Serum Peptidome Profiling Revealed Platelet Factor 4 as a Potential Discriminating Peptide Associated With Pancreatic Cancer

Authors

Georg Martin Fiedler[1*], Alexander Benedikt Leichtle[1*], Julia Kase[1*], Sven Baumann[1], Uta Ceglarek[1], Klaus Felix[3], Tim Conrad[2], Helmut Witzigmann[4], Arved Weimann[5], Christof Schütte[2], Johann Hauss[4], Markus Büchler[3], Joachim Thiery[1]

Authors´ Affiliations

1) Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University Hospital Leipzig, Liebigstr. 27, D-04103 Leipzig, Germany
2) Department of Mathematics, Free University of Berlin, Arnimallee 6, D-14195 Berlin, Germany
3) Department of General Surgery, University of Heidelberg, Im Neuenheimer Feld 110, D-69120 Heidelberg, Germany
4) Clinic of Visceral Surgery, University Hospital Leipzig, Liebigstr. 20, D-04103 Leipzig, Germany
5) Clinic of General Visceral Surgery, Hospital St. Georg Leipzig, Delitzscher Str. 141, D-04129 Leipzig, Germany

[*] These authors contributed equally to this study

Requests for reprints

Joachim Thiery, M.D.

Professor of Medicine, Head

Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics

University Hospital Leipzig

Medical Faculty, University Leipzig

Liebigstr. 27, D-04103 Leipzig

Germany

Phone: +49341-9722200

Fax: +49341-9722379

Email: thiery@medizin.uni-leipzig.de

Grant support

This work was supported by a grant from the BMBF/SAB and by a grant from Microsoft Research, Cambridge to Joachim Thiery, and by a "formel-1 grant" of the Medical Faculty of the University Leipzig to Sven Baumann and Alexander Leichtle.

Running Title

PF4 as a supporting marker of pancreatic cancer.

Key Words

Peptidome profiling, pancreatic cancer, platelet factor 4, MALDI-TOF MS, biomarker

Translational Relevance

Based on MALDI-TOF peptidome profiling, the present study identifies platelet factor 4 (PF4) as a discriminating serum marker in pancreatic cancer patients. In conjunction with the conventional tumor markers CA 19-9 and CEA, PF4 strongly improves the diagnostic power of tumor marker testing. This might be of special relevance in the differential diagnosis of pancreatic cancer and pancreatitis.

Abstract

Purpose: Mass spectrometry-based serum peptidome profiling is a promising tool to identify novel disease-associated biomarkers, but is limited by preanalytical factors and the intricacies of complex data processing. Therefore, we investigated whether standardized sample protocols and new bioinformatic tools combined with external data validation improve the validity of peptidome profiling for the discovery of pancreatic cancer associated serum markers.

Experimental Design: For discovery study, two sets of sera from patients with pancreatic cancer (n=40) and healthy controls (n=40) were obtained from two different clinical centers. For external data validation, we collected an independent set of samples from patients (n=20) and healthy controls (n=20). Magnetic beads (MB) with different surface functionalities were used for peptidome fractionation followed by MALDI-TOF MS. Data evaluation was carried out comparing two different bioinformatic strategies. Following proteome database search the matching candidate peptide was verified by MALDI-TOF MS after specific antibody-based immunoaffinity chromatography and independently confirmed by an ELISA assay.

Results: Two significant peaks (m/z 3884; 5959) achieved a sensitivity of 86.3% and specificity of 97.6% for the discrimination of patients and healthy controls in the external validation set. Adding peak m/z 3884 to conventional clinical tumor markers (CA 19-9 and CEA) improved sensitivity and specificity as shown by ROC analysis ($AUROC_{combined}$=1.00). Mass spectrometry based m/z 3884 peak identification and following immunological quantitation revealed platelet factor 4 as the corresponding peptide.

Conclusions: MALDI-TOF MS based serum peptidome profiling allowed the discovery and validation of platelet factor 4 as a new discriminating marker in pancreatic cancer.

Introduction

Pancreatic cancer is the fourth leading cause of cancer death in the United States. Most patients diagnosed with pancreatic cancer develop clinical symptoms late in the course of the disease (1). Therefore, only 20% will be amenable to potentially curative therapy and only 3% to 5% of patients survive 5 years or more (2). Earlier diagnosis of the disease and early relapse monitoring are probably the best available options to improve patient survival (3). Currently, no single clinical chemical marker meets the sensitivity and specificity criteria required for screening or stratification purposes (4). Established serum markers such as Carbohydrate Antigen CA 19-9 or Carcinoembryonic Antigen (CEA) are useful to monitor the course of disease on and off treatment, but they lack the prerequisites for screening and to estimate the prognosis of a patient (2, 5).

Peptidome-based studies using high-throughput spectrometric methods such as matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) promise to be valuable for the identification of new ‘disease signatures‘ and cancer associated biomarkers, especially combined with hitherto known biomarker patterns in a multivariate approach (6-13). However, there is a controversy regarding the diagnostic potential and reliability of the clinical proteomics approach (14-18). It was perceived, that standardization of preanalytical and analytical factors as well as improvements in bioinformatic tools are important preconditions for translating serum peptidomics from bench to bedside (16, 19-21). For instance, it was shown that many preanalytical factors have major impact on the results of biomarker discovery and limit the use of pre-existing sample banks (22-26). Recently, we developed a standardized sample protocols and new bioinformatic tools for spectral data pre-processing and peak selection to enhance the sensitivity of data analysis (proteomics.net) (22, 27).

The present clinical study investigated the impact of our standardized sample protocol and new bioinformatic tools combined with external data validation on the efficiency of MALDI-TOF based

peptidome profiling for the discovery and clinical replication of novel serum markers of pancreatic cancer.

Materials and Methods

Patients and samples

A total of 120 patients with pancreatic cancer and controls were recruited for this study. For the discovery study sera were obtained from two different clinical centers (University Hospital Leipzig (UHL, set A) and Heidelberg (UHH, set B)). Consequently, we obtained two sets from patients with pancreatic cancer (Ap (n=20), Bp (n=20)) and two sets of healthy controls (Ac (n=20), Bc (n=20)). Following finalization the discovery study, additional blood samples of patients with pancreatic cancer and of healthy controls were collected at UHL for independent external validation (Cp (n=20) and Cc (n=20)). Subjects were adjusted according to age and gender (Table 1). Additionally, serum samples of 26 patients suffering from acute pancreatitis were collected as inflammatory control group.

Blood sampling from patients was performed before initiation of specific therapy. Diagnosis of pancreatic cancer was confirmed by histological examination in all cases. Healthy controls showed no evidence of actual disease proven by physical examination and routine laboratory testing (differentials, CRP, creatinine, transaminases, alkaline phosphatise, γ-glutamyl transferase, bilirubin, tumor markers (CA 19-9, CEA)). Serum samples were collected and stored (-80°C) by standardized techniques and protocol (22).

The study was approved by the local ethics committees and fulfils the requirements of the Helsinki declaration. All subjects gave informed consent to participate in the study.

Chemicals, standards and consumables

Gradient grade acetonitrile, ethanol, and HPLC-water were obtained from J.T. Baker (Phillipsburg, NJ, USA); p.a. trifluoroacetic acid (TFA) and acetone were purchased from Sigma-Aldrich (Taufkirchen, Germany). The peptide- and protein MALDI-TOF calibration standards I and α-cyano-4-hydroxycinnamic acid (HCCA) were purchased from Bruker Daltonics (Bremen, Germany). Automated magnetic bead preparations were performed using 96 well plates, TubePlates from Biozym (Hessisch Oldendorf, Germany), polypropylene tubes (low profile) from Abgene (Surrey, UK), and

modular reservoir quarter modules from Beckman (Fullerton, USA). For sample storage 450 µL CryoTubes™ were purchased from Sarstedt (Nümbrecht, Germany). Multifly needle sets and polypropylene serum monovettes with clotting activators were also obtained from Sarstedt.

Peptidome Separation

All serum samples of the discovery set were processed at one time and analyzed simultaneously to avoid procedure-dependent errors. The external validation set was prepared, processed and analyzed separately.

Peptidome separation of the samples was performed using the ClinPro Tools profiling purification kits from Bruker Daltonics. Magnetic particles with defined surface functionalities (magnetic bead-immobilized metal ion affinity chromatography (MB-IMAC Cu), magnetic bead-hydrophobic interaction (MB-HIC C8) and weak cation exchange (MB-WCX)) were processed by the ClinPro Tools liquid handling robot according to the manufacturer's protocol (Bruker Daltonics). Serum specimens were thawed on ice for 30 min and immediately processed according to our standardized protocol for serum peptidomics (22).

Mass Spectrometry

A linear MALDI-TOF mass spectrometer (Autoflex I, Bruker Daltonics) was used for the peptidome profiling. Daily mass calibration was performed using the standard calibration mixture of peptides and proteins in a mass range of 1-10 kDa. Mass spectra were recorded and processed using AutoXecute tool of the flexControl acquisition software (Ver. 2.0; Bruker Daltonics).

For candidate biomarker discovery using MS/MS analysis a MALDI-TOF/TOF mass spectrometer (Ultraflex III, Bruker Daltonics) was used.

Mass Spectrometry Data Analysis

Bioinformatic processing

For data analysis two bioinformatic tools were applied, Bruker Daltonics ClinPro Tools (CPT) 2.0.365 Software and proteomics.net, a novel statistical driven pre-processing, peak-finding and analyzing pipeline (27). Mass spectra were generated equally for CPT and proteomics.net. Device-dependent raw data served as data base for CPT. For proteomics.net data were transformed into device-independent raw data using an ASCII-Converter (ASCII=American Standard Code for Information Interchange) provided by Bruker Daltonics.

The CPT workflow starts by spectra loading of two selected classes (e.g. pancreatic cancer and healthy controls of one set). The used CPT software package includes an automated raw data pre-treatment workflow, comprising baseline subtraction with 80% baseline flatness (Convex Hull), normalization of spectra according to the total ion count, an alignment of peaks with a signal-to-noise ratio (S/N) > 3 to prominent peaks with S/N > 100 and a peak picking procedure resulting in peaks defined as dynamic m/z ranges. Savitzky-Golay smoothing was deactivated to avoid the blurring of peaks (CPT manual Version 2.0, Bruker Daltonics). Peak statistics are performed using Welch's t-test without multiple testing correction. Finally, the software provides a list of peaks sorted along statistical difference between two classes, which was used for further data analysis (28, 29).

Furthermore, we introduced proteomics.net as a novel web-based bioinformatic platform for peptidome data analysis in cooperation with the Department for Mathematics of the Free University Berlin and Microsoft Research (Cambridge, UK) (27). The fundamental difference to former approaches is based on the omission of post-analysis spectra alignment and of S/N-cutoffs. This results in an up to 10-fold improvement of detection sensitivity (27). Following single spectra processing proteomics.net-software creates a master-peak list by detecting clusters using statistical distribution analysis. Single spectra are matched to the master-peak list and the peak-features are derived. For candidate creation, we selected peaks with significant feature (via e.g. Jensen-Shannon divergence) in patient and control groups. P-values were calculated for every significant peak feature using an Extreme Values Distribution method. Terminal classification of spectra was done by established

classification algorithms (e.g. Support Vector Machine from WEKA). An internal validation (20%-leave-out internal 5 cycle cross-validation) was realized to avoid model-specific overfitting.

Following data analysis and peak calculation we obtained peak lists according to statistical significance for every study set and both software tools (Figure 1). Since CPT performs no correction of the multiple testing problem, we finally focused on the 50 best-discriminating peaks.

Data analysis

In the discovery study each disease group were cross-tested against both control groups to create lists of candidate markers applying both bioinformatic approaches (Figure 1). Subsequently, we filtered discriminating candidate peaks present in all comparisons. In the following external validation study we applied patterns of these candidate peaks alone and in combination with the tumor markers CA 19-9 and CEA to classify the independent samples of the external validation set (Table 1 (C)). Supplementary, we performed ROC analyses using the SPSS Software (Version 15, SPSS Inc., Chicago IL, USA) to detect the gain in sensitivity and specificity introduced by the additional candidate peaks.

In silico search of candidate biomarkers

An in silico search against SwissProt database was performed using TagIdent tool from ExPASy Proteomics Server with following criteria: m/z 7767.6, error tolerance 350 ppm, taxonomy: Homo sapiens.

Antibody based confirmation of MB-MALDI-TOF MS based marker

Antibody capture beads MB-IAC Prot G (Magnetic Bead based Immunoaffinity Chromatography on immobilized Protein G) (Bruker Daltonics) were used for the antibody-based confirmation of the specific candidate peaks. Selected sample aliquots were prepared according to manufactures instructions. Briefly, 15 µL of the MB-IAC Prot G beads were incubated for 60 min at room

temperature with 10µl anti-PF4 antibodies (conc. 1 µg/µl). Thereafter, 5 µl of a MB-IMAC Cu prepared serum- and a PF4 standard solution were added and incubated for 120 min at room temperature. Samples were eluted using 10 µl elution buffer (EB) and incubated for 20 min at room temperature. MALDI target preparation was performed as mentioned above.

Immunoassay for Quantification of platelet factor 4

Platelet factor 4 (PF4) levels were measured in the serum samples using an Asserachrom PF4 Antibody based ELISA (Roche). The serum samples were diluted 1/2100 v/v with dilution buffer and analyzed according to manufactures instructions on a Tecan MT-plate reader (Tecan, Crailsheim, Germany) at a wavelength of 450 nm.

Immunoassays for Quantification of CA 19-9 and CEA

CA 19-9 and CEA were measured in serum samples by an electrochemiluminescence immunoassay (ECLIA; Roche Diagnostics, Mannheim, Germany) on ’Modular‘ analytics E 170 analyzer (Roche) according to the manufacturer’s instructions.

Results

Detection of discriminating peaks

In the discovery phase 80 samples (subsets A: 20 cancer (Ap) and 20 controls (Ac); subsets B: 20 cancer (Bp) and 20 controls (Bc)) were included for peptidome profiling (Table 1). We obtained approximately 750 signals for each single serum specimen using MB-IMAC-Cu, MB-WCX and MB-HIC C8, respectively. Each sample was 4-fold processed to improve the data reproducibility (22). Therefore, 960 peptidome profiles in the mass range of 1000-10.000 Dalton were generated. In total, about 60.000 mass signals were available for further bioinformatic analysis.

Data processing using ClinPro Tools (CPT)

The cancer subsets Ap and Bp were cross tested against the healthy control subsets Ac and Bc for significant differences in the peptide profile using the Welch´s t-test on the basis of peak areas (Figure 1). As result four candidate lists of the 50 most significant peaks were generated. This was separately done for each of the three bead functionalities. In total 150 promising candidates were selected for each group comparison. In summary, as a result of the discovery phase 600 candidates were selected for further validation. In the next step, all candidates, which could not be found in all of the four candidate lists, were excluded from further evaluation. As results, only seven significant peaks, which fulfilled this criterion, remained. In a following verification procedure, unfortunately, all candidates had to be excluded as discriminators due to their inconstant rectification as shown for m/z 1945 in Figure 2. Consequently, further data processing by genetic algorithm and support vector machine using CPT was omitted.

Proteomics.net

Based on the same design the cancer subsets Ap and Bp were cross tested against the healthy control subsets Ac and Bc for significant differences in the peptide profile using protemics.net. This bioinformatic tool calculates p-values using an Extreme Values Distribution. In the four candidate lists

of each functionality 8 to 19 significant different peaks could be identified (in total 57 peaks) of which six peaks (m/z 1003, 1021, 3194, 3884, 4055, and 5959) were significant in at least two comparisons (Supplemental Table 1). Signals up to an m/z 1500 were excluded from further data analysis due to ion suppression effects of the MALDI matrix resulting in disturbing high background noise. Therefore, only four peaks (m/z 3194, 3884, 4055, and 5959) could be confirmed as discriminating peaks in the following verification procedure. These four peaks were applied for linear SVM classification analysis. The peak pattern m/z 3884 and 5959 (Figure 3) showed the best discriminating power with a sensitivity of 86.3% and specificity of 97.6%, adding a further third or fourth peak did not significantly improve the sensitivity and specificity. Therefore, we decided to limit our further analyses to the peak couple m/z 3884 and m/z 5959.

Models for AUROC analysis

The SVM data were not applicable for AUROC analysis. Therefore, we performed all AUROC calculations and model generations on the quantitative data of the peak heights and serum tumor marker levels using the SPSS 16 statistical package (SPSS Inc., Chicago, IL, USA).

The four model variables of interest were the established tumor markers CA19-9 and CEA as well as the aforementioned peaks m/z 3884 and m/z 5959 as non-discrete quantitative variables for a given patient P as follows:

W(P)= Numerical value of the serum concentration of CEA in ng/mL for patient (P)

X(P)= Numerical value of the serum concentration of CA19-9 in ng/mL for patient (P)

Y(P)= Numerical value of the peak height of m/z 3884 for patient (P)

Z(P)= Numerical value of the peak height of m/z 5959 for patient (P)

We built simple factorial models out of these variables, which resulted in composite scores:

Model for tumor markers alone: score 1 = W(P) * X(P)

Model for peaks alone: score 2 = Y(P) * Z(P)

Model for tumor markers and peaks combined: score 3 = (Y(P) * Z(P)) / (W(P) * X(P))

Model for tumor markers and peak m/z 5959: score 3a = Z(P) / (W(P) * X(P))

Model for tumor markers and peak m/z 3884: score 3b = Y(P) / (W(P) * X(P))

These models were equally applied to the discovery and to the validation set.

AUROCs in the discovery set

Based on these quantitative data of the discovery set, the AUROC of the serum tumor marker concentrations alone (score 1) was 0.925 (95% CI: 0.856-0.994), the AUROC consisting of peak heights of peaks m/z 3884 and m/z 5959 (score 2) was 0.734 (95% CI: 0.620-0.848), and the AUROC of the combination of peak heights and serum tumor marker concentrations (score 3) was 0.960 (95% CI: 0.922-0.997).

AUROCS in the external validation set

The aim of the external replication was to validate the discriminatory power of candidate m/z 3884 and 5959 in an independently collected sample subset of 20 patients with pancreatic cancer (Cp) and 20 healthy controls (Cc) alone and in combination with the established tumor markers CA 19-9 and CEA. Computing receiver-operator-characteristics curves (ROC), CA 19-9 and CEA (score 1) revealed an area under the curve (AUROC) of 0.868. Adding peak m/z 5959 to CA 19-9 and CEA (score 3a) increased the sensitivity, but decreased the specificity resulting in an unimproved AUROC of 0.868. However, adding peak m/z 3884 to CA 19-9 and CEA (score 3b) led to an increase in sensitivity and specificity with an AUROC of 1.0 (Figure 4).

Identification of biomarkers

The MALDI-TOF/TOF MS analysis of the purified MB-IMAC CU eluate revealed besides the signal m/z 3884 a second prominent peak at m/z 7767, which suggests that both peaks are differentially charged ions from the same molecule. However, no sufficient fragment spectra could be obtained for structural identification and database search. High resolution MS proved m/z 3884.3 as the double

charged ion of the signal at m/z 7767.6. Database search using TagIdent tool from SwissProt database revealed platelet factor 4 as the potential underlying peptide. Next, we employed anti-PF4 MB-IAC Prot G particle based MALDI-TOF analysis to confirm m/z 3884 as double charged ion of PF4. As shown in Figure 5 the signals at m/z 7767 as well as m/z 3884 could be unambiguously identified as single and double charged ions of PF4.

Validation of platelet factor 4 by ELISA techniques

For the direct validation and quantification of the identified PF4 in all serum samples of the present study ELISA technique was used. The PF4-concentrations of healthy controls (Median/2.5th/97.5th percentile: 7.3/3.3/13.8 kU/ml) and patients with pancreatic cancer (Median/2.5th/97.5th percentile: 5.6/0.8/12.3 kU/ml) differed significantly ($p=0.001$) and confirmed the MB-MALDI-TOF MS results. Supplementary, 26 serum samples from patients with acute pancreatitis were analyzed as inflammatory controls. The PF4-concentrations of patients with acute pancreatitis (Median/2.5th/97.5th percentile: 8.7/4.6/15.4 kU/ml) were slightly elevated compared to healthy controls and significantly ($p<0.001$) higher compared to patients with pancreatic cancer (Supplemental Figure 1). The AUROCs of immunologically determined PF4 concentrations for the discrimination between healthy and patients with pancreatic cancer was 0.833 (95% CI: 0.725-0.941), for the discrimination between pancreatic cancer and acute pancreatitis 0.829 (95% CI: 0.720-0.938).

Discussion

In this study, we identified and confirmed platelet factor 4 (PF4) as a potential marker peptide for pancreatic cancer using MALDI-TOF MS based clinical serum peptidome profiling with special considering of preanalytical preconditions and bioinformatic intricacies. The additional application of PF4 strongly improves the diagnostic power of the conventional serum tumor marker panels consisting of CA-19-9 and CEA.

Clinical proteomics and peptidomics have rapidly grown over the past years, especially in the discovery of potential biomarkers for cancer diagnosis (6-13). However, missing standardization of preanalytical factors, methodological shortcomings and bioinformatic artifacts led to controversies regarding the applicability of these techniques in clinical settings (14-18). To achieve the objective of extracting true positive marker peptides from a haystack of interference-based candidates, well-designed, bias-free, and prospective investigations are demanded, in which sample collection and storage are highly standardized as well as appropriate bioinformatic tools for analysis of putative informative peptides are applied (10, 30). Therefore, we collected, stored and processed the serum samples of the study according to a feasible and highly standardized preanalytical protocol to minimize any sampling-related disturbances (22). Following MALDI-TOF peptidome analysis, we applied a proprietary bioinformatic software (ClinPro Tools) and a recently developed bioinformatic approach (´protemics net`) for spectra analysis, which promised to improve the detection of candidate peaks (31). Primarily, we used the proprietary ClinPro Tools software with data analysis based on an averaged mass spectrum generation for each sample set. In the discovery phase, we could initially select 600 candidates. However, following cross validation and verification no single candidate marker remained. Secondarily, we applied ‘protemics.net’, which was specially developed for the demands of large scale peptidome profiling studies (27). This bioinformatic tool supports the processing of each single peptidome profile. Therefore, it allows data analysis with respect to the variance and the statistical distribution of each single peak even below the common noise level and avoids artifactual peak findings as well as data overfitting (27). Using the ´proteomics.net` software in total 57 candidate

peaks were obtained in the discovery phase. Following cross-testing and verification four reproducible candidate peptides remained of which peak m/z 3884 and 5959 showed the best discriminating power with a sensitivity of 86.3% and specificity of 97.6%.

To assure the reproducibility of the results, independent validation studies are necessary (17), but often not practical due to the limited access to comparable patient subsets of similar source and possible storage time dependent degradation of the samples (21). Hence, following discovery study, we collected an independent sample set for external validation of our first results adhering to the same preanalytical, analytical, and bioinformatic protocols. Complementing the suggestions proposed by Diamandis (20) and Pepe et al. (32), this procedure provided the possibility to immediately sort out irreproducible peaks and confirm the true-positives. As a result of this external validation, the two peaks m/z 3884 and 5959 could be confirmed as potential candidate markers. Our data also support the requirement of a multi-center study design to detect reproducible candidate markers and to rule out center-specific influences on the peptidome profiles (33).

The diagnostic power of the single peaks m/z 3884 and 5959 as well as their pattern were proved alone and in combination with the conventional tumor markers CA 19-9 and CEA by performing ROC analysis. The selectivity of the conventional tumor marker model (AUROC 0.868) was comparable to previous data (34). Introducing peak m/z 5959 sensitized the model, but lowered the specificity, resulting in unimproved discriminatory power (AUROC 0.868). However, addition of peak m/z 3884 enabled the correct assignment of the whole evaluation set and increased the selectivity by 13.2% (AUROC 1.000). This information surplus might be attributable to disease-associated alterations not covered by the conventional tumor markers.

Interestingly, the mass signal m/z 3884 was also found as potential discriminating peak in patients suffering from pancreatic cancer in a study by Koopmann et al. using SELDI-TOF MS analysis (35), but the underlying peptide was not identified. Using MALDI-TOF/TOF MS analysis, we could prove m/z 3884 as double charged ion of m/z 7767 by high resolution MS. A following database search revealed platelet factor 4 (PF4) as most likely candidate peptide. Using a specific G-protein coupled

antibody-based MALDI-TOF approach, we could identify the underlying peptide of candidate m/z 3884 as PF4. The subsequent direct immunological quantification of PF4 in all study samples corroborated our mass spectrometric findings and proved PF4 levels significantly decreased ($p=0.001$) in patients suffering from pancreatic cancer. To exclude an inflammatory response effect, we analyzed the PF4 concentration in serum samples of patients suffering from acute pancreatitis. The PF4 concentrations of pancreatitis patients were significantly higher ($p<0.001$) compared to the cancer patients and even slightly elevated compared to the healthy controls. This finding resembles results in previous studies of inflammatory bowel disease (36). Thus, it is unlikely that concomitant inflammation in pancreatic cancer is causing the decreased PF4 levels. In patients suffering from prostate cancer PF4 was also significantly decreased compared to controls (37). Therefore, the PF4 decrease itself seems not to be a pancreatic cancer specific effect, but it adds information to the conventional serum tumor marker panels consisting of CA 19-9 and CEA and thereby strongly improves the sensitivity and specificity of the laboratory tumor marker testing in patients suffering from pancreatic cancer. The application of a monoclonal PF4 ELISA specific for m/z 3884 (m/z 7767, respectively) might further enhance these results.

PF4 is a member of the C-X-C chemokine family (CXCL4) and is present in α granules of all mammalian platelets as well as in granules of mast cells (38, 39). The implication of PF4 in tumor growth and vascularisation is still in discussion and possible mechanisms of action are only partially elucidated (40). Recent evidence suggests that PF4 might pleiotropically both mark and mediate the expansion of pancreatic malignancies (41, 42). The cancer associated reduction of PF4 serum concentration in patients with pancreatic cancer might be explained by recent observations of Villanueva et al. (43, 44). They could show that differential exoprotease activities might contribute to cancer-type specific serum peptidome degradation. In pancreatic cancer, several matrix metalloproteinases (MMPs) are upregulated and partially secreted into the blood (45, 46). This has recently been shown for MMP-9 (47), which is also capable of degrading PF4 (45).

In conclusion, we identified and replicated PF4 as an additional discriminating marker in pancreatic cancer, which improves in combination with the conventional markers CA 19-9 and CEA the diagnostic power of tumor marker testing. This might be of additional relevance for the differential diagnosis of pancreatic cancer and pancreatitis. Further investigations are necessary to enlighten the complexity of PF4 action in pancreatic cancer and to evaluate the clinical impact of PF4 as novel and additional tumor marker in prospective clinical studies.

Disclosure of Potential Conflicts of Interest

The authors declare that no competing interests exist.

Acknowledgements

We gratefully thank Dr. Markus Kostrzewa and Dr. Thomas Elssner from Bruker Daltonics for skillful technical support regarding our mass-spectrometric platform. We also thank Dr. André Hagehülsmann from Microsoft Research for his outstanding support in planning and implementation of the proteome data base.

References

1. Lowenfels AB, Maisonneuve P. Epidemiology and risk factors for pancreatic cancer. Best Pract Res Clin Gastroenterol 2006; 20:197-209.
2. Michl P, Pauls S, Gress TM. Evidence-based diagnosis and staging of pancreatic cancer. Best Pract Res Clin Gastroenterol 2006; 20:227-51.
3. Chen R, Yi EC, Donohoe S, et al. Pancreatic cancer proteome: the proteins that underlie invasion, metastasis, and immunologic escape. Gastroenterology 2005; 129:1187-97.
4. Misek DE, Patwa TH, Lubman DM, Simeone DM. Early detection and biomarkers in pancreatic cancer. J Natl Compr Canc Netw 2007; 5:1034-41.
5. Hess V, Glimelius B, Grawe P, et al. CA 19-9 tumour-marker response to chemotherapy in patients with advanced pancreatic cancer enrolled in a randomised controlled trial. Lancet Oncol 2008; 9:132-8.
6. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003; 422:198-207.
7. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: written in blood. Nature 2003; 425:905.
8. Rai AJ, Chan DW. Cancer proteomics: Serum diagnostics for tumor marker discovery. Ann N Y Acad Sci 2004; 1022:286-94.
9. Coombes KR, Morris JS, Hu J, Edmonson SR, Baggerly KA. Serum proteomics profiling--a young technology begins to mature. Nat Biotechnol 2005; 23:291-2.
10. Diamandis EP, van der Merwe DE. Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. Clin Cancer Res 2005; 11:963-5.
11. Hortin GL, Jortani SA, Ritchie JC, Jr., Valdes R, Jr., Chan DW. Proteomics: a new diagnostic frontier. Clin Chem 2006; 52:1218-22.
12. Beretta L. Proteomics from the clinical perspective: many hopes and much debate. Nat Methods 2007; 4:785-6.
13. Grantzdorffer I, Carl-McGrath S, Ebert MP, Rocken C. Proteomics of pancreatic cancer. Pancreas 2008; 36:329-36.
14. Check E. Proteomics and cancer: running before we can walk? Nature 2004; 429:496-7.
15. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst 2004; 96:353-6.
16. Hortin GL. Can mass spectrometric protein profiling meet desired standards of clinical laboratory practice? Clin Chem 2005; 51:3-5.
17. Diamandis EP. Oncopeptidomics: a useful approach for cancer diagnosis? Clin Chem 2007; 53:1004-6.
18. Liotta LA, Petricoin EF. Putting the "Bio" back into Biomarkers: Orienting Proteomic Discovery toward Biology and away from the Measurement Platform. Clin Chem 2008; 54:3-5.
19. Banks RE, Stanley AJ, Cairns DA, et al. Influences of blood sample processing on low-molecular-weight proteome identified by surface-enhanced laser desorption/ionization mass spectrometry. Clin Chem 2005; 51:1637-49.
20. Diamandis EP. Peptidomics for cancer diagnosis: present and future. J Proteome Res 2006; 5:2079-82.
21. Banks RE. Preanalytical Influences in Clinical Proteomic Studies: Raising Awareness of Fundamental Issues in Sample Banking. Clin Chem 2008; 54:6-7.
22. Baumann S, Ceglarek U, Fiedler GM, Lembcke J, Leichtle A, Thiery J. Standardized approach to proteome profiling of human serum based on magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clin Chem 2005; 51:973-80.
23. Findeisen P, Sismanidis D, Riedl M, Costina V, Neumaier M. Preanalytical impact of sample handling on proteome profiling experiments with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clin Chem 2005; 51:2409-11.

24. Rai AJ, Vitzthum F. Effects of preanalytical variables on peptide and protein measurements in human serum and plasma: implications for clinical proteomics. Expert Rev Proteomics 2006; 3:409-26.
25. Fiedler GM, Baumann S, Leichtle A, et al. Standardized peptidome profiling of human urine by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clin Chem 2007; 53:421-8.
26. Bruegel M, Planert M, Baumann S, et al. Standardized peptidome profiling of human cerebrospinal fluid by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Proteomics 2008.
27. Conrad TOF, Leichtle A, Hagehülsmann A, et al. Beating the Noise: New Statistical Methods for Detecting Signals in MALDI-TOF Spectra below Noise Level. In: M. R. Berthold RG, and I. Fischer, editor. Proceeding of CompLife 2006: Springer-Verlag Berlin Heidelberg; 2006. p. 119-28.
28. Ketterlinus R, Hsieh SY, Teng SH, Lee H, Pusch W. Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. Biotechniques 2005; Suppl:37-40.
29. Bruker. clinprotools 2.0 User Manual: Bruker Daltonik GmbH; 2005.
30. McLerran D, Grizzle WE, Feng Z, et al. Analytical Validation of Serum Proteomic Profiling for Diagnosis of Prostate Cancer: Sources of Sample Bias. Clin Chem 2008; 54:44-52.
31. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics 2004; 3:367-78.
32. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001; 93:1054-61.
33. van der Merwe DE, Oikonomopoulou K, Marshall J, Diamandis EP. Mass spectrometry: uncovering the cancer proteome for diagnostics. Adv Cancer Res 2007; 96:23-50.
34. Chen R, Pan S, Brentnall TA, Aebersold R. Proteomic profiling of pancreatic cancer for biomarker discovery. Mol Cell Proteomics 2005; 4:523-33.
35. Koopmann J, Zhang Z, White N, et al. Serum diagnosis of pancreatic adenocarcinoma using surface-enhanced laser desorption and ionization mass spectrometry. Clin Cancer Res 2004; 10:860-8.
36. Meuwis MA, Fillet M, Geurts P, et al. Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling. Biochem Pharmacol 2007; 73:1422-33.
37. Lam YW, Mobley JA, Evans JE, Carmody JF, Ho SM. Mass profiling-directed isolation and identification of a stage-specific serologic protein biomarker of advanced prostate cancer. Proteomics 2005; 5:2927-38.
38. Maurer AM, Zhou B, Han ZC. Roles of platelet factor 4 in hematopoiesis and angiogenesis. Growth Factors 2006; 24:242-52.
39. McLaren KM, Pepper DS. The immunoelectronmicroscopic localization of human platelet factor 4 in tissue mast cells. Histochem J 1983; 15:795-800.
40. Bikfalvi A. Platelet factor 4: an inhibitor of angiogenesis. Semin Thromb Hemost 2004; 30:379-85.
41. Soucek L, Lawlor ER, Soto D, Shchors K, Swigart LB, Evan GI. Mast cells are required for angiogenesis and macroscopic expansion of Myc-induced pancreatic islet tumors. Nat Med 2007; 13:1211-8.
42. Theoharides TC. Mast cells and pancreatic cancer. N Engl J Med 2008; 358:1860-1.
43. Villanueva J, Shaffer DR, Philip J, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006; 116:271-84.
44. Villanueva J, Martorella AJ, Lawlor K, et al. Serum peptidome patterns that distinguish metastatic thyroid carcinoma from cancer-free controls are unbiased by gender and age. Mol Cell Proteomics 2006; 5:1840-52.

45. Van den Steen PE, Proost P, Wuyts A, Van Damme J, Opdenakker G. Neutrophil gelatinase B potentiates interleukin-8 tenfold by aminoterminal processing, whereas it degrades CTAP-III, PF-4, and GRO-alpha and leaves RANTES and MCP-2 intact. Blood 2000; 96:2673-81.
46. Patten LC, Berger DH. Role of proteases in pancreatic carcinoma. World J Surg 2005; 29: 258-63.
47. Tian M, Cui YZ, Song GH, et al. Proteomic analysis identifies MMP-9, DJ-1 and A1BG as overexpressed proteins in pancreatic juice from pancreatic ductal adenocarcinoma patients. BMC Cancer 2008; 8:241.

Figure Legends:

Figure 1: Schematic overview of the entangled peak selection procedure including reciprocal cross validation and external confirmation. From the discovery sets (A) and (B), consisting of patients (p) and controls (c) differentiating peak lists (boxes) are linearly and reciprocally derived. Comparison of these peak lists reveals the common candidate peak (D), which is applied as a discriminator on the external validation set. This discriminator was subsequently analyzed with respect to its discriminatory power and structure.

Figure 2: Peaks identified by CPT as significant were contrarily orientated (example peak m/z 1945, which is higher in the control subset Ac (A), but lower in the control subset Bc (B), compared to the corresponding controls).

Figure 3: Peak m/z 3884 in the controls (A) and the pancreatic carcinoma patients (B) and peak m/z 5959 in the controls (C) and pancreatic carcinoma patients (D) of the external validation set C, as revealed by proteomics.net.

Figure 4: ROC curves of the external validation set. The figure displays ROC curves of the external validation set. A shows sensitivity and specificity for the combination of the two tumor markers CA 19-9 and CEA, resulting in an AUROC of 0.868, B shows the first model plus marker m/z 5959, resulting also in an AUROC of 0.868, C shows the first model plus marker m/z 3884, resulting in an AUROC of 1.00.

Figure 5: Identification of m/z 3884 and m/z 7767, respectively. A: IMAC Cu eluate of a characteristic serum sample after standard preparation, B: IMAC Cu eluate of a PF4 standard after incubation with anti-PF4 MB-IAC Prot G particles, C: IMAC Cu eluate of a serum sample after incubation with anti-PF4 MB-IAC Prot G particles.

Supplemental Figure 1: Boxplots of immunologically determined Platelet Factor 4 (PF4) serum levels. PF4 serum levels are significantly decreased ($p \leq 0.001$) in patients suffering from pancreatic carcinoma compared to healthy controls as well as pancreatitis patients.

Figure 1

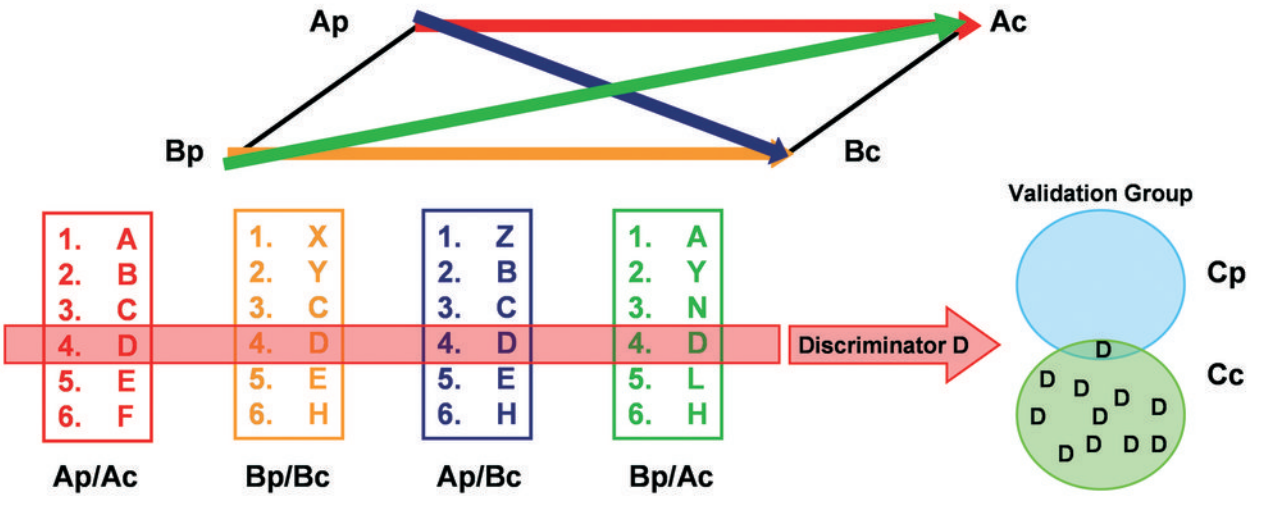
Ap
Ac
Bp
Bc
1. A
2. B
3. C
4. D
5. E
6. F
1. X
2. Y
3. C
4. D
5. E
6. H
1. Z
2. B
3. C
4. D
5. E
6. H
1. A
2. Y
3. N
4. D
5. L
6. H
Ap/Ac
Bp/Bc
Ap/Bc
Bp/Ac
Discriminator D
Validation Group
Cp
Cc
D
D
D
D
D
D
D
D
D
D
D

Figure 2

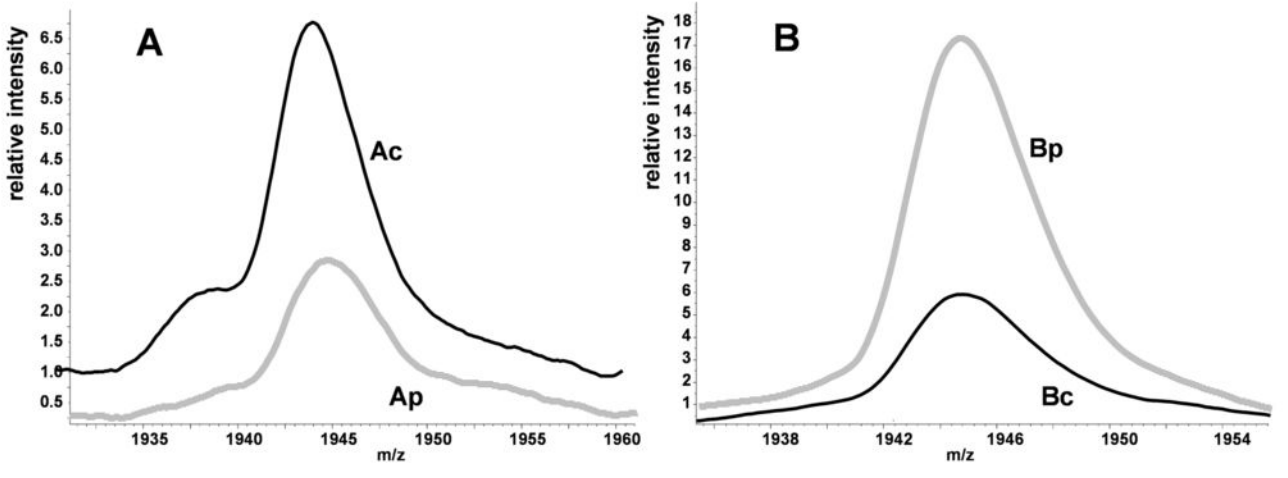
A
relative intensity
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
Ac
Ap
1935
1940
1945
1950
1955
1960
m/z
B
relative intensity
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Bp
Bc
1938
1942
1946
1950
1954
m/z

Figure 3

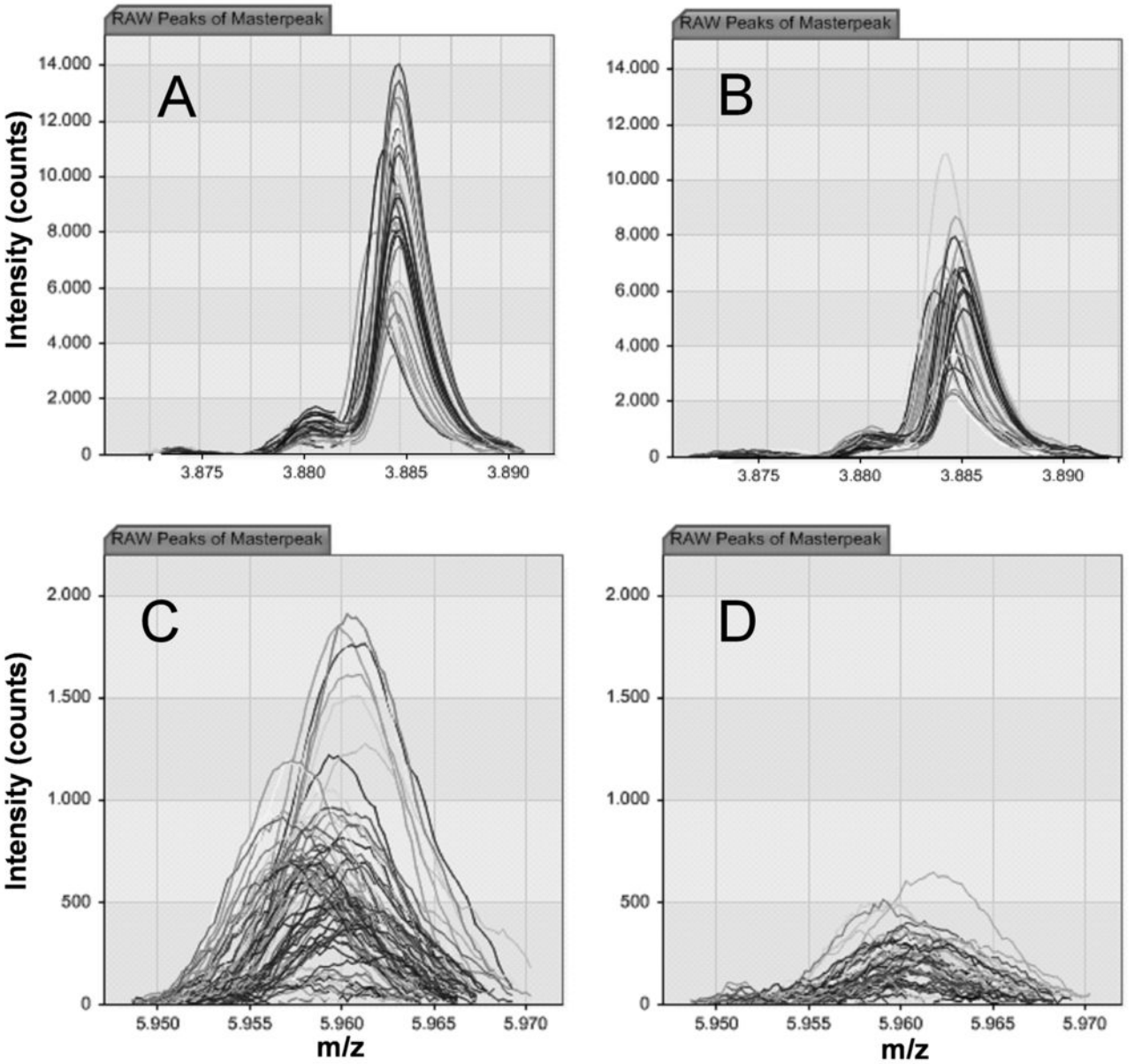
RAW Peaks of Masterpeak
A
B
C
D
Intensity (counts)
14.000
12.000
10.000
8.000
6.000
4.000
2.000
0
2.000
1.500
1.000
500
3.875
3.880
3.885
3.890
5.950
5.955
5.960
5.965
5.970
m/z

Figure 4

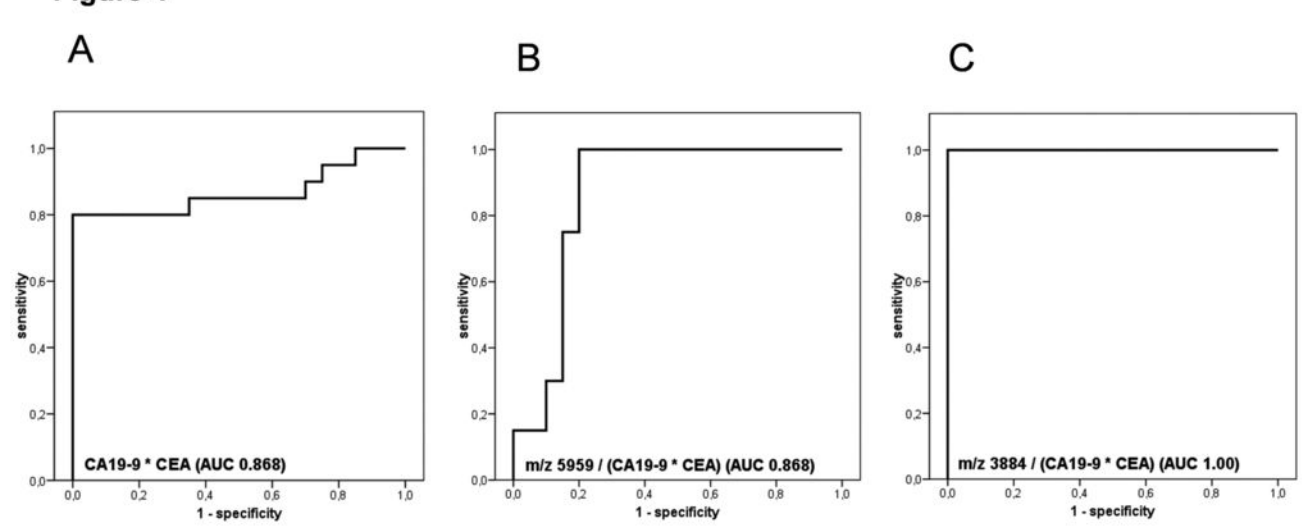
A
B
C
sensitivity
1 - specificity
0,0
0,2
0,4
0,6
0,8
1,0
CA19-9 * CEA (AUC 0.868)
m/z 5959 / (CA19-9 * CEA) (AUC 0.868)
m/z 3884 / (CA19-9 * CEA) (AUC 1.00)

Figure 5

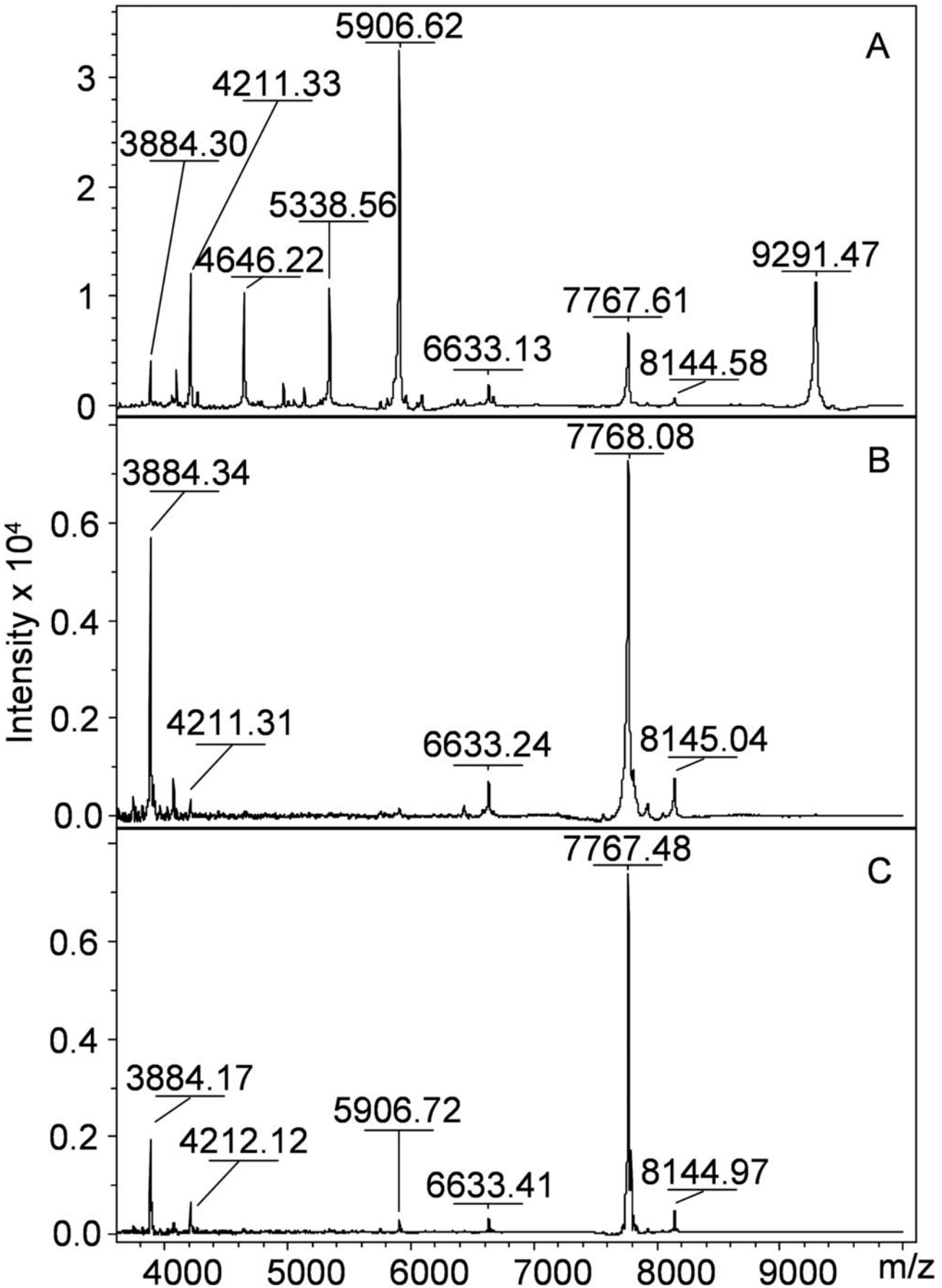
A
5906.62
4211.33
3884.30
5338.56
4646.22
9291.47
7767.61
6633.13
8144.58
B
7768.08
3884.34
4211.31
6633.24
8145.04
C
7767.48
3884.17
5906.72
4212.12
6633.41
8144.97
Intensity x 10⁴
0
1
2
3
0.0
0.2
0.4
0.6
4000
5000
6000
7000
8000
9000
m/z

TABLE 1.
Title: Clinical characteristics of patients with pancreatic cancer and controls.

Study Complex	Sample Sets	Parameter	Pancreatic Cancer	Controls
Discovery Study				
Set A	**UHL***	n	20	20
		male/female	10/10	10/10
		mean age (y)	59.3	50.1
		age range (y)	46-71	37-71
Set B	**UHH†**	n	20	20
		male/female	10/10	15/5
		mean age (y)	59.3	56.9
		age range (y)	47-70	41-85
External Validation				
Set C	**UHL***	n	20	20
		male/female	10/10	10/10
		mean age (y)	63.8	52.2
		age range (y)	33-72	32-70

*University Hospital Leipzig (UHL); †University Hospital Heidelberg (UHH).