association of hypolipidemia with some malignancies[37], but the peripheral blood lipidome may also be affected by other factors[38].

Based on previous literature[39,40], we hypothesize that the observed alterations in lipid concentrations are the overall result of complex processes in the human body, including the accumulation of lipids in plasma to favor tumor growth. Malignant cell proliferation requires excess lipids to build membranes, organelles, and participate in signaling processes[41]. The statistical analysis showed that the following seven lipids downregulated in patients' plasma had the most significant effect on the differentiation between cases and controls: CE 16:0, Cer 42:1, LPC 18:2, PC 36:2, PC 36:3, SM 32:1, and SM 41:1. CE 16:0 may be a potential source of palmitic acid for cancer cells. Palmitic acid can be converted to phospholipids, sphingolipids, glycerolipids, and other fatty acids essential for cancer cell survival. Louie et al. showed that cancer cells and tumors robustly incorporate
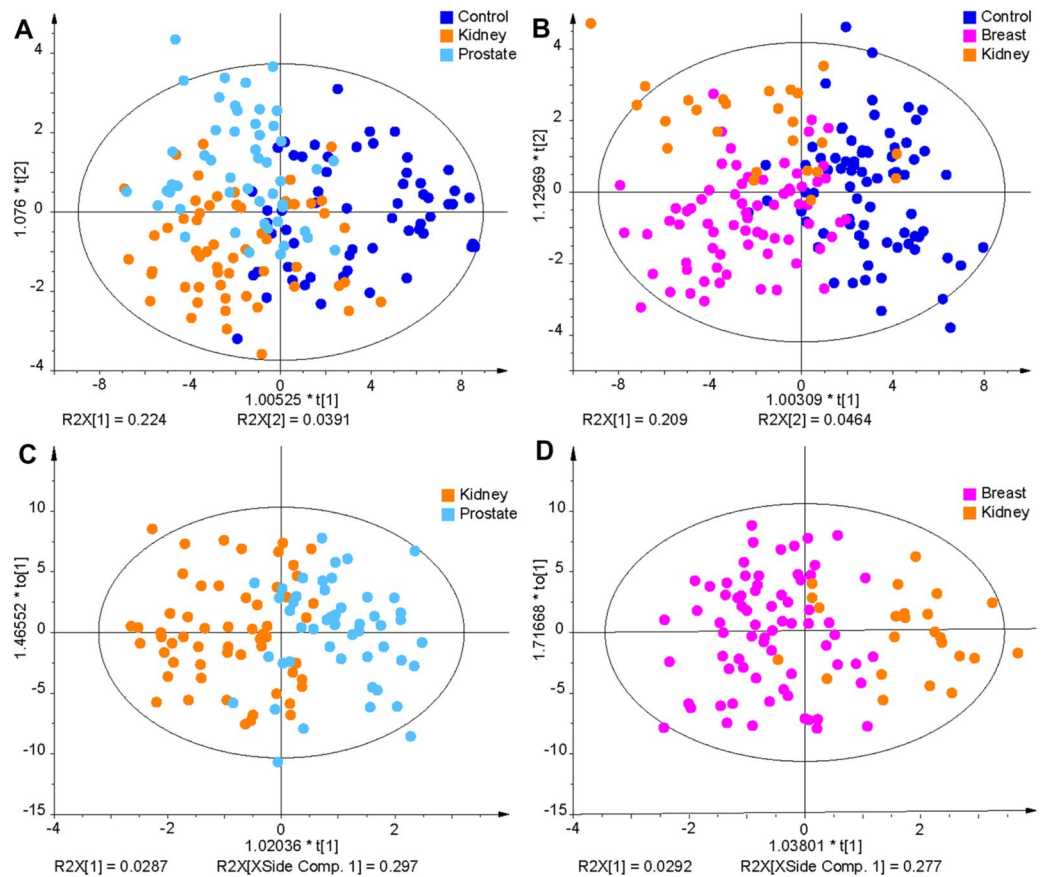
Figure 8. OPLS-DA models for the differentiation of the sample type like cancer type and control samples (**A**) males and (**B**) females using the concentrations of the 138 lipids determined in the discovery phase with UHPSFC/MS for the training set. OPLS-DA models for the differentiation of (**C**) prostate and kidney cancer samples for males and (**D**) breast and kidney cancer samples for females.

exogenous palmitic acid and remodel it into other oncogenic lipid species[42]. Sphingolipids play an essential role in signal transduction pathways that regulate cell growth/death, migration, and senescence. Cancer cells can dysregulate enzymes involved in the metabolism of sphingolipids, resulting in the suppression of apoptosis, e.g., through downregulation of pro-cell death sphingolipids, such as ceramides[43,44]. LPC act as bioactive proinflammatory signaling molecules[45]. LPC are also substrates for LPA synthesis, and subsequently, LPA can modulate the immunological response and promote tumor cell growth[46]. PC are major components of membranes necessary for cell proliferation and survival. The alteration in PC metabolism is a potential signature of tumor progression and could be a good target for therapy[47]. The molecular biology techniques or MS-based approaches with stable isotope-labeled metabolites in future biological studies could shed more light on the lipid metabolism in cancer. Current data based on the analysis of lipids in human plasma could not explain the complete biological mechanism of ongoing processes in cancer.

Some previous studies have reported dysregulation of lipids in several diseases[17,48–50]. Unfortunately, molar concentrations are often not provided for measured lipids, therefore, the correlation of their conclusions with other studies is difficult. In any case, the future prospective study should include a group of cases with other nonmalignant conditions to determine whether the lipidomic analysis has the clinical utility for their differentiation from not only from healthy controls, but also from other diseases. It will also be worth investigating the discrimination potential of seven lipids versus the broader lipidomic profiling. In the present study, the accuracy decrease for cancer patients versus healthy controls is relatively low, but the comparison with patients with other nonmalignant diseases has not yet been performed.

In conclusion, the present data indicate the potential of lipid profiling in cancer screening, at least for breast, kidney, and prostate cancers. The use of individual MDA models to distinguish healthy control samples and the single cancer type results in higher accuracy than MDA models that include multiple cancer types. The use of IS for each lipid class allows the quantitation of lipid species and the comparison of lipid concentrations between different laboratories. Subsequent prospective studies are necessary for seven lipid species identified as potential biomarkers for cancer screening.

Methods

Human samples.
A retrospective study was performed on 481 human plasma samples. A total of 192 control samples and 289 cancer samples from patients with breast, kidney, or prostate cancer were collected. The criteria for healthy controls were that they did not have any type of cancer during the life time and age over 18 years. For cancer patients, the disease was histologically confirmed by needle biopsy or by examining the surgical resection specimen. Both cancer patients and healthy controls were of the same ethnicity (Caucasian), collected in the same place (University Hospital in Olomouc) and processed in the same way. No other exclusion criteria were applied. The clinical information for all patients and controls is summarized in Fig. 1 and Supplementary Tables S1 and S2. The sample set was divided into training (using to build OPLS-DA models) and validation (indicates the possible use for samples with unknown classification) sets. Each fourth sample was assigned to the validation set, to obtain a distribution of the 75% of samples belonging to the training set and 25% of the samples to the validation set. Patients had no treatment before blood collection. Human plasma was collected in 9 mL lithium-heparin collection tubes and then centrifuged. The supernatant was transferred, aliquoted, and stored at − 80 °C until further processing for lipidomic analysis.

Ethics declaration.
The study was approved by the ethical committee of the University Hospital Olomouc. All subjects signed an informed consent. All methods were carried out in line with Ethical Principles for Medical Research Involving Human Subjects (Declaration of Helsinki).

Study phases.
The lipidome of 481 plasma samples was measured by UHPSFC/MS in the discovery phase. To ensure that UHPSFC/MS results are reproducible, the same extracts were measured again several months later corresponding to the qualification phase. The sequence of sample measurements was randomized to exclude any measurement bias. The data set was independently processed and the results were compared to the discovery phase. Furthermore, the extracts were also measured with shotgun MS to exclude any bias caused by the employed method, independently processed, and compared with UHPSFC/MS results.

Chemicals.
Solvents for analysis, such as acetonitrile, 2-propanol, methanol (HPLC/MS grade), water (UHPLC/MS grade), and hexane, were purchased from Honeywell (Riedel-da Haën, CHROMASOLV™ LC–MS Ultra, Hamburg, Germany), distributed by Fisher Scientific (Waltham, Massachusetts, USA). Chloroform stabilized with 0.5–1% ethanol was purchased from Sigma-Aldrich (St. Louis, MO, USA) or Merck (Darmstadt, Germany), respectively. Ammonium acetate was purchased from Fisher Scientific. Deionized water for liquid–liquid extraction was obtained from a Milli-Q Reference Water Purification System (Molsheim, France). Carbon dioxide of 4.5 grade (99.995%) was purchased from Messer Group (Bad Soden, Germany). Non-endogenous lipids were used as internal standards (IS) for quantitative analysis, i.e., MG 19:1/0:0/0:0, DG 12:1/0:0/12:1, and TG 19:1/19:1/19:1 from Nu-ChekPrep (Elysian, MN, USA); CE 16:0 D7, Cer d18:1/12:0, cholesterol D7, LPC 17:0/0:0, LPE 14:0/0:0, PC 14:0/14:0, PC 22:1/22:1, PE 14:0/14:0, PI 15:0/18:1 D7, SM d18:1/12:0, PS 14:0/14:0, PA 14:0/14:0, PG 14:0/14:0, LPG 14:0/0:0, HexCer d18:1/12:0, Hex2Cer d18:1/12:0, and SHexCer d18:1/12:0 from Avanti Polar Lipids (Alabaster, AL, USA). The concentrations of stock solutions of individual IS and the volumes needed to prepare the IS mixture are summarized in Supplementary Table S15.

Lipidomic analysis.
For the extraction of lipids, a modified Folch procedure was employed, which was previously validated[51]. The same sample extracts were analyzed with UHPSFC/MS and shotgun MS. Human serum (25 μL) and the mixture of IS (17.5 μL) were homogenized in 3 mL of chloroform/methanol (2:1, v/v) for 10 min in an ultrasonic bath (40 °C). When the samples reached ambient temperature, 600 μL of water was added, and the mixture was vortexed for 1 min. After 3 min of centrifugation (3000 rpm), the aqueous layer was removed, and the organic layer was evaporated under a gentle stream of nitrogen. The residue was dissolved in a mixture of 500 μL of chloroform/2-propanol (1:1, v/v), carefully vortexed and filtered (0.2 μm syringe filter). The extract was diluted 1:20 with the mixture of hexane/2-propanol/chloroform (7:1.5:1.5, v/v/v) for UHPSFC/MS analysis and 1:8 with chloroform/methanol/2-propanol (1:2:4, v/v/v) mixture containing 7.5 mM of ammonium acetate and 1% of acetic acid for shotgun MS analysis.

UHPSFC/MS measurements were carried out on an Acquity Ultra Performance Convergence Chromatography (UPC2) system hyphenated to the hybrid quadrupole traveling wave ion mobility time-of-flight mass spectrometer Synapt G2-Si from Waters using the commercial interface kit (Waters, Milford, MA, USA). Chromatographic settings were used with minor improvements from the previously published method[51,52]. UHPSFC analyses were measured on the Viridis BEH column (100 × 3 mm, 1.7 μm) using a linear gradient with supercritical CO_2 and as a modifier, MeOH with 30 mM ammonium acetate and 1% water: 0 min—1% modifier, 5 min—51% modifier, 6.5 min—51% modifier, 6.8 min—1% modifier. The total run time including the equilibration was 7.5 min. The automatic back-pressure regulator was set to 1800 psi, the column temperature to 60 °C, the flow to 1.9 mL/min, and the injection volume was 1 μL. The injection needle was washed after each injection with hexane/2-propanol/water (2:2:1, v/v/v). The make-up solvent was MeOH with 30 mM ammonium acetate and 1% water with the flow rate of 0.25 mL/min. The following parameters were set for MS measurements: the positive ion electrospray ionization (ESI) in the sensitivity mode, the mass range of m/z 150–1200, the capillary voltage of 3 kV, the sampling cone of 20 V, the source offset of 90 V, the source temperature of 150 °C, the desolvation temperature of 500 °C, the cone gas flow of 50 L/h, the desolvation gas flow of 1000 L/h, and the nebulizer gas flow of 4 bar. The analysis was done in the continuum mode with a scan time of 0.1 s and the lock mass scanning. Leucine enkephalin peptide was used as a lock mass with the scan time of 0.1 s and the interval of 30 s interval. The lock mass was scanned but the mass correction was not automatically applied. All samples were measured in duplicates. Noise reduction was performed on the raw files using the Waters compression tool.

Data files were lock mass corrected and converted into centroid data using the exact mass measure tool from Waters. The MarkerLynx software from Waters was used for data preprocessing. Further data processing was done by LipidQuant 1.0 software[53].

Shotgun experiments were performed on a 6500 QTRAP quadrupole linear ion trap mass spectrometer (Sciex, Concord, ON, Canada) equipped with an ESI probe using the characteristic precursor ion (PIS) and neutral loss (NL) scan events[54]. Raw data files were processed with the Sciex LipidView Software from Sciex in order to obtain a summary table of m/z versus intensity for each scan mode (NL and PIS) of all samples. Raw data were prefiltered by applying the following settings in the positive ion mode, a tolerance mass window of 0.5 Da, a minimum intensity threshold of 0.1%, and a minimum signal-to-noise ratio of 3 after smoothing. The summary tables of m/z versus intensity for all samples were exported as txt files and further processed by the LipidQuant 1.0 software.

Data processing.
LipidQuant 1.0 is a Microsoft Excel based script used for the automated data processing of txt files[53] including m/z values versus intensities for all samples. The experimental m/z values were compared with the theoretical m/z values from the embedded database for lipid identification, depending on the retention time window or scan type defining the lipid class. The lipid quantitation was performed by calculating the ratio of the intensities of the target lipid and the internal standard and multiplying with the known concentration of the internal standard. Isotopic correction type II[55] was automatically applied and a summary table containing lipid concentrations in all samples was generated. Zero filling for missing values was applied by setting the number for 80% of the minimum measured concentration for a given lipid species for all samples. If the concentration was not determined for more than 25% of the samples, then the lipid species was excluded from the data set. The data set was divided into training and validation set by assigning each fourth sample to the validation set. Clinical information for the samples, like gender and pathological state, was revealed, and samples were assigned. Final tables containing the lipid concentrations for all samples and fulfilling all defined criteria were used for MDA and other statistical tools.

Statistical analysis.
MDA was performed with SIMCA software, version 13.0 (Umetrics, Sweden). The lipid species were defined as variables, and the samples as observations. The data set was preprocessed by applying logarithmic transformation, pareto scaling, and centering. Data preprocessing should facilitate the normal distribution of lipid concentrations and that low abundant lipid species contribute similarly to the MDA as high abundant lipid species. PCA was performed to evaluate for outliers, estimate measurement quality by checking the clustering of QC samples, and evaluate the clustering of sample groups depending on the pathological state. OPLS-DA is a statistical tool for visualizing differences between sample groups of known classification. OPLS-DA was built using the training set and then used for the sample prediction of the validation set. For both PCA and OPLS-DA, the score scatter plots for the first two components are visualized, although more components may contribute to the model. The number of components for PCA and OPLS-DA models was determined by selecting the option autofit in the SIMCA software, where only components are considered of significance according to cross-validation rules. The cross-validation is automatically applied following Eastment et al. for PCA[56] and Martens et al. for OPLS-DA[57]. The data set is divided into seven groups, omitting one group, building the model and predicting the excluded group. This is repeated for each group, and the results of predictions reveal the number of significant components, which is provided in Supplementary Table S4. OPLS-DA revealed differences in lipidome by using gender as a classifier. As a consequence, the data sets for females and males were treated separately for investigation of the prediction performance.

Microsoft Excel was used for the calculation of average lipid concentrations obtained for all sample groups, fold change, T-value, and p-value. For the calculation of p-value, a two-sided two-sample T-test assumed unequal variances (Welch test) for the samples obtained from healthy controls and patients with kidney, breast, or prostate cancer. p-values < 0.05 were considered as significant, but p-values were further evaluated according to the Bonferroni correction. All statistical parameters for all lipids are summarized in Supplementary Tables S3 and S6 below the lipid concentrations measured in individual samples and Supplementary Table S10. Another parameter indicating some relevance to differentiate samples from healthy controls and cancer patients, is the variable of importance (VIP) value obtained for each OPLS-DA plot. The most regulated and statistically significant lipid species with a fold change $\pm 20\%$, a p-value < 0.05, and a VIP value > 1 for all methods and phases are summarized in Supplementary Table S10. Box plots were used to better visualize lipid species concentrations depending on the health state. The box plots were constructed in R free software environment (https://www.r-project.org) using readxl and ggplot2 packages. In each box plot, the median was presented by a horizontal line, the box represented the first and the 3rd quartile values, and whiskers stood for 1.5*IQR from the median, and each measurement was plotted as a jittered point value. The receiver operating characteristics curves were generated by using the packages readxl and AUC in R. The dendrograms were also constructed in R[58]. For the circular dendrograms, the Euclidean distances were calculated, and then the upgma function from the phangorn library was used for clustering (the Ward agglomeration method was selected). Circular dendrograms were generated and surrounded by the heatmap (ggtree and gheatmap functions – ggtree library). For the presentation of the heatmap, all concentrations were min–max scaled.

Data availability
All data relevant for the conclusions presented conclusions are provided in the manuscript or in the supplementary tables. Raw files of all measurements can be provided on request from the corresponding author.

References

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424. https://doi.org/10.3322/caac.21492 (2018).
2. National Cancer Institute. Cancer staging. Cancer.gov. https://www.cancer.gov/about-cancer/understanding/what-is-cancer. Updated 9 Feb 2021.
3. Culp, M. B., Soerjomataram, I., Efstathiou, J. A., Bray, F. & Jemal, A. Recent global patterns in prostate cancer incidence and mortality rates. *Eur. Urol.* **77**, 38–52. https://doi.org/10.1016/j.eururo.2019.08.005 (2020).
4. Ahmad, A. Breast Cancer Statistics: Recent Trends In *Breast Cancer Metastasis and Drug Resistance: Challenges and Progress* (ed. Ahmad, A.) 1–7 (Springer, 2019).
5. Znaor, A., Lortet-Tieulent, J., Laversanne, M., Jemal, A. & Bray, F. International variations and trends in renal cell carcinoma incidence and mortality. *Eur. Urol.* **67**, 519–530. https://doi.org/10.1016/j.eururo.2014.10.002 (2015).
6. Li, P. *et al.* Regional geographic variations in kidney cancer incidence rates in European countries. *Eur. Urol.* **67**, 1134–1141. https://doi.org/10.1016/j.eururo.2014.11.001 (2015).
7. Descotes, J.-L. Diagnosis of prostate cancer. *Asian J. Urol.* **6**, 129–136. https://doi.org/10.1016/j.ajur.2018.11.007 (2019).
8. Jafari, S. H. *et al.* Breast cancer diagnosis: Imaging techniques and biochemical markers. *J. Cell. Physiol.* **233**, 5200–5213. https://doi.org/10.1002/jcp.26379 (2018).
9. Capitanio, U. & Montorsi, F. Renal cancer. *Lancet* **387**, 894–906. https://doi.org/10.1016/S0140-6736(15)00046-X (2016).
10. Ljungberg, B. *et al.* EAU guidelines on renal cell carcinoma: the 2010 update. *Eur. Urol.* **58**, 398–406. https://doi.org/10.1016/j.eururo.2010.06.032 (2010).
11. Lennon, A. M. *et al.* Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. *Science* **369**, eabb9601. https://doi.org/10.1126/science.abb9601 (2020).
12. Amin, M. B. *et al.* The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J. Clin.* **67**, 93–99. https://doi.org/10.3322/caac.21388 (2017).
13. Tanos, R. & Thierry, A. R. Clinical relevance of liquid biopsy for cancer screening. *Transl. Cancer Res.* **7**, S105–S129. https://doi.org/10.21037/tcr.2018.01.31 (2018).
14. Chen, M. & Zhao, H. Y. Next-generation sequencing in liquid biopsy: Cancer screening and early detection. *Hum. Genomics* **13**, 10. https://doi.org/10.1186/s40246-019-0220-8 (2019).
15. Cohen, J. D. *et al.* Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* **359**, 926. https://doi.org/10.1126/science.aar3247 (2018).
16. Zhang, L., Han, X. & Wang, X. Is the clinical lipidomics a potential goldmine?. *Cell Biol. Toxicol.* **34**, 421–423. https://doi.org/10.1007/s10565-018-9441-1 (2018).
17. Kordalewska, M. *et al.* Multiplatform metabolomics provides insight into the molecular basis of chronic kidney disease. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **1117**, 49–57. https://doi.org/10.1016/j.jchromb.2019.04.003 (2019).
18. Yang, K. & Han, X. Lipidomics: Techniques, applications, and outcomes related to biomedical sciences. *Trends Biochem. Sci.* **41**, 954–969. https://doi.org/10.1016/j.tibs.2016.08.010 (2016).
19. Wolrab, D., Jirásko, R., Chocholoušková, M., Peterka, O. & Holčapek, M. Oncolipidomics: Mass spectrometric quantitation of lipids in cancer research. *Trends Anal. Chem.: TrAC* **120**, 115480. https://doi.org/10.1016/j.trac.2019.04.012 (2019).
20. Buszewska-Forajta, M. *et al.* Lipidomics as a diagnostic tool for prostate cancer. *Cancers* https://doi.org/10.3390/cancers13092000 (2021).
21. Holčapek, M., Liebisch, G. & Ekroos, K. Lipidomic analysis. *Anal. Chem.* **90**, 4249–4257. https://doi.org/10.1021/acs.analchem.7b05395 (2018).
22. Liebisch, G. *et al.* Lipidomics needs more standarization. *Nat. Metab.* **1**, 745–747. https://doi.org/10.1038/s42255-019-0094-z (2019).
23. Wolrab, D. *et al.* Lipidomic profiling of human serum enables detection of pancreatic cancer *Nat. Com.*, revision, preprint available at medRxiv: https://www.medrxiv.org/content/10.1101/2021.01.22.21249767v1 (2021).
24. Holčapek, M., Cífková, E., Lísa, M., Jirásko, R., Wolrab, D., Hrnčiarová, T. *A Method of Diagnosing Pancreatic Cancer Based on Lipidomic Analysis of a Body Fluid.* European patent EP3514545 (granted), date of filling 22.1.2018.
25. Holčapek, M., Wolrab, D., Jirásko, R., Cífková, E. *A Method of Diagnosing Cancer Based on Lipidomic Analysis of a Body Fluid.* EP18174963.1 (pending), filing date 29.5.2018.
26. Sales, S. *et al.* Gender, contraceptives and individual metabolic predisposition shape a healthy plasma lipidome. *Sci. Rep.* **6**, 27710. https://doi.org/10.1038/srep27710 (2016).
27. West, A. L. *et al.* Lipidomic analysis of plasma from healthy men and women shows phospholipid class and molecular species differences between sexes. *Lipids* **56**, 229–242. https://doi.org/10.1002/lipd.12293 (2021).
28. Chen, X. *et al.* Plasma lipidomics profiling identified lipid biomarkers in distinguishing early-stage breast cancer from benign lesions. *Oncotarget* **7**, 36622–36631. https://doi.org/10.18632/oncotarget.9124 (2016).
29. Cala, M. P. *et al.* Multiplatform plasma metabolic and lipid fingerprinting of breast cancer: A pilot control-case study in Colombian Hispanic women. *PLoS ONE* **13**, e0190958. https://doi.org/10.1371/journal.pone.0190958 (2018).
30. Cui, M., Wang, Q. & Chen, G. Serum metabolomics analysis reveals changes in signaling lipids in breast cancer patients. *Biomed. Chromatogr.* **30**, 42–47. https://doi.org/10.1002/bmc.3556 (2016).
31. Guo, Y. *et al.* Simultaneous Quantification of Serum Multi-Phospholipids as Potential Biomarkers for Differentiating Different Pathophysiological states of lung, stomach, intestine, and pancreas. *J. Cancer* **8**, 2191–2204. https://doi.org/10.7150/jca.19128 (2017).
32. Lin, L. *et al.* LC-MS-based serum metabolic profiling for genitourinary cancer classification and cancer type-specific biomarker discovery. *Proteomics* **12**, 2238–2246. https://doi.org/10.1002/pmic.201200016 (2012).
33. Yu, Z. *et al.* Global lipidomics identified plasma lipids as novel biomarkers for early detection of lung cancer. *Oncotarget* **8**, 107899–107906. https://doi.org/10.18632/oncotarget.22391 (2017).
34. Guo, Y. *et al.* Probing gender-specific lipid metabolites and diagnostic biomarkers for lung cancer using Fourier transform ion cyclotron resonance mass spectrometry. *Clin. Chim. Acta* **414**, 135–141. https://doi.org/10.1016/j.cca.2012.08.010 (2012).
35. Zhou, X. *et al.* Identification of plasma lipid biomarkers for prostate cancer by lipidomics and bioinformatics. *PLoS ONE* **7**, e48889. https://doi.org/10.1371/journal.pone.0048889 (2012).
36. Patel, N., Vogel, R., Chandra-Kuntal, K., Glasgow, W. & Kelavkar, U. A novel three serum phospholipid panel differentiates normal individuals from those with prostate cancer. *PLoS ONE* **9**, e88841. https://doi.org/10.1371/journal.pone.0088841 (2014).
37. Bielecka-Dbrowa, A., Hannam, S., Rysz, J. & Banach, M. Malignancy-associated dyslipidemia. *Open Cardiovasc. Med. J.* https://doi.org/10.2174/1874192401105010035 (2011).
38. Bergheanu, S. C. *et al.* Lipidomic approach to evaluate rosuvastatin and atorvastatin at various dosages: Investigating differential effects among statins. *Curr. Med. Res. Opin.* **24**, 2477–2487. https://doi.org/10.1185/03007990802321709 (2008).

39. Tiwary, S., Berzofsky, J. A. & Terabe, M. Altered lipid tumor environment and its potential effects on NKT cell function in tumor immunity. *Front. Immun.* https://doi.org/10.3389/fimmu.2019.02187 (2019).
40. Gil-De-Gómez, L., Balgoma, D. & Montero, O. Lipidomic-based advances in diagnosis and modulation of immune response to cancer. *Metabolites* **10**, 1–15. https://doi.org/10.3390/metabo10080332 (2020).
41. Guo, R. *et al.* The function and mechanism of lipid molecules and their roles in the diagnosis and prognosis of breast cancer. *Molecules* https://doi.org/10.3390/molecules25204864 (2020).
42. Louie, S. M., Roberts, L. S., Mulvihill, M. M., Luo, K. & Nomura, D. K. Cancer cells incorporate and remodel exogenous palmitate into structural and oncogenic signaling lipids. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1566–1572**, 2013. https://doi.org/10.1016/j.bbalip.2013.07.008 (1831).
43. Cui, M., Wang, Q. & Chen, G. Serum metabolomics analysis reveals changes in signaling lipids in breast cancer patients. *Biomed. Chromatogr.* **30**, 42–47. https://doi.org/10.1002/bmc.3556 (2016).
44. Sheridan, M. & Ogretmen, B. The role of ceramide metabolism and signaling in the regulation of mitophagy and cancer therapy. *Cancers* https://doi.org/10.3390/cancers13102475 (2021).
45. Law, S. H. *et al.* An updated review of lysophosphatidylcholine metabolism in human diseases. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms20051149 (2019).
46. Rolin, J. & Maghazachi, A. A. Effects of lysophospholipids on tumor microenvironment. *Cancer Microenviron.* **4**, 393–403. https://doi.org/10.1007/s12307-011-0088-1 (2011).
47. Ahn, H. S. *et al.* Convergence of plasma metabolomics and proteomics analysis to discover signatures of high-grade serous ovarian cancer. *Cancers* **12**, 1–20. https://doi.org/10.3390/cancers12113447 (2020).
48. Liu, J. *et al.* Serum metabolomic patterns in patients with autoimmune thyroid disease. *Endocr. Pract.* **26**, 82–96. https://doi.org/10.4158/EP-2019-0162 (2020).
49. Kartsoli, S., Kostara, C. E., Tsimihodimos, V., Bairaktari, E. T. & Christodoulou, D. K. Lipidomics in non-alcoholic fatty liver disease. *World J. Hepatol.* **12**, 436–450. https://doi.org/10.4254/wjh.v12.i8.436 (2020).
50. Marczak, L. *et al.* Mass spectrometry-based lipidomics reveals differential changes in the accumulated lipid classes in chronic kidney disease. *Metabolites* https://doi.org/10.3390/metabo11050275 (2021).
51. Wolrab, D., Chocholoušková, M., Jirásko, R., Peterka, O. & Holčapek, M. Validation of lipidomic analysis of human plasma and serum by supercritical fluid chromatography–mass spectrometry and hydrophilic interaction liquid chromatography–mass spectrometry. *Anal. Bioanal. Chem.* **412**, 2375–2388. https://doi.org/10.1007/s00216-020-02473-3 (2020).
52. Lísa, M. & Holčapek, M. High-throughput and comprehensive lipidomic analysis using ultrahigh-performance supercritical fluid chromatography-mass spectrometry. *Anal. Chem.* **87**, 7187–7195. https://doi.org/10.1021/acs.analchem.5b01054 (2015).
53. Wolrab, D. *et al.* LipidQuant tool for automated data processing in lipid class separation-mass spectrometry workflows. *Bioinformatics* https://doi.org/10.1093/bioinformatics/btab644 (2021).
54. Lísa, M., Cífková, E., Khalikova, M., Ovčačíková, M. & Holčapek, M. Lipidomic analysis of biological samples: Comparison of liquid chromatography, supercritical fluid chromatography and direct infusion mass spectrometry methods. *J. Chromatogr. A* **1525**, 96–108. https://doi.org/10.1016/j.chroma.2017.10.022 (2017).
55. Wang, M., Wang, Ch. & Han, X. Selection of internal standards for accurate quantification of complex lipid species in biological extracts by electrospray ionization mass spectrometry—What, how and why?. *Mass Spectrom. Rev.* **36**, 693–714. https://doi.org/10.1002/mas.21492 (2016).
56. Eastment, H. T. & Krzanowski, W. J. Cross-validatory choice of the number of components from a principal component analysis. *Technometrics* **24**, 73–77. https://doi.org/10.2307/1267581 (1982).
57. Martens, H. & Naes, T. M. v. C., Wiley, Chichester (1989).
58. Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinform.* **69**, e96. https://doi.org/10.1002/cpbi.96 (2020).

Acknowledgements

Author contributions

D.W., R.J., and M.H prepared the concept of the study. D.W. and O.P. performed the sample preparation. D.W. analyzed samples by UHPSFC/MS and I.B. by shotgun MS, D.W. processed data, D.W. and J.I. performed statistical analysis. H.S. and B.M. obtained and provided plasma samples and clinical information. D.W. prepared the first draft of the manuscript and Figures. M.H. was responsible for funding and supervision of this study. All co-authors read, reviewed, edited, and approved the manuscript.

Competing interests

M.H., R.J., and D.W. are listed as inventors on the patent A method of diagnosing cancer based on lipidomic analysis of a body fluid (EP18174963.1, filing date 29. 5. 2018) related to this work.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-99586-1.

Correspondence and requests for materials should be addressed to M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.