

## Panel of serum metabolites discriminates cancer patients and healthy participants of lung cancer screening – a pilot study

Małgorzata Roś-Mazurczyk<sup>1</sup>, Anna Wojakowska<sup>1</sup>, Łukasz Marczak<sup>2</sup>, Krzysztof Polański<sup>3</sup>, Monika Pietrowska<sup>1</sup>, Joanna Polanska<sup>4</sup>, Rafał Dziadziuszko<sup>5</sup>, Jacek Jassem<sup>5</sup>, Witold Rzyman<sup>5</sup> and Piotr Widlak<sup>1</sup>✉

<sup>1</sup>Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice, Poland; <sup>2</sup>Polish Academy of Science, Institute of Bioorganic Chemistry, Poznań, Poland; <sup>3</sup>University of Warwick, Coventry, Great Britain; <sup>4</sup>Silesian University of Technology, Gliwice, Poland; <sup>5</sup>Medical University of Gdańsk, Gdańsk, Poland

**Introduction.** Blood biomarkers may support early diagnosis of lung cancer by enabling pre-selection of candidates for computed tomography screening or discrimination between benign and malignant screening-detected nodules. We aimed to identify features of serum metabolome distinguishing individuals with early-detected lung cancer from healthy participants of the lung cancer screening program. **Methods.** Blood samples were collected in the course of a low-dose computed tomography screening program performed in the Gdańsk district (Northern Poland). The analysis included 31 patients with screening-detected lung cancer and the pair-matched group of 92 healthy controls. The gas chromatography coupled to mass spectrometry (GC/MS) approach was used to identify and quantify small metabolites present in serum. **Results.** There were several metabolites detected in the sera whose abundances discriminated patients with lung cancer from controls. Majority of the differentiating components were down-regulated in cancer samples, including amino acids, carboxylic acids and tocopherols, whereas benzaldehyde was the only compound significantly upregulated. A classifier including nine serum metabolites allowed separation of cancer and control samples with 100% sensitivity and 95% specificity. **Conclusions.** Signature of serum metabolites discriminating between cancer patients and healthy participants of the early lung cancer screening program was identified using a GC/MS metabolomics approach. This signature, though not validated in an independent dataset, deserves further investigation in a larger cohort study.

**Key words:** early detection; lung cancer screening; metabolomics; serum biomarkers

**Received:** 31 January, 2017; **revised:** 27 June, 2017; **accepted:** 17 July, 2017; **available on-line:** 12 August, 2017

✉ e-mail: piotr.widlak@io.gliwice.pl

**Abbreviations:** GC-MS, gas chromatography coupled with mass spectrometry; LD-CT, low-dose computed tomography; MS, mass spectrometry

### INTRODUCTION

Lung cancer is the leading cause of cancer mortality, responsible for about one-fifth of cancer-related deaths worldwide. The majority of lung cancer cases are diagnosed at advanced stages and have a grim prognosis (the average 5-year survival of about 10–15%). However, in the case of disease detected at early stages, prognosis is

much better (the average 5-year survival in the range of 65–85%). Thus, in addition to primary prevention (i.e., tobacco smoking control), screening for the early detection of lung cancer might be the major strategy to reduce lung cancer mortality (Hoffman *et al.*, 2000; Jemal *et al.*, 2010; Torre *et al.*, 2012). Several diagnostic tools allowing early lung cancer detection have been investigated within the past decades, but none have found their routine application in clinical practice. Nevertheless, the low-dose computed tomography (LD-CT) screening in a high-risk group had shown a 20% reduction in lung cancer-specific mortality as compared with conventional chest X-ray examination (Aberle *et al.*, 2011). Hence, lung cancer screening based on LD-CT is now the most efficacious strategy for lung screening, with a perspective for world-wide cancer mortality reduction. However, relatively low positive predictive value and sensitivity of this test may lead to “over-diagnosis”. In our own experience, around 75% of patients with screening-detected lung abnormalities underwent unnecessary diagnostic work-up, including around 25% of patients subjected to further invasive procedures (Rzyman *et al.*, 2013). For these reasons, complementation of CT-based screening with other tests allowing effective and reliable pre-selection of individuals for LD-CT examination, or better discrimination between benign and malignant nodules detected by LD-CT, seems a critical issue for practical application of this strategy (Priola *et al.*, 2013; Rzyman *et al.*, 2015). Blood is the most available source of biomarkers potentially enhancing the power of early lung cancer detection or differentiating lung nodules. Several components of blood, including circulating tumor cells, circulating tumor DNA, micro RNA, autoantibodies and specific serum/plasma proteins have been analyzed in the search for such biomarkers (Hassanein *et al.*, 2012; Hassanein *et al.*, 2011; Sozzi *et al.*, 2014), but none has yet been adopted in the clinics.

The overall response of human organism to pathological conditions is mirrored in different molecular fractions of body fluids, including the metabolome. In recent years, monitoring of cancer-related metabolites in blood has been an emerging approach to detection and diagnosis of different malignancies (Spratlin *et al.*, 2009). Several studies have demonstrated that profiling serum or plasma samples by mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy could reveal metabolites whose blood levels discriminate patients with lung cancer from healthy individuals or from patients with non-malignant lung diseases. Such differentiating

compounds included phospholipids, carboxylic acids, amino acids, sugars and many other small metabolites (Jordan *et al.*, 2010; Rocha *et al.*, 2011; Hori *et al.*, 2011; Guo *et al.*, 2012; Wang *et al.*, 2013; Deja *et al.*, 2014; Liu *et al.*, 2014; Chen *et al.*, 2015). More recently, two relatively large studies using NMR-based analysis of plasma or serum metabolome revealed a promising diagnostic potential of multicomponent lung cancer signatures built of different types of small metabolites (Puchades-Carrasco *et al.*, 2016; Louis *et al.*, 2016). Another study, using MS-based approaches, revealed a large set of metabolites whose serum levels discriminated lung cancer patients from matched controls, and allowed for building multicomponent cancer classifiers (Mazzone *et al.*, 2016). However, lung cancer patients enrolled in the above-mentioned studies included both, early and advanced cancer cases, and no study has yet been performed using material obtained exclusively from high-risk subjects participating in the LD-CT screening. Hence, potential relevance of proposed biomarkers for early detection of lung cancer remains to be verified. Here, we assessed the applicability of a GC-MS-based approach to identify a signature of serum metabolites discriminating between patients with screening-detected lung cancer and healthy participants of the LD-CT screening program.

## MATERIALS AND METHODS

**Study subjects.** Material for this study was collected in the course of the Pomeranian Lung Cancer Screening Program performed by Gdansk Medical University between 2008 and 2010. This program enrolled over 8000 participants and offered LD-CT examination for current or former smokers with at least a 20 pack-year history, aged from 50 to 75 years. Blood samples were collected from about 3600 participants. The study group involved material from 31 participants who were finally diagnosed with lung cancer (i.e., 0.9% of the screened group) (Table 1). Each cancer case was accompanied by three controls, with no detected malignancy, matched according to sex, age and smoking history, who were selected from the participants of the LD-CT screening program (92 cases). The study was approved by the Ethics Committee of Gdansk Medical University (approval number NKEBN/42/2009), and each participant provided a written informed consent indicating her/his voluntary

**Table 1. Characteristics of the donor groups.**

Group	Controls (n=92)	Cancer cases (n=31)
Clinical stage	N.A.	
– IA		20
– IB		2
– IIA		0
– IIB		3
– IIIA		6
Histopathology	N.A.	
– adenocarcinoma		21
– squamous cell carcinoma		9
– not otherwise specified NSCLC		1
Sex		
– female	40	14
– male	52	17
Age (years) [median]	52–73 [59]	52–72 [60]
Smoking (pack-year) [median]	10–43 [24]	15–60 [20]

N.A., not applicable

participation in the project and provision of blood samples for future research.

**Sample preparation.** Peripheral blood was collected into a 5 mL BD Vacutainer Tube, incubated for 30 min. at room temperature to allow clotting, and then centrifuged at 1000×g for 10 min. to remove the clot. The serum was aliquoted and stored at –70°C prior to analysis. 25 µl of serum was added to 200 µl mixture of MeOH:H<sub>2</sub>O (1:1 v/v), vortexed for 20 min., centrifuged for 10 min. at 18000 g, and then the supernatant 1 was collected into a new tube. The pellet was re-suspended in 200 µl mixture of CH<sub>2</sub>Cl<sub>2</sub>:MeOH (3:1 v/v). The mixture was placed in the ultrasonic bath for 5 min, vortexed for 10 min and centrifuged for 10 min at 18000×g, and then the supernatant 2 was collected into a new tube. Both supernatant fractions were combined and evaporated in a vacuum concentrator.

**GC/MS analysis.** Dried extracts were derivatized directly before GC/MS. Each sample was mixed with 25 µl of methoxyamine hydrochloride in pyridine (20 mg/ml) and vortexed (950 rpm) for 90 min. at 37°C, and then 80 µl of N-methyl-N-trimethylsilyl-trifluoro-acetamide was added to the mixture and vortexed (950 rpm) for 30 min. at 37°C. The GC/MS analysis was performed with Agilent 7890A gas chromatograph (Agilent Technologies) combined with Pegasus 4D GCxGC-TOFMS mass spectrometer (Leco). Compounds were separated using the DB-5 bonded-phase fused-silica capillary column (30 m length, 0.25 mm inner diameter, 0.25 µm film thickness) (J&W Scientific Co.); the GC oven temperature program was as follows: 2 min. at 70°C, raised by 8°C/min. to 300°C and held for 16 min. at 300°C (the total time of GC analysis was 46.75 min). Helium was used as the carrier gas at a flow rate of 1 ml/min. One microliter of each sample was injected in a splitless mode. The initial injector temperature was 20°C for 0.1 min., then temperature raised to 350°C at the 600°C/min. rate. The septum purge flow rate was 3 ml/min. and the purge was turned on after 60 s. The transfer line and ion source temperatures were set to 250°C. In-source fragmentation was performed with 70 eV energy.

**Analysis of spectra.** Mass spectra were recorded in the mass range of 35–650 *m/z*. All spectra were subjected to automatic peak detection, deconvolution, retention index calculation and library search by Leco ChromaTOF-GC software (v4.51.6.0). The alkane series mixture (C-10 to C-36) was used to correct retention time (Rt) and to determine the retention index (RI) for each compound. Automated identification of metabolites was based on the Replib, Mainlib and Fiehn libraries; the quality threshold was set for similarity index (SI) above 700 and retention index ±10. The unique quantification masses were specified for each component and the samples were reprocessed in order to obtain accurate peak areas for the deconvoluted components. The obtained profiles were normalized against the sum of chromatographic peak area (using the TIC approach). All peaks that were identified as artifacts (column bleed, alkanes, plasticizer, derivatization reagents) and peaks corresponding to unidentified compounds were excluded from further analysis.

**Statistical analysis.** Quantitative spectral data were log transformed and missing values were imputed using the k-nearest-neighbor approach with the standardized Euclidean metric on a per-group basis (only metabolites present in more than 2/3 of samples in each set were used for further analyses). For each compound, the normality of the distribution of abundance and the homogeneity of variances were assessed using the Lilliefors

test and Bartlett's test, respectively. Then, the significance of differences between groups was estimated using either the two-sample T test (with a correction for heteroscedasticity, if necessary) or the nonparametric U Mann-Whitney test; the Benjamini-Hochberg approach was applied to the  $p$ -values for multiple testing correction ( $q$ -value). A multivariable logistic regression model was constructed to find the signature describing the relation between metabolite abundances and patient status. A stepwise procedure combined with the Bayesian Information Criterion (BIC), R<sup>2</sup> and  $p$ -value of likelihood ratio test was used for model selection. The contribution of individual predictors was measured using the Wald test. The optimal threshold for the discriminating function was found by maximizing the value of Youden's index based on the receiver operating characteristics (ROC) and the random guess line. The standard classification performance indices (sensitivity, specificity, positive predictive value, negative predictive value) were

calculated as appropriate ratios of false/right negatives/positives.

**Bioinformatics analysis.** Metabolomic pathways were identified using the Metabolite Set Enrichment Analysis (accessed on 10.2016 at <http://www.msea.ca/MSEA/faces/Home.jsp>); a statistical significance of resulting over-representation was estimated using the hypergeometric test.

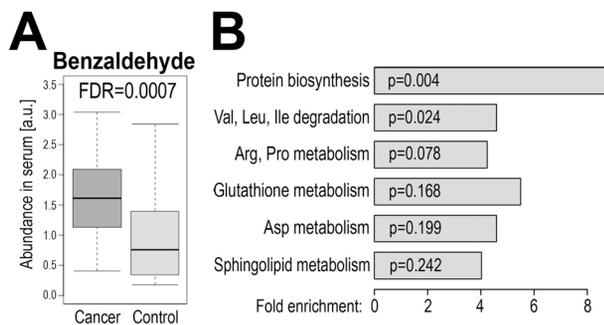
## RESULTS

The GC/MS approach, a standard analytical tool in metabolomics study (Spratlin *et al.*, 2009), was used to characterize profile of metabolites in serum samples collected from participants of the LD-CT screening program for early detection of lung cancer. The study included 31 patients with screening-detected lung cancer and 92 matched controls. In general, there were 195 unique metabolites identified in the analyzed samples.

**Table 2. Compounds with differential abundance between early lung cancer and control samples**

Metabolite	Type of compound	Lung cancer		Healthy controls		Difference between groups		
		mean [a.u.]	C.V.	mean [a.u.]	C.V.	cancer/control ratio	$p$ -value	$q$ -value
Compounds upregulated in lung cancer samples								
<u>Benzaldehyde</u>	cyclic compound	1.60	0.41	0.96	0.80	1.67	<0.00001	0.0007
<u>Hydroxypyruvic acid</u>	carboxylic acid	0.53	1.07	0.35	0.98	1.51	0.122	0.427
<u>Urea</u>	amide	574	0.98	486	0.86	1.18	0.242	0.576
Compounds downregulated in lung cancer samples								
<u>Glycolic acid</u>	carboxylic acid	0.44	0.88	0.79	0.72	0.56	0.0007	0.020
$\beta$ -Hydroxybutyric acid	carboxylic acid	2.20	0.95	3.88	0.98	0.57	0.003	0.068
<u>L-Isoleucine (Ile)</u>	amino acid	14.61	0.72	23.20	0.75	0.63	0.0002	0.006
<u>Gluconic acid lactone</u>	carboxylic acid ester	0.34	0.82	0.52	0.95	0.64	0.038	0.242
$\gamma$ -Tocopherol	polycyclic compound	0.28	0.41	0.41	0.69	0.69	0.009	0.124
Indol-3-acetic acid	carboxylic acid	0.44	0.59	0.62	0.76	0.71	0.018	0.179
Ketoleucine	carboxylic acid	0.43	0.62	0.59	0.73	0.72	0.004	0.071
Hexadecanol	fatty alcohol	0.35	0.53	0.47	1.17	0.74	0.038	0.264
$\alpha$ -Tocopherol	polycyclic compound	6.00	0.48	7.82	0.50	0.77	0.032	0.241
L-Valine (Val)	amino acid	42.65	0.37	55.05	0.53	0.77	0.017	0.201
L-Proline (Pro)	amino acid	56.39	0.42	72.43	0.49	0.78	0.028	0.241
Inositol	sugar	0.17	0.61	0.22	0.71	0.79	0.048	0.266
<u>Allyl laurate</u>	carboxylic acid ester	0.17	0.23	0.22	0.48	0.79	0.012	0.149
L-Glycine (Gly)	amino acid	53.53	0.77	67.12	0.59	0.80	0.009	0.124
Phosphoryethanolamine	amine	0.51	0.49	0.63	0.49	0.81	0.023	0.212
Citrulline	amino acid	0.65	0.78	0.79	0.61	0.82	0.035	0.242
Pregn-4-ene-3,11,20-trione	polycyclic compound	0.74	0.88	1.35	2.72	0.55	0.180	0.504
Pipecolinic acid	amino acid	28.26	0.99	50.15	1.78	0.56	0.118	0.454
Parabanic acid	nitrogenous acid	1.03	0.75	1.63	1.36	0.63	0.105	0.431
$\alpha$ -Hydroxybutyric acid	carboxylic acid	5.44	0.35	8.51	0.96	0.64	0.397	0.677
Dihydrouracil	heterocyclic compound	3.78	0.76	5.88	1.23	0.64	0.191	0.523
Cholesta-4,6-dien-3-ol	polycyclic compound	1.17	0.76	1.80	3.16	0.65	0.387	0.677
<u>Linolenic acid</u>	fatty acid	12.63	0.37	14.51	0.47	0.87	0.188	0.506
<u>L-Phenylalanine (Phe)</u>	amino acid	6.34	0.42	7.17	0.63	0.88	0.702	0.819

Compounds present in the cancer classifier are underlined; compounds with differences below and above the threshold of statistical significance ( $p \geq 0.05$ ) are separated by horizontal lines; C.V., coefficient of variation; a.u., arbitrary units.



**Figure 1. Metabolites discriminating early lung cancer patients from healthy controls.**

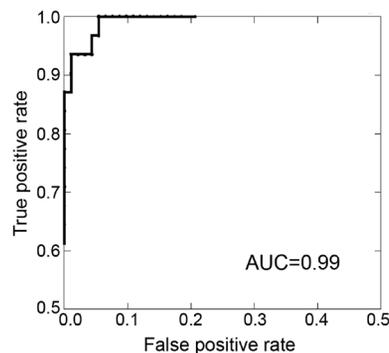
**(Panel A)** Serum level of benzaldehyde; boxplots show minimum, lower quartile, median, upper quartile and maximum values (abundance in arbitrary units). **(Panel B)** Metabolite sets enrichment overview of discriminating compounds, showing a relative over-representation (fold enrichment) of pathways associated with cancer-downregulated metabolites and its statistical significance (only pathways showing fold enrichment  $\geq 4$  are presented).

Out of them, 102 compounds detected and quantified in the majority of the samples in each group were used for further analyses and testing of a cancer classifier (Supplementary Table S1 at [www.actabp.pl](http://www.actabp.pl)). In general, the inter-individual variability in levels of the detected metabolites was similar in both groups: the mean coefficient of variation was 0.65 and 0.83 in cancer and control samples, respectively.

First, we looked for individual metabolites discriminating groups of cancer patients and healthy controls. We found 17 metabolites that showed significant difference between both groups ( $p < 0.05$ ), including 16 compounds with abundances lower in cancer compared to control samples (fold change 0.56 to 0.82; Table 2). The only compound significantly upregulated in cancer samples (fold change 1.67) was benzaldehyde (Fig. 1A). However, only three compounds, namely benzaldehyde, isoleucine and glycolic acid, retained statistical significance after correction for multiple testing ( $q < 0.05$ ). Additionally, there were six metabolites showing at least 50% down-regulation in cancer samples and one metabolite showing at least 50% upregulation (fold change  $< 1.50$  or  $> 0.67$ ), yet observed differences were below the threshold of statistical significance ( $p > 0.05$ ; Table 2).

Metabolites whose abundances in serum discriminated patients with early lung cancer and healthy donors included mostly carboxylic acids and amino acids (Table 2). To reveal systemic information about potential functional importance of these differences, the identified differentiating compounds ( $p < 0.05$ ) were annotated to metabolic pathways using the Metabolite Set Enrichment Analysis (Fig. 1B). This type of analysis allowed identification of “over-represented” pathways associated with metabolites discriminating between cancer and control samples (i.e., pathways associated with the types of compounds that were more numerous than expected by chance). It is noteworthy that primary pathways associated with compounds downregulated in cancer samples included those involved in protein metabolism; there were two pathways that had shown a statistical significance of overrepresentation: protein biosynthesis, and Val, Leu and Ile degradation.

Finally, we built and tested a multicomponent signature that allowed separation of cancer and control samples. The optimum classifier included nine metabolites: benzaldehyde, hydroxypyruvic acid and urea (cancer-up-



**Figure 2. ROC characteristic of the separator including nine serum metabolites discriminating the cancer and control samples.**

regulated compounds), and glycolic acid, isoleucine, gluconic acid lactone, allyl laurate, phenylalanine and linolenic acid (cancer-downregulated compounds; Table 2). This classifier allowed separation of cancer and healthy samples with 100% sensitivity, 95% specificity, 86% positive predictive value and 100% negative predictive value. The ROC characteristic of such a separator is shown in Fig. 2. The obtained classifier was not validated with an independent dataset, yet high indices observed in the discovery set apparently inspire for its testing in further studies.

## DISCUSSION

General features of cancer metabolism that could be mirrored in blood metabolome include enhanced glycolysis and gluconeogenesis in combination with suppressed Krebs cycle and lipid catabolism. These features of cancer metabolism were also demonstrated in blood of lung cancer patients (Rocha *et al.*, 2011; Hori *et al.*, 2011; Louis *et al.*, 2016). This characteristic of blood metabolome was accompanied by a decreased level of different amino acids, including Ala, Cys, Glu, His, Met, Pro, Thr, Trp, Tyr and Val (Rocha *et al.*, 2011; Puchades-Carrasco *et al.*, 2016; Mazzone *et al.*, 2016). Metabolites specifically associated with glycolysis and gluconeogenesis were not differentiating in the current study, likely due to inclusion of only low-stage, early-detected cancer patients. On the other hand, decreased levels of amino acids (Gly, Ile, Pro, Val) and carboxylic/fatty acid (and their derivatives) were compatible with the known characteristic of the cancer metabolome. Among compounds with decreased abundance in the samples of early-detected lung cancer were vitamin E species ( $\alpha$ -tocopherol and  $\gamma$ -tocopherol). The use of antioxidant vitamins in lung cancer chemoprevention has long been speculated, but there is no evidence that increased intake of vitamin E is associated with reduced risk of lung cancer in smokers (Willett *et al.*, 1984; Wright *et al.*, 2000; Mahabir *et al.*, 2008; Virtamo *et al.*, 2014; Wu *et al.*, 2015). Similarly, plasma level of  $\alpha$ -tocopherol does not seem to be associated with lung cancer risk (Comstock *et al.*, 2008). On the other hand, however,  $\alpha$ -tocopherol was downregulated while  $\gamma$ -tocopherol and  $\delta$ -tocopherol were upregulated in serum of lung cancer patients when compared to healthy controls (Mazzone *et al.*, 2016). Hence, results of the latter study may suggest some association between metabolism of vitamin E and the development of lung cancer. The general characteristic of early lung cancer signature developed in the present study seems to be coherent with the already known features of lung cancer metabo-

lome, yet only a small fraction of specific compounds discriminating cancer and control samples matched metabolites reported in previous studies.

We found that benzaldehyde was the only compound showing significantly increased abundance in samples collected from screening-detected lung cancer cases. Benzaldehyde was previously shown to be present among volatile compounds detected in exhaled breath of lung cancer patients (Bajtarevic *et al.*, 2009). Importantly, this compound was specific for cancer patients (i.e. was not detected in breath of healthy volunteers), and its presence was not related to the smoking behaviour. Moreover, signature composed of different volatile compounds, including benzaldehyde, allowed discrimination between lung cancer patients and healthy controls with 50–80% sensitivity and 100% specificity. Hence, an increased level of benzaldehyde observed in serum of lung cancer patients apparently confirms the relevance of this compound in lung cancer detection.

A unique feature of our metabolomics study is the use of material derived solely from screening-detected cancer cases and corresponding controls collected in the course of the LD-CT lung cancer screening program in a population of high-risk smokers. Owing to a relatively low number of study samples, our results should be confronted with results of a previous large-cohort studies including general population of lung cancer patients. A comprehensive study using the MS-based approaches, which involved 94 cancer patients and 190 matched controls, was recently published by Mazzone *et al.* (Mazzone *et al.*, 2016). This study revealed tocopherols among essential components of signatures discriminating lung cancer patients from healthy controls, either in a general lung cancer population (downregulated  $\alpha$ -tocopherol) or in the group of patients with squamous cell cancer (upregulated  $\gamma$ -tocopherol and  $\delta$ -tocopherol). Moreover, all amino acids that showed reduced abundance in sera of cancer patients analyzed in the current study (Gly, Ile, Phe, Pro, Val), were also cancer-downregulated in the former study. The results of two other large metabolomics studies based on NMR spectroscopy involving 296 cancer cases *vs.* 114 controls (Puchades-Carrasco *et al.*, 2016) and 357 cancer cases *vs.* 347 controls (Louis *et al.*, 2016) were not confirmed in our study. This inconsistency could be attributed to both, different features of analytical approaches and different donor characteristics. Nevertheless, we are aware of some limitations of the study presented here, which include a relatively low number of donors and its “snapshot” design (no follow-up of controls is available). Our results should therefore be considered as exploratory.

## CONCLUSIONS

In conclusion, we developed a signature based on a set of serum metabolites that discriminate cancer cases from the matched healthy subjects in a unique series of early lung cancer screening participants. Remarkably, several components of this signature were associated with the known features of cancer metabolism revealed in previous studies that included a general population of lung cancer patients. Hence, a further validation study is warranted to confirm the robustness of our data and to assess a potential clinical utility of the signature. Such study has been recently initiated in conjunction with a large LD-CT screening program carried by our group.

## Acknowledgements

This work was supported by the National Science Centre, Grant 2013/11/N/NZ7/00770, and the National Centre for Research and Development, Grant MOLTEST2013 (POIG.01.01.02-20-080/09) and Grant MOLTEST-BIS (DZP/PBS3/247184/2014). The computation analyses were carried out using the GeCONiI infrastructure (POIG.02.03.01-24-099/13).

## REFERENCES

- Aberle DR, Adams AM, Berg CD, *et al.* (2011) National lung screening trial research team reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* **365**: 395–409. <http://doi:10.1056/NEJMoa1102873>
- Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, Ligor T, Filipiak W, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Haidenberger A, Buszewski B, Miekisch W, Schubert J, Amann A (2009) Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC Cancer* **9**: 348. <http://doi:10.1186/1471-2407-9-348>
- Chen Y, Ma Z, Min L, Li H, Wang B, Zhong J, Dail L (2015) Biomarker identification and pathway analysis by serum metabolomics of lung cancer. *Biomed Res Int* **2015**: 183624. <http://doi:10.1155/2015/183624>
- Comstock GW, Alberg AJ, Huang HY, Wu K, Burke AE, Hoffman SC, Norkus EP, Gross M, Cutler RG, Morris JS, Spate VL, Helzlsouer KJ (2008) The risk of developing lung cancer associated with antioxidants in the blood: ascorbic acids, carotenoids, alpha-tocopherol, selenium, and total peroxyl radical absorbing capacity. *Am J Epidemiol* **168**: 831–840. <http://doi:10.1093/aje/kwn328>
- Deja S, Porebska I, Kowal A, Zabek A, Barg W, Pawelczyk K, Stanimirova I, Daszykowski M, Korzeniewska A, Jankowska R, Mlynarz P (2014) Metabolomics provide new insights on lung cancer staging and discrimination from chronic obstructive pulmonary disease. *J Pharm Biomed Anal* **100**: 369–380. <http://doi:10.1016/j.jpba.2014.08.020>
- Guo Y, Wang X, Qiu L, Qin X, Liu H, Wang Y, Li F, Wang X, Chen G, Song G, Li F, Guo S, Li Z (2012) 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. <http://doi:10.1016/j.cca.2012.08.010>
- Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP (2012) The state of molecular biomarkers for the early detection of lung cancer. *Cancer Prev Res (Phila)* **5**: 992–1006. <http://doi:10.1158/1940-6207.CAPR-11-0441>
- Hassanein M, Rahman JS, Chaurand P, Massion PP (2011) Advances in proteomic strategies toward the early detection of lung cancer. *Proc Am Thorac Soc* **8**: 183–188. <http://doi:10.1513/pats.201012-069MS>
- Hoffman PC, Mauer AM, Vokes EE (2000) Lung cancer. *Lancet* **355**: 479–485. [http://doi:10.1016/S0140-6736\(00\)82038-3](http://doi:10.1016/S0140-6736(00)82038-3)
- Hori S, Nishiumi S, Kobayashi K, Shinohara M, Hatakeyama Y, Kotani Y, Hatano N, Maniwa Y, Nishio W, Bamba T, Fukusaki E, Azuma T, Takenawa T, Nishimura Y, Yoshida M (2011) A metabolomic approach to lung cancer. *Lung Cancer* **74**: 284–292. <http://doi:10.1016/j.lungcan.2011.02.008>
- Jemal A, Center MM, DeSantis C, Ward EM (2010) Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol Biomarkers Prev* **19**: 1893–1907. <http://doi:10.1158/1055-9965.EPI-10-0437>
- Jordan KW, Adkins CB, Su L, Halpern EF, Mark EJ, Christiani DC, Cheng LL (2010) Comparison of squamous cell carcinoma and adenocarcinoma of the lung by metabolomic analysis of tissue-serum pairs. *Lung Cancer* **68**: 44–50. <http://doi:10.1016/j.lungcan.2009.05.012>
- Liu J, Mazzone PJ, Cata JP, Kurz A, Bauer M, Mascha EJ, Sessler DI (2014) Serum free fatty acid biomarkers of lung cancer. *Chest* **146**: 670–679. <http://doi:10.1378/chest.13-2568>
- Louis E, Adriaenssens P, Guedens W, Bigirumurame T, Baeten K, Vanhove K, Vandeurzen K, Darquennes K, Vansteenkiste J, Dooms C, Shkedy Z, Mesotten L, Thomeer M (2016) Detection of lung cancer through metabolic changes measured in blood plasma. *J Thor Oncol* **11**: 516–523. <http://doi:10.1016/j.jtho.2016.01.011>
- Mahabir S, Schendel K, Dong YQ, Barrera SL, Spitz MR, Forman MR (2008) Dietary alpha-, beta-, gamma- and delta-tocopherols in lung cancer risk. *Int J Cancer* **123**: 1173–1180. <http://doi:10.1002/ijc.23649>
- Mazzone PJ, Wang XF, Beukemann M, Zhang Q, Seeley M, Mohney R, Holt T, Pappan KL (2016) Metabolite profiles of the serum of patients with non-small cell carcinoma. *J Thor Oncol* **11**: 72–88. <http://doi:10.1016/j.jtho.2015.09.002>

- Priola AM, Priola SM, Giaj-Levra M, Basso E, Veltri A, Fava C, Cardinale L. (2013) Clinical implications and added costs of incidental findings in an early detection study of lung cancer by using low-dose spiral computed tomography. *Clin Lung Cancer* **14**: 139–148. [http://doi: 10.1016/j.clcc.2012.05.005](http://doi:10.1016/j.clcc.2012.05.005)
- Puchades-Carrasco L, Jantus-Lewintre E, Pérez-Rambla C, García-García F, Lucas R, Calabuig S, Blasco A, Dopazo J, Camps C, Pineda-Lucena A (2016) Serum metabolomic profiling facilitates the non-invasive identification of metabolic biomarkers associated with the onset and progression of non-small cell lung cancer. *Oncotarget* **7**: 12904–12916. [http://doi: 10.18632/oncotarget.7354](http://doi:10.18632/oncotarget.7354)
- Rocha CM, Carrola J, Barros AS, Gil AM, Goodfellow BJ, Carreira IM, Bernardo J, Gomes A, Sousa V, Carvalho L, Duarte IF (2011) Metabolic signatures of lung cancer in biofluids: NMR-based metabolomics of blood plasma. *J Proteome Res* **10**: 4314–4324. [http://doi: 10.1021/pr200550p](http://doi:10.1021/pr200550p)
- Rzyman W, Dziedzic R, Jelitto-Górska M, Biadacz I, Książek J, Siebert J, Zdrojewski T, Studniarek M (2015) Results of an open-access lung cancer screening program with low-dose computed tomography: the Gdańsk experience. *Pol Arch Med Wewn* **125**: 232–239.
- Rzyman W, Jelitto-Górska M, Dziedzic R, Biadacz I, Książek J, Chwirot P, Marjanski T (2013) Diagnostic work-up and surgery in participants of the Gdansk lung cancer screening programme: the incidence of surgery for non-malignant conditions. *Interact Cardiovasc Thorac Surg* **17**: 969–973. [http://doi: 10.1093/icvts/ivt388](http://doi:10.1093/icvts/ivt388)
- Sozzi G, Boeri M (2014) Potential biomarkers for lung cancer screening. *Transl Lung Cancer Res* **3**: 139–1348. [http://doi: 10.3978/j.issn.2218-6751.2014.06.04](http://doi:10.3978/j.issn.2218-6751.2014.06.04)
- Spratlin JL, Serkova NJ, Eckhardt SG (2009) Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* **15**: 431–440. [http://doi: 10.1158/1078-0432.CCR-08-1059](http://doi:10.1158/1078-0432.CCR-08-1059)
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics 2012. *CA Cancer J Clin* **65**: 87–108. [http://doi: 10.3322/caac.21262](http://doi:10.3322/caac.21262)
- Virtamo J, Taylor PR, Kontto J, Männistö S, Utriainen M, Weinstein SJ, Huttunen J, Albanes D (2014) Effects of  $\alpha$ -tocopherol and  $\beta$ -carotene supplementation on cancer incidence and mortality: 18-year postintervention follow-up of the Alpha-tocopherol, Beta-carotene Cancer Prevention Study. *Int J Cancer* **135**: 178–185. [http://doi: 10.1002/ijc.28641](http://doi:10.1002/ijc.28641)
- Wang L, Tang Y, Liu S, Mao S, Ling Y, Liu D, He X, Wang X (2013) Metabonomic profiling of serum and urine by (1)H NMR-based spectroscopy discriminates patients with chronic obstructive pulmonary disease and healthy individuals. *PLoS One* **8**: e65675. [http://doi: 10.1371/journal.pone.0065675](http://doi:10.1371/journal.pone.0065675)
- Willett WC, Polk BF, Underwood BA, Stampfer MJ, Pressel S, Rosner B, Taylor JO, Schneider K, Hames CG. (1984) Relation of serum vitamins A and E and carotenoids to the risk of cancer. *N Engl J Med* **310**: 430–434
- Wright GS, Gruidl ME (2000) Early detection and prevention of lung cancer. *Curr Opin Oncol* **12**: 143–148
- Wu QJ, Xiang YB, Yang G, Li HL, Lan Q, Gao YT, Zheng W, Shu XO, Fowke JH (2015) Vitamin E intake and the lung cancer risk among female nonsmokers: a report from the Shanghai Women's Health Study. *Int J Cancer* **136**: 610–617. [http://doi: 10.1002/ijc.29016](http://doi:10.1002/ijc.29016)