

## Supplemental:

### Metabolic biomarker signature to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis

*Julia Mayerle<sup>1#</sup>, Holger Kalthoff<sup>2#</sup>, Regina Reszka<sup>3</sup>, Beate Kamlage<sup>3</sup>, Erik Peter<sup>3</sup>, Bodo Schniewind<sup>2</sup>, Sandra González Maldonado<sup>4</sup>, Christian Pilarsky<sup>5</sup>, Claus-Dieter Heidecke<sup>6</sup>, Philipp Schatz<sup>3</sup>, Marius Distler<sup>7</sup>, Jonas A. Scheiber<sup>1</sup>, Ujjwal M. Mahajan<sup>1</sup>, F. Ulrich Weiss<sup>1</sup>, Robert Grützmann<sup>5§</sup>, Markus M. Lerch<sup>1§</sup>*

<sup>1</sup>Department of Medicine A, University Medicine, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany

<sup>2</sup>Institut for Experimental Cancer Research (IET), Section for Molecular Oncology, UKSH, Campus Kiel, Kiel, Germany

<sup>3</sup>Metanomics Health GmbH, Berlin, Germany

<sup>4</sup>metanomics GmbH, Berlin, Germany

<sup>5</sup> Department of Surgery, University Hospital, Erlangen, Germany

<sup>6</sup>Department of General, Visceral, Thoracic and Vascular Surgery University Medicine Greifswald, Ernst-Moritz-Arndt University, Greifswald, Germany

<sup>7</sup> Clinic and Outpatient Clinic for Visceral-, Thorax- and Vascular Surgery, Medizinische Fakultät, TU Dresden, Dresden, Germany

<sup>8</sup>Medizinische Klinik und Poliklinik II, Klinikum der LMU München-Grosshadern, Germany

# equally contributing first authors, § equally contributing senior authors

#### Address for correspondence:

Markus M. Lerch, MD, FRCP

Department of Medicine A

University Medicine Greifswald

Ferdinand-Sauerbruchstrasse

17475 Greifswald

Phone: ++49-3834-86 7230

Fax: ++49-3834-86 7234

Email: lerch@uni-greifswald.de

## **Suppl. Material and Methods**

The study was designed and conducted according to the reporting recommendations for Tumor Marker Diagnostic Studies (REMARK) guidelines [[1]]. A total of 914 patients with pancreatic cancer, chronic pancreatitis, liver cirrhosis, healthy blood donors and preoperative patients with non-pancreatic or liver disease were prospectively enrolled in the study. All patients gave their written informed consent and the local ethics review boards approved the protocol at all three participating centers (Az. 110/99). Patients were consecutively recruited from University referral centers in Greifswald (80 pancreatic cancer patients, 120 chronic pancreatitis patients, 100 liver cirrhosis patients and 117 blood donors), Kiel (34 pancreatic cancer patients, 2 chronic pancreatitis patients, and 64 blood donors) and Dresden (157 pancreatic cancer patients, 160 chronic pancreatitis patients, and 80 controls (non-pancreatic disease). The non-pancreatic controls had the following patients' characteristics: all controls underwent general anaesthesia. The cohort comprised 52.2% male subjects, mean age was 64 years  $\pm$  14.4 SD, 22.5% were diabetics and mean BMI was calculated with 26.8  $\pm$  4.8 SD. 20 patients underwent vascular surgery, 18 patients received a hernia repair, 3 were resected for goiter and 39 received various other small surgical procedures under general anaesthesia. None was operated in metabolically deranged state. The final diagnosis of pancreatic cancer was established by a pathologist based on histology examination (either on resection specimen, on pancreatic fine needle aspiration or biopsy of liver metastasis) and PDAC ruled out in the disease groups by a medium follow up time of 24 months. EDTA-blood samples were taken before cancer treatment. For pancreatitis patients inclusion criteria for all cohorts included written informed consent given by patients and preexisting chronic pancreatitis. The diagnosis of chronic pancreatitis was made if one or more of the following criteria were met and no other diagnosis was more likely [2]: recurrent bouts of pancreatic pain with documented rise in amylase or lipase activity for a duration of more than one year and radiological evidence supporting the diagnosis, pancreatic calcifications, histological proof of chronic pancreatitis, unequivocal changes in pancreatic duct morphology, severely abnormal pancreatic function tests with maldigestion. Calcifications were identified on CT-scan, diabetes was diagnosed as suggested by the WHO definition and exocrine insufficiency was determined by either fecal elastase measurement or concurrent pancreatic enzyme supplementation. Suppl. table 7 gives a detailed summary on the patients' characteristics of our cohort.

Plasma was prepared by centrifugation at 1500 x g for 15 min at 4 to 8°C. The plasma supernatant was removed, mixed in a new centrifugation tube by inverting, split into 0.5-ml aliquots and stored at -80°C in 2-ml screw cap tubes. Samples were shipped on dry ice to the analytical laboratory. Serum sampling was performed according to the protocol above except BD

Vacutainer<sup>®</sup> Rapid Serum Tubes (RST) were used for sampling. Exclusion criteria were other concomitant malignant diseases, curative treatment for malignant disease within the last 2 years before recruitment, concomitant cystic diseases of the pancreas, pregnancy or patients unable to give informed consent

CA19-9 and total bilirubin determination were performed centralized at a certified clinical laboratory applying a cut-off of 37 U/ml as a classifier for CA19-9 and 1.2 mg/dL for bilirubin. According to the National Cancer Institute Early Detection Research Network (EDRN) [[3]] we started with an exploratory study (discovery, n=201) from two different centers. In a second phase (training set) we recruited 474 subjects with two different matrixes (Greifswald, serum: 80 PDAC, 79 CP, 80 liver cirrhosis, 77 blood donors; Dresden, plasma: 78 PDAC, 80 CP). The plasma samples of the training study (n=158) were used as a training set. A third independent cohort (Dresden, plasma: 79 PDAC, 80 CP, 80 non-pancreatic controls) was employed as testing set for validation of the biomarker signature. For details see flow chart (**figure 1**). Before analysis of the test set metabolites and algorithms were defined. Power analysis was performed to estimate an adequate sample size using representative metabolite profiling standard deviations that were determined in earlier studies [4]. The primary aim of the study was to determine a 20% metabolite concentration difference on a 5% significance level with approximately 72-99% power in patient samples. Metabolic difference was defined as absolute or relative difference in concentrations of individual metabolites.

### **Metabolite profiling:**

#### **MxP<sup>®</sup> Broad Profiling and MxP<sup>®</sup> Steroids:**

Three types of mass spectrometry analyses were applied. GC–MS (gas chromatography-mass spectrometry; Agilent 6890 GC coupled to an Agilent 5973 MS System, Agilent, Waldbronn, Germany) and LC–MS/MS (liquid chromatography-MS/MS; Agilent 1100 HPLC-System, Agilent, Waldbronn, Germany, coupled to an Applied Biosystems API4000 MS/MS-System, Applied Biosystems, Darmstadt, Germany) were used for broad profiling [[4]]. SPE -LC–MS/MS (Solid phase extraction-LC–MS/MS; Symbiosis Pharma, Spark, Emmen, Netherlands, coupled to an Applied Biosystems API4000 MS/MS-System, Applied Biosystems, Darmstadt, Germany) was used for the determination steroid levels. Fractionation and derivatization of samples and detection technologies have been previously described [5] [6] [7] [8] [9]] [7]. Proteins were removed from plasma samples (60 µl) by precipitation. Subsequently polar and non-polar fractions were separated for both GC–MS and LC–MS/MS analyses by adding water and a mixture of ethanol and dichloromethane. For GC–MS analyses, the non-polar fraction was

treated with methanol under acidic conditions to yield the fatty acid methyl esters derived from both free fatty acids and hydrolyzed complex lipids. The polar and non-polar fractions were further derivatized with O-methyl-hydroxylamine hydrochloride (20 mg/ml in pyridine) to convert oxo-groups to O-methyloximes and subsequently with a silylating agent (N-Methyl-N-(trimethylsilyl) trifluoroacetamide) before GC–MS analysis. For LC–MS/MS analyses, both fractions were dried and subsequently reconstituted in appropriate solvent mixtures. HPLC (High performance LC) was performed by gradient elution using methanol/water/formic acid on reversed phase separation columns. Steroids and their related metabolites were measured by online SPE-LC–MS/MS. Quantification was performed by using stable isotope-labeled standards.

### **MxP<sup>®</sup> Lipids:**

Total lipids were extracted from plasma by liquid/liquid extraction using chloroform/methanol. The lipid extracts were subsequently fractionated by normal phase liquid chromatography (NPLC) into eleven different lipid groups according to [[5] [8]]. The fractions were analyzed by LC-MS/MS using electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) with detection of specific multiple reaction monitoring (MRM) transitions for cholesterol esters (CE), sphingomyelins (SM), and ceramides (CER) respectively. Sphingosines and sphingosine-1-phosphates (SP) were analyzed by LC-MS/MS using electrospray ionization (ESI) with detection of specific multiple reaction monitoring (MRM) transitions as described by [[8]]. The lipid classes monoacylglycerides (MAG), triacylglycerides (TAG), phosphatidylcholines (PC), phosphatidylserines (PS), phosphatidylinositols (PI), lysophosphatidylcholines (LPC), diacylglycerides (DAG), free fatty acids (FFA) were measured by GC-FID (Flame Ionization Detector). The fractions were analyzed by GC-FID after derivatization with TMSH (Trimethyl sulfonium hydroxide), yielding the fatty acid methyl esters (FAME) corresponding to the acyl moieties of the class-separated lipids. The concentrations of FAME from C14 to C24 were determined in each fraction.

### **Data normalization, data set alignment, metabolite levels and nomenclature**

Metabolite profiling based on a semi-quantitative analytical platform results in relative metabolite levels (“ratio”) to a defined reference. To support this concept and to allow an alignment of different analytical batches, two different reference sample types were run in parallel throughout the whole process. First, a project pool was generated from aliquots of all samples and measured with four replicates within each analytical sequence that comprised 24 samples. For

all semi-quantitatively analyzed metabolites, the results of each analyte from each sample were normalized against the median of the corresponding analyte in the pool reference samples within each analytical sequence to provide pool-normalized ratios. This process step compensated for inter- and intra-instrumental variation, i.e. variability that occurs when different analytical sequences are analyzed by different devices. This corresponds to a single-point calibration using multiple aliquots of the same level. As the calibration samples are of the same matrix as the project samples, they account best for all sequence to sequence deviations that may occur within sample preparation and measurement. Furthermore, metabolite levels of the pooled sample are representative for metabolite levels of the study, leading to most accurate semi-quantitative values. Moreover, the slope of each metabolite signal, including unknown analytes, relative to sample amount has been determined during method development. Only metabolites with a slope justifying a single-point calibration approach were taken into account for the presented study.

Second, to allow for an experiment-to-experiment alignment of semi-quantitative data, MxPool™ (a large pool of a commercial human EDTA plasma suited for alignment of MxP® studies) was analyzed with 12 replicated samples in the identification and validation studies, but not in the exploratory study and the pool-normalized ratios were further normalized to the median of the MxPool™ samples, i.e. ratios from this study are on the same level and therefore comparable with data from other studies normalized to other aliquots of the same MxPool™.

The limit of detection and the dynamic range of the semi-quantitative measurements were determined by dilution and spiking experiments during method development. The term “additional” (add.) was applied to metabolite names when the quantification may be influenced by metabolites exhibiting identical analytical characteristics with respect to the quantitation method. A rigorous quality control was performed on peak, analyte and sample level and has been described previously [10].

We were only willing to proceed from the training to the test study if in the training study absolute performance of our biomarker signature alone or in combination with CA19-9 would either show diagnostic improvement of greater 10% in comparison to CA19-9 or Sensitivity be > 80% and Specificity:  $\geq 85\%$  [+/- range of 95% CI] or AROC:  $\geq 0.85$ .

Metabolite identification was carried out by authentic standards where applicable. In case the authentic standard is not commercially available, in-depth MS-experiments were carried out and interpreted resulting into a highly reliable structure annotation from our scientific knowledge. Such experiments include determination of exact mass by fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), labeling experiments, changes in derivatization or ionization-modes.

*Definition of high priority metabolites:* High priority metabolites are all metabolites semi quantitatively detected above the limit of quantification for the majority of analyzed samples. Semi-quantitative ratios are peak data normalized to median of pooled samples within each analytical sequence in order to correct for device variability (=> ratios relative to pool). The pool was generated from aliquots of each sample. Samples of the training and test data set were analyzed in two batches including MxPool™. For the linking of semi-quantitative data a second normalization step was introduced by further normalizing pool-normalized data to median of MxPool™ within each analytical batch. Total quantified data remained in absolute concentrations during batch alignment. Suppl. table 8 gives individual ratios of each of the metabolites of the biomarker signature. Suppl. figures 3A-3J give corresponding scatter plots.

## **Statistics**

### **Data set analysis and normalization**

Prior to statistical analysis, log<sub>10</sub> transformation of ratios was conducted so that the data distribution becomes approximately normal. SIMCA-P version 13.0 (Umetrics AB, Umea, Sweden), TIBCO® Spotfire® 3.3.1 and R 2.8.1 were used for data analyses and visualizations. Initially, an exploratory multivariate analysis [Principal Component Analysis (PCA)] was applied to log<sub>10</sub> transformed ratios scaled to unit variance.

A simple linear model (ANOVA, package nlme) with “disease”, “age”, “body mass index”, “sex” and “sample storage time” as fixed effects was fitted to the data. Significance level was set to 5%. The multiple test problem was addressed by calculating the false discovery rate (FDR) using the Benjamini & Hochberg method [[11]].

To classify patients depending on their metabolic profiles a penalized logistic regression was fitted via using the Elastic Net Algorithm using the R package glmnet [[12]]. Equal penalties were used for both the L1 and the L2 norm. The performance as measured by the sensitivity at a fixed specificity of 85% was determined on the training data by tenfold cross-validation. This corresponded to a cutoff of 0.384. Afterwards the cutoff was applied without retraining on the test data and the performance measured in terms of sensitivity and specificity.

Confidence levels for the AUC and the specificity at a fixed sensitivity were calculated using the binormal model for the ROC Curve. When the sensitivity is fixed at a particular value PPV, NPV and accuracy become monotone functions of the specificity and confidence intervals for these estimates are obtained by transformation of the confidence interval for the specificity.

Confidence intervals for sensitivity, specificity and accuracy were obtained for the pre-specified cutoff in the training data by the method of Clopper and Pearson for the binomial distribution. For PPV and NPV the confidence intervals were obtained by the method of Gart and Nam [[13]] for

ratios of binomial parameters as implemented in the R package pairwise CI [[14]]. When comparing our marker and CA19-9 on the test data differences in sensitivity and specificity were tested for with the McNemar test.

## References:

- 1 McShane LM, Hayes DF. Publication of tumor marker research results: the necessity for complete and transparent reporting. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2012;30:4223-32.
- 2 Mayerle J, Hoffmeister A, Werner J, Witt H, Lerch MM, Mossner J. Chronic pancreatitis-- definition, etiology, investigation and treatment. *Deutsches Arzteblatt international* 2013;110:387-93.
- 3 Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, et al. Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute* 2001;93:1054-61.
- 4 Jung K, Reszka R, Kamlage B, Bethan B, Stephan C, Lein M, et al. Tissue metabolite profiling identifies differentiating and prognostic biomarkers for prostate carcinoma. *International journal of cancer Journal international du cancer* 2013;133:2914-24.
- 5 Christie WW. Rapid separation and quantification of lipid classes by high performance liquid chromatography and mass (light-scattering) detection. *Journal of lipid research* 1985;26:507-12.
- 6 Mutch DM, Fuhrmann JC, Rein D, Wiemer JC, Bouillot JL, Poitou C, et al. Metabolite profiling identifies candidate markers reflecting the clinical adaptations associated with Roux-en-Y gastric bypass surgery. *PloS one* 2009;4:e7905.
- 7 Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L. Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J* 2000;23:131-42.
- 8 Schmidt H, Schmidt R, Geisslinger G. LC-MS/MS-analysis of sphingosine-1-phosphate and related compounds in plasma samples. *Prostaglandins Other Lipid Mediat* 2006;81:162-70.
- 9 van Ravenzwaay B, Cunha GC, Leibold E, Looser R, Mellert W, Prokoudine A, et al. The use of metabolomics for the discovery of new biomarkers of effect. *Toxicol Lett* 2007;172:21-8.
- 10 Meller S, Meyer HA, Bethan B, Dietrich D, Maldonado SG, Lein M, et al. Integration of tissue metabolomics, transcriptomics and immunohistochemistry reveals ERG- and gleason score-specific metabolomic alterations in prostate cancer. *Oncotarget* 2016;7:1421-38.
- 11 Klipper-Aurbach Y, Wasserman M, Braunspiegel-Weintrob N, Borstein D, Peleg S, Assa S, et al. Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. *Medical hypotheses* 1995;45:486-90.
- 12 Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software* 2010;33:1-22.
- 13 Gart JJ, Nam J. Approximate interval estimation of the ratio of binomial parameters: a review and corrections for skewness. *Biometrics* 1988;44:323-38.
- 14 Scherer R, Schaarschmidt F, Prescher S, Priesnitz KU. Simultaneous confidence intervals for comparing biodiversity indices estimated from overdispersed count data. *Biometrical journal Biometrische Zeitschrift* 2013;55:246-63.