Diacetylspermine is a novel pre-diagnostic serum biomarker for non-small cell lung cancer and has additive performance with pro-SFTPB

Wikoff, et al

SUPPLEMENTARY METHODS

Study Population

Participants in this nested case-control study were selected from the Carotene and Retinol Efficacy Trial (CARET) described previously (see Goodman et al.[1] for details). Briefly, CARET was a multicenter, randomized, double-blinded, placebo-controlled trial conducted to assess the safety and efficacy of daily supplementation with 30 mg of β-carotene plus 25,000 IU of retinyl palmitate in reducing lung cancer incidence in persons at high risk for the disease. A total of 18,314 participants were enrolled from two high risk populations: heavy smokers (N = 14,254) and asbestos-exposed workers (N = 4,060). Eligible for the heavy smoker population were men and women, 50-69 years of age, who were either current or former smokers (quit no more than 6 years prior) with at least 20 pack-years of cigarette smoking. The asbestos-exposed population was comprised of men aged 45-69 who were smoking at baseline or quit within 15 years prior and had either a history of occupational asbestos exposure documented by at least 5 years of employment in a protocol defined high-risk trade or a history of occupational exposure in any job, and baseline chest x-ray positive for changes consistent with non-malignant asbestos-related disease. The CARET intervention was halted in 1996, with participant follow-up for cancer and mortality outcomes continued until 2005 with a total of 1,445 lung cancers observed when follow up ended. Reports of cancer were confirmed though review of clinical records and pathology reports following a detailed protocol. Blood draws were conducted at baseline and every other year thereafter for most participants. A common blood collection and processing protocol was followed at each study center. Participants were not required to fast prior to blood collection, but elapsed time from the most recent meal was recorded. Blood was collected in foil-covered vacutainers and allowed to stand at room temperature for 30 to 120 minutes prior to centrifugation. Several serum aliquots ranging from 0.5 to 2 mL were created

and stored at -20°C for up to two weeks and then transferred to central 70°C freezers for long term storage. All CARET participants provided informed consent at recruitment and throughout follow-up, and the institutional review boards at each of the six study centers, located in Seattle (WA), Portland (OR), San Francisco (CA), Irvine (CA), Baltimore (MD), and New Haven (CT), approved all study procedures.

Sample extraction and preparation

The mass spectrometry analysis, including sample preparation and interpretation of chromatographic and mass spectral peaks, was performed at University of California Davis.

Serum samples were extracted in three volumes of ice cold methanol, mixed, and maintained at -20°C for 30min, followed by centrifugation for 10 min at 14K RCF. Supernatants were transferred to clean tubes, and re-centrifuged. Supernatants were again transferred to clean tubes and dried in a spinVac, resuspended in 80:20 acetonitrile:water, sonicated to dissolve, centrifuged for 5 min at 14K RCF, and transferred to an autosampler tube.

Chromatography and Mass Spectrometry

A BEH HILIC column 1.7 μ, dimension 2 x 150 mm (Waters) was used. The temperature was maintained at 50°C. Solvent A was 10mM ammonium acetate pH 5.0 in water and Solvent B was 10mM ammonium acetate pH 5.0 in 90:10 acetonitrile:water. The flow rate was 400μl/min beginning with 100%B to 70%B over 14 minutes, and then to 45%B over ½ minute and the flow rate was increased to 450μl/min and held at 45%B for 2 minutes, then returned to 100%B where it was held for 6 minutes for re-equilibration. The total injection-to-injection time was 23 minutes. The autosampler was maintained at 4°C. The HPLC system was an Agilent 1200 and the mass spectrometer was an Agilent quadrapole time-of-flight (QTOF) 6530 with a JetStream ion source; two reference ions were infused separately for internal, continuous mass correction during data collection.

Before each run, the instrument source was cleaned, and the instrument was tuned and calibrated. Data were collected at a rate of 2 spectra/sec. The nebulizer flow was 35L/min with the drying gas at 9L/min, and temperature of 350°C; JetStream gas temperature was 200°C. Data were collected separately in positive and negative ion modes, and were accurate to 3ppm or better.

Data analysis and compound identification

Data in instrument specific format were converted to mzXML format. The program XCMS was used for non-linear alignment of the data in the time domain and automatic integration and extraction of the peak intensities[2]; isotope assignment and adduct annotation was performed with CAMERA.[3] For the positive ion mode HILIC LC/MS data there were 2,161 observations after background removal. Accurate mass information (within 2-3ppm) was obtained from profiling and used to search metabolite databases. Three metabolomics databases were used for compound identification: METLIN, LIPIDMAPS, and HMDB; when needed, PUBCHEM and SciFinder Scholar were used for chemical searches. These hits provided a list of *possible* identities of the metabolites. For the DAS, the biomarker identified in this study, it was necessary to synthesize the compound as it was not available commercially (see below).

We then used the synthetic standard for comparison to the chromatographic and mass spectral properties of the unknown, and found a match on the following criteria: (1) Accurate mass (2) MS/MS pattern (CID) (3) isotope pattern (4) column retention time. The unknown compound was identified by repeating the LC conditions used in the profiling experiment, and collecting collision-induced-dissociation MS/MS fragmentation data on the ion using the QTOF mass spectrometer. This process was repeated for the synthetic chemical standard, and both the compound retention time and MS/MS fragmentation pattern were compared to the unknown. For all four criteria, there was an excellent match between the synthetic DAS standard and the unknown.

Synthesis of diacetylspermine

N^1,N^{12}-Diacetylspermine (CAS: 61345-83-3) was synthesized according to the previously published method.[4] Briefly, 100 mg of spermine was added 3ml of 1,4-dioxane in a 10ml beaker, with stirring and gentle heating for about 15 minutes until dissolved. Methoxydiacetamide was solubilized by gentle warming, and 150 µl was added drop-wise to the spermine/dioxane with stirring. A crystalline precipitate was allowed to form overnight and collected by centrifugation, washed with dioxane and dissolved in ethanol; crystallization was effected by addition of an equimolar volume of HCl/ethanol, followed by ethylacetate. The crystals were washed with dioxane and the crystallization was repeated. The crystals were air dryed followed by drying under vacuum. Identity of the synthetic DAS standard was confirmed by NMR. 1H NMR data was obtained on a 600 MHz Bruker Avance III spectrometer equipped with a 5 mm CPTCI cryoprobe, using a 30 degree hard pulse, 4 dummy scans and 16 scan averages. Powdered sample was dissolved in 99.96 % D2O (Sigma Aldrich) to a final concentration of 2.5 mg/mL.

Supplementary Table 1.

Serum DAS levels in the discovery set.

		Control subjects				Case subjects		
	N	Median	IQR	P value	N	Median	IQR	P value
Total	199	0.0508	0.0359-0.0753	-	100	0.0771	0.0423-0.1358	-
Sex								
Female	50	0.0480	0.0310-0.0717	0.1596	25	0.0695	0.0326-0.1020	0.2170
Male	149	0.0535	0.0394-0.0758		75	0.0773	0.0436-0.1563	
Age								
< 55	38	0.0453	0.0275-0.0542	0.0018	19	0.0785	0.0406-0.1017	0.1793
55-65	104	0.0494	0.0377-0.0709		52	0.0667	0.0406-0.1082	
≥ 65	57	0.0651	0.0409-0.0875		29	0.1128	0.0586-0.1595	
Fasting time (hours)								
< 3	73	0.0471	0.0353-0.0703	0.1137	28	0.0732	0.0389-0.1904	0.9474
3-6	76	0.0563	0.0400-0.0837		44	0.0722	0.0451-0.1421	
≥ 6	50	0.0497	0.0323-0.0709		27	0.0850	0.0418-0.1184	
BMI (kg/m^2)								
< 25	53	0.0502	0.0326-0.0748	0.0824	29	0.0764	0.0442-0.1206	0.8331
25-30	90	0.0499	0.0349-0.0689		32	0.0746	0.0423-0.1306	
≥ 30	55	0.0592	0.0400-0.0863		39	0.0785	0.0418-0.1570	
Smoking Status								
Current	121	0.0504	0.0338-0.0731	0.4753	61	0.0773	0.0452-0.1336	0.7651
Former	78	0.0532	0.0394-0.0777		39	0.0768	0.0381-0.1406	
Pack-years								
<40	70	0.0523	0.0393-0.0721	0.0191	19	0.0679	0.0381-0.1358	0.7217
40-50	50	0.0430	0.0303-0.0607		29	0.0850	0.0498-0.1498	
≥ 50	79	0.0574	0.0416-0.0844		52	0.0724	0.0439-0.1216	
Exposure population								
Asbestos-exposed worker	56	0.0515	0.0359-0.0789	0.8622	31	0.1358	0.0606-0.1972	0.0080

Heavy Smoker	143	0.0508	0.0359-0.0739		69	0.0679	0.0416-0.1036	
Histology								
Adenocarcinoma	-	-	-	-	40	0.0662	0.0387-0.1065	0.3592
SCC	-	-	-	-	30	0.0818	0.0564-0.1417	
Other NSCLC	-	-	-	-	30	0.0974	0.0361-0.1621	
Stage								
I and II	-	-	-	-	14	0.0776	0.0498-0.1044	0.3056
III and IV	-	-	-	-	69	0.0842	0.0455-0.1566	
Unknown	-	-	-	-	17	0.0608	0.0328-0.1200	
Months from blood collection to diagnosis (months)								
0-6	-	-	-	-	48	0.1005	0.0473-0.1695	0.0110
>6-12	-	-	-	-	52	0.0634	0.0409-0.1041	

IQR: interquartile range. P values were calculated by Mann Whitney test or Kruskal-Wallis test.

Supplementary Table 2.

Performance of pro-SFTPB in the discovery set.

Variable	Group	P value	AUC	95% CI
Total		0.0001	0.635	0.571-0.700
Sex	Female	0.0023	0.714	0.595-0.832
	Male	0.0056	0.613	0.537-0.689
Smoking Status	Current Smoker	< 0.0001	0.682	0.603-0.761
	Former Smoker	0.1954	0.574	0.468-0.680
Histology	AdenoCa	0.0015	0.677	0.574-0.779
	SCC	0.1601	0.591	0.474-0.709
	Other NSCLC	0.1080	0.605	0.481-0.728
Stage	I and II	0.7138	0.537	0.357-0.717
	III and IV	0.0001	0.663	0.588-0.738
Timing of blood draw	0-6 months	0.0142	0.626	0.532-0.720
	>6-12 months	0.0031	0.645	0.556-0.734

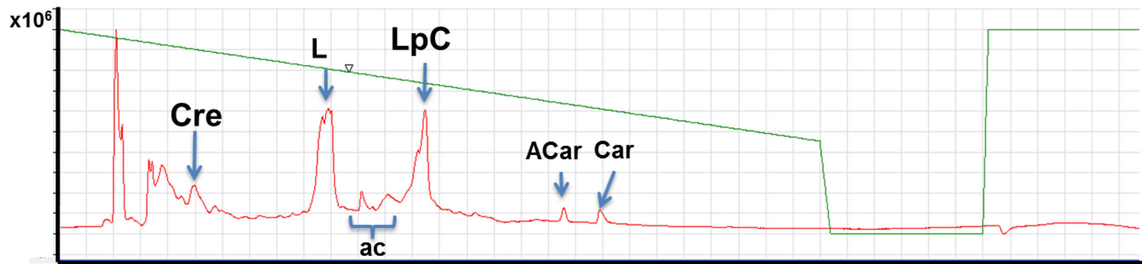
Supplementary Table 3.

Performance of pro-SFTPB in the validation set.

Variable	Group	P value	AUC	95% CI
Total		< 0.0001	0.699	0.639-0.760
Sex	Female	0.0014	0.696	0.578-0.813
	Male	< 0.0001	0.705	0.635-0.775
Smoking Status	Current Smoker	< 0.0001	0.710	0.636-0.784
	Former Smoker	0.0010	0.693	0.590-0.797
Histology	AdenoCa	0.0172	0.633	0.530-0.737
	SCC	< 0.0001	0.756	0.659-0.853
	Other NSCLC	0.0006	0.718	0.604-0.832
Stage	I and II	0.0051	0.693	0.570-0.817
	III and IV	< 0.0001	0.686	0.607-0.766
Timing of blood draw	0-6 months	< 0.0001	0.741	0.650-0.831
	>6-12 months	< 0.0001	0.675	0.595-0.755

Supplementary Figure 1.

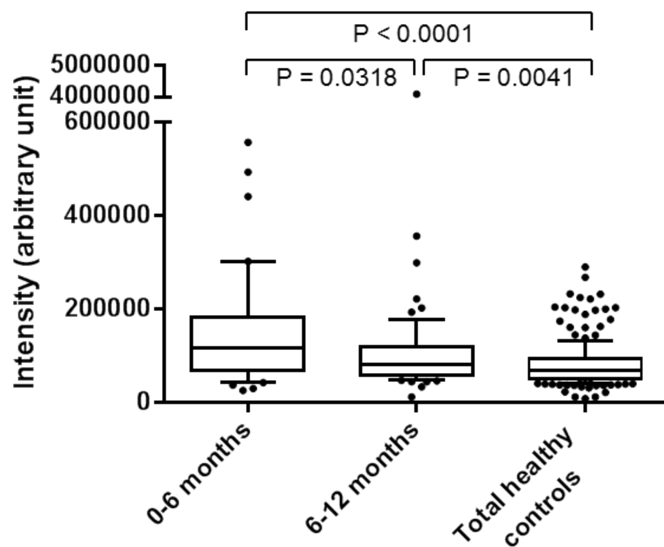
Total ion chromatogram (TIC) highlighting selected profiled compounds.



Cre: creatinine; L: large lipids; ac: acetate clusters; LpC: lysophosphatidylcholines; ACar: acetylcarnitine; Car: carnitine. Green line: gradient profile.

Supplementary Figure 2.

DAS levels and timing of blood draw in the Validation set.



P values were calculated by Mann Whitney test.

REFERENCES

1. Goodman GE, Thornquist MD, Balmes J, et al: The Beta-Carotene and Retinol Efficacy Trial: incidence of lung cancer and cardiovascular disease mortality during 6-year follow-up after stopping beta-carotene and retinol supplements. J Natl Cancer Inst 96:1743-50, 2004

2. Benton HP, Wong DM, Trauger SA, et al: XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. Anal Chem 80:6382-9, 2008

3. Kuhl C, Tautenhahn R, Bottcher C, et al: CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. Anal Chem 84:283-9, 2012

4. Samejima K, Hiramatsu K, Takahashi K, et al: Identification and determination of urinary acetylpolyamines in cancer patients by electrospray ionization and time-of-flight mass spectrometry. Anal Biochem 401:22-9, 2010