

**Patterns of carbon-bound exogenous compounds in lung cancer patients
and association with disease pathophysiology**

Thomas Kunzke^{1#}, Verena M. Prade^{1#}, Achim Buck¹, Na Sun¹, Annette Feuchtinger¹, Marco Matzka¹,
Isis E. Fernandez^{2,3}, Wim Wuyts⁴, Maximilian Ackermann^{5,6}, Danny Jonigk⁷, Michaela Aichler¹,
Ralph A. Schmid⁸, Oliver Eickelberg⁹, Sabina Berezowska^{10,11*}, Axel Walch^{1*}

1. Research Unit Analytical Pathology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany
2. Department of Internal Medicine V, Comprehensive Pneumology Center (CPC-M), Member of the German Center for Lung Research (DZL), Ludwig-Maximilians-Universität München, Munich, Germany
3. Comprehensive Pneumology Center, Institute of Lung Biology and Disease, Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL), Munich, Germany
4. Unit for Interstitial Lung Diseases, Department of Respiratory Medicine, University Hospitals Leuven, Leuven, Belgium
5. Institute of Functional and Clinical Anatomy, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany
6. Institute of Pathology and Molecular Pathology, Helios University Clinic Wuppertal, University of Witten/Herdecke, Wuppertal, Germany
7. Institute of Pathology, Hannover Medical School, Hannover, Germany and Member of the German Center for Lung Research (DZL), Biomedical Research in Endstage and Obstructive Lung Disease Hannover (BREATH)
8. Department of General Thoracic, Inselspital University Hospital Bern, Bern, Switzerland
9. Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, University of Pittsburgh, Pittsburgh, USA
10. Institute of Pathology, University of Bern, Bern, Switzerland
11. Department of Laboratory Medicine and Pathology, Institute of Pathology, Lausanne University Hospital and Lausanne University, Lausanne, Switzerland

These authors contributed equally to this work

* Corresponding authors:

Axel Walch, MD; Ingolstädter Landstraße 1, 85764 München, Germany; Phone:+49 89 3187-2739;
Fax: +49 89 3187-3349; E-mail: axel.walch@helmholtz-muenchen.de

Sabina Berezowska, MD; Bugnon 25, CH-1011 Lausanne; Phone:+41 21 314-7211; Fax: +41 21 314-7115; E-mail: sabina.berezowska@chuv.ch

The authors declare no potential conflicts of interest.

Running title:

Exogenous compounds in lung cancer patients

ABSTRACT

Asymptomatic anthracosis is the accumulation of black carbon particles in adult human lungs. It is a common occurrence, but the pathophysiological significance of anthracosis is debatable. Using in situ high mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry imaging analysis, we discovered noxious carbon-bound exogenous compounds, such as polycyclic aromatic hydrocarbons (PAHs), tobacco-specific nitrosamines, or aromatic amines, in a series of 330 lung cancer patients in highly variable and unique patterns. The characteristic nature of carbon-bound exogenous compound had a strong association with patient outcome, tumor progression, the tumor immune microenvironment, PD-L1 expression, and DNA damage. Spatial correlation network analyses revealed substantial differences in the metabolome of tumor cells compared to tumor stroma depending on carbon-bound exogenous compounds. Overall, the bioactive pool of exogenous compounds is associated with several changes in lung cancer pathophysiology and correlates with patient outcome. Given the high prevalence of anthracosis in the lungs of adult humans, future work should investigate the role of carbon-bound exogenous compounds in lung carcinogenesis and lung cancer therapy.

Keywords: lung cancer, anthracosis, exogenous compounds, patient outcome, mass spectrometry imaging

Significance

This study identifies a bioactive pool of carbon-bound exogenous compounds in patient tissues associated with several tumor biological features, contributing to an improved understanding of drivers of lung cancer pathophysiology.

INTRODUCTION

Asymptomatic anthracosis is the macroscopically and histologically visible black discoloration resulting from the deposition of black carbon particles in various anatomical locations of human lungs. Associated with age, environmental pollution, and smoking load, anthracosis can serve as an index of lifetime exposure to exogenous factors (1). Studies have shown lungs of heavy smokers to exhibit more pronounced anthracosis (2) and an association with lung carcinogenesis or cancer progression (3,4). Others have shown cigarette smoke not to be a risk factor for anthracosis and found no epidemiological or etiological link with lung cancer (5).

The highly complex and heterogeneous chemical composition of black carbon particles comprises numerous organic and inorganic compounds, including carbon, silica, aluminum and iron oxide (5,6). Black carbon particles can bind potentially toxic or carcinogenic compounds present in air pollutants, soot, dust, or tobacco smoke (2,7,8). Furthermore, black carbon particles are carriers of toxic chemicals, such as polycyclic aromatic hydrocarbons (PAHs) and nicotine-derived nitrosamine ketones, to the lung, immune cells, and systemic blood circulation (9). The toxins that were originally thought to be chemically inert can be retained, released or metabolized over a long period of time (10). Although extensively researched, the effects and interplay of carbon particles in anthracosis and exogenous compounds within their natural cellular and extracellular context of human lung tissue are unexplored and challenging due to the complex histological interrelationships.

Mass spectrometry imaging (MSI) has gained significant relevance in biomedical research and reveals the discrete distribution of compounds and their related metabolites. MSI has high molecular specificity and allows comprehensive, multiplexed detection and localization of thousands of endogenous metabolites directly in tissues (11). In a very recent study, mass spectrometry imaging was applied on mice to characterize the *in situ* organ distribution of intratracheal instilled and intravenously injected carbon particles, revealing surface-adsorbed aromatic hydrocarbons (12). The toxicological and pathological findings based on studies of the molecular and cellular processes

induced by toxins are important to achieve a mechanistic understanding. One of the strengths of MSI is its ability to directly overlay molecular information with tissue sections to correlatively compare molecular and histologic information. Therefore, MSI can provide novel insights into the effects and interactions of anthracosis, compounds and endogenous metabolites within their natural cellular and extracellular context in human lung tissue.

Using *in situ* high mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) MSI analysis, we report carbon-bound exogenous compounds in a series of 330 lung cancer patients. The spatial distribution of compounds like PAHs, tobacco-specific nitrosamines, or aromatic amines, as well as their impact on patient outcome, tumor progression, composition of intratumoral immune cells, programmed death-ligand 1 (PD-L1) expression, and DNA damage is examined. Furthermore, we investigate metabolic differences between tumor cells and the tumor microenvironment and illuminate the relationship of concentration and composition of black carbon pigments in lung cancer patients.

MATERIALS AND METHODS

Squamous-cell carcinoma (SQCC) patients and tissues

We retrospectively analyzed 330 consecutive patients with primary resected SQCC, diagnosed at the Institute of Pathology, University of Bern without previous or concomitant diagnosis of SQCC of other organs, to reliably exclude metastatic lung disease, as previously described (13). All patients provided written informed consent, and the study was done in accordance with the Declaration of Helsinki. The cohort was assembled according to pathology files, and validated according to clinical files. The histology of all cases was re-assessed according to current WHO-guidelines for diagnosis of SQCC (14). All tumors were re-staged according to the Union for International Cancer Control (UICC) 2017, 8th edition TNM-classification (15). Overall survival (OS) was defined as the time from the resection to death of any cause. For baseline characteristics, see Supplemental Table S1. The study was approved by the Cantonal Ethics Commission of the Canton of Bern (KEK 200/14). A tissue microarray was constructed from formalin-fixed and paraffin-embedded (FFPE) tissue blocks as described before (13). In short, slides were scanned and digitally annotated with subsequent automatic transfer of the punches to a TMA-receptor block, which was used for further analysis. Additionally, full tissue sections were used for comparison between tissue microarray cores and full tissue sections.

Idiopathic pulmonary fibrosis (IPF) patients and tissues

IPF tissues were collected at the Institute of Pathology, Hannover Medical School, Germany (FFPE), as previously described (16). All patients provided written informed consent, and the study was done in accordance with the Declaration of Helsinki. All experiments were performed in accordance with relevant guidelines and regulations (ethical votes #1691-2013 or #3381-2016, Hannover Medical School). In brief, the specimens for primary surgical resection were obtained from patients diagnosed with lung IPF (n=10) and preserved as FFPE material.

In addition, explanted lung tissue from patients with IPF (n=4) and healthy organ donors (n=4) were inflated with air to a transpulmonary pressure of 30 cm H₂O, then deflated to 10 cm H₂O while freezing in liquid nitrogen vapor; frozen samples were stored at -80°C. This study was approved by the hospital ethics and university biosafety committees in Leuven, Belgium (ML6385). IPF tissues and healthy lung tissues were collected at KU Leuven, as previously described (17). All patients provided written informed consent, and the study was done in accordance with the Declaration of Helsinki. For baseline characteristics, see Supplemental Table S2.

Quantification of anthracotic pigment

Tissues were counterstained with nuclear red stain (Fluka, 60700, 0.1%). Stained tissue sections were scanned using an AxioScan.Z1 digital slide scanner (Zeiss, Jena, Germany) equipped with a 20x magnification objective. Quantification of the amount of anthracotic pigments was determined by digital image analysis using the software Definiens Developer XD2 (Definiens AG, Germany), following a previously published procedure (18). The calculated parameter was the ratio of pigment area respective to total tissue area for each core.

High mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry imaging (MSI)

High mass resolution MALDI FT-ICR MSI was performed as previously described (19,20). In brief, FFPE sections (4 µm) or fresh frozen sections (12 µm) were mounted onto indium-tin-oxide (ITO)-coated glass slides (Bruker Daltonik, Bremen, Germany). The air-dried tissue sections were spray-coated with 10 mg/mL of 9-aminoacridine hydrochloride monohydrate matrix (Sigma-Aldrich, Munich, Germany) in methanol (70%) using the SunCollect™ sprayer (Sunchrom, Friedrichsdorf, Germany). Prior to matrix application, FFPE tissue sections were incubated additionally for 1 h at 70°C and deparaffinized in xylene (2 × 8 min). Spray-coating of the matrix was conducted in eight passes (ascending flow rates 10 µL/min, 20 µL/min, and 30 µL/min for layers 1–3 and for layers 4–8 with 40 µL/min), utilizing 2 mm line distance and a spray velocity of 900 mm/min.

Metabolites were detected in negative-ion mode on a 7 T Solarix XR FT-ICR mass spectrometer (Bruker Daltonik) equipped with a dual electrospray ionization MALDI (ESI-MALDI) source and a SmartBeam-II Nd: YAG (355 nm) laser. Mass spectra were acquired within m/z 50–1,100 and a lateral resolution of 50 μm . L-Arginine was used for external calibration in the ESI mode. The SCiLS lab software 2020b was used to export the picked peaks of the mass spectra as processed and root mean square normalized imzML files.

The SPACiAL workflow was used as previously described to automatically annotate tumor and stroma regions in SQCC tissues (21). In short, after MALDI-MSI analysis, the 9-aminoacridine matrix was removed with ethanol (70%) for 5 min from tissue sections, followed by immunohistochemical staining. Double staining of the tissue microarray was performed using pan-cytokeratin (monoclonal mouse pan-cytokeratin plus [AE1/AE3+8/18], 1:75, catalogue no. CM162, Biocare Medical, US, RRID: AB_10582491) and vimentin (abcam, clone ab92547, 1:500, RRID: AB_10562134). Regions positive for pan-cytokeratin were defined as tumor. Regions negative for pan-cytokeratin but positive for vimentin were defined as stroma.

Discovery and visualization of exogenous and endogenous compounds

In order to discover and visualize exogenous and endogenous compounds, mass spectra in and near anthracotic pigments were extracted using the SCiLS lab software 2020b. Annotations were performed using Kyoto Encyclopedia of Genes and Genomes (KEGG, RRID: SCR_012773) (22), The Human Metabolome Database (HMDB, RRID: SCR_007712) (23) and Hoffmann analytes (24).

We performed a stringent annotation of molecules using the following inclusion criteria: (1) The molecular mass of endogenous and exogenous compounds must be between 50 Da and 1100 Da; (2) The signal to noise ratio must be above 2; (3) For exogenous compounds, literature evidence must exist for their presence in tobacco smoke. Exclusion criteria were: (1) Signals that were annotated as isotopes were excluded; (2) As previously published, substances with HMDB descriptions containing

a reference to drugs, pesticides, or other implausible descriptions were excluded (21). M-H, M-H₂O-H, and M+Cl as negative adducts with a mass tolerance of 4 ppm were prioritized.

On tissue measurement of benzo[a]pyrene

Benzo[a]pyrene was purchased from Sigma Aldrich (St. Louis, MO, USA) and diluted in xylene. 1 μ l benzo[a]pyrene solution was spotted onto human fresh frozen lung tissue sections between the absolute amounts of 60 nmol – 0.6 nmol benzo[a]pyrene. Matrix application and high mass resolution MALDI FT-ICR MSI was performed as described before. Stack plot was created by flexImaging (v. 5.0, Bruker), and overlaid peak spectra were illustrated in mMass (v. 5.5.0). Curve fitting was performed with GraphPad Prism (v. 9.2.0).

Immunohistochemical staining

Immunohistochemical staining for cluster of differentiation 3 (CD3), cluster of differentiation 8 (CD8), and programmed death-ligand 1 (PD-L1) was performed as previously described (13) on consecutive sections. In brief, an automated immunostainer (Bond III, Leica Bio-systems, Muttentz, Switzerland) with anti-CD3 (Abcam Cambridge, UK, clone SP7, 1:400, RRID: AB_443425), anti-CD8 (Dako, clone C8/144B, 1:100, RRID: AB_2075537), and anti-PD-L1 (Cell Signaling Technology, clone E1L3N, 1:400, RRID: AB_2687655) was used. The numbers of CD8+ and CD3+ tumor infiltrating lymphocytes were determined using image analysis (Aperio Image Scope) and adjusted for core completeness. PD-L1 expression was assessed as the intensity of membranous staining by a pathologist (SB).

Immunofluorescence analysis of γ H2AX

Immunofluorescence analysis of γ H2AX expression was achieved using primary antibodies against pH2A.X (Cell Signaling, Nr.: # 2577, 1:400, RRID: AB_2118010) and pan-cytokeratin (monoclonal mouse pan-cytokeratin plus [AE1/AE3+8/18], 1:75, catalogue no. CM162, Biocare Medical, US,

RRID: AB_10582491) on consecutive sections. Slides were digitized at 20× objective magnification using an Axio Scan.Z1 (Zeiss). Quantification was performed by digital image analysis in Definiens Developer XD2, following a previously published procedure (18). The quantified parameter was the ratio of γ H2AX and cytokeratin positive cells to the total number of cytokeratin positive cells.

Statistical analysis

Correlations were calculated using pairwise Spearman's rank-order correlation (Python 3.7, SciPy 1.2.0, RRID: SCR_008058). Spearman *P*-values were adjusted with Benjamini/Hochberg correction (Python 3.7, StatsModels 0.9.0). To determine significant differences between UICC TNM stages, Kruskal–Wallis test by ranks (Python 3.7, SciPy 1.2.0) and post hoc Dunn's multiple comparison test (Python 3.7, scikit-posthocs 0.6.1) were used in conjunction with Benjamini/Hochberg correction. Cutoff-optimized survival analyses were performed using a Kaplan–Meier Fitter and log-rank test (Python 3.7, lifelines 0.24.8). Cutoff-optimized in this context means that the threshold for low and high abundance of a compound was chosen such that the *P*-value in the resulting Kaplan-Meier curve is minimal, while ensuring robust results for similar cutoffs.

We investigated the association between the survival time of patients and several predictor variables. The Cox proportional-hazards model is a regression model commonly used in medical research for this purpose. The multivariate analysis was performed using the Cox proportional hazards model (Python 3.8, lifelines 0.25.7) using the same cutoffs as for Kaplan–Meier Fitter. Categorical data were used for the Cox proportional-hazards model. Compounds that passed the non-proportional test were included in the model (Python 3.8, lifelines 0.25.7). All survival calculations were based on overall survival (OS).

Spatial correlation networks

Correlation networks were created with Cytoscape (v. 3.8.0, RRID: SCR_003032) (25). All networks were visualized using the edge weighted spring embedded layout and the absolute value of the

correlation coefficient calculated as described above. Compounds with at least one significant correlation are shown ($P < 0.05$).

Circular plots

Circular plots were generated using Circos (v.0.69.8, RRID: SCR_011798) (26). The metabolites of interest and correlations with the exogenous compounds were extracted from the spatial correlation networks. Pathway information for each metabolite was extracted from KEGG (22). If available, common compound name abbreviations were retrieved from HMDB (23) or KEGG databases.

RESULTS

Carbon pigment is common not only in normal lung tissue, but also in lung squamous cell carcinoma (SQCC).

Carbon deposits can be seen macroscopically (Figure 1A), often found in the center of the tumor (27), beyond parenchymal (Figure 1B) and pleural anthracosis (Figure 1C). They are found intratumorally dispersed with varying degrees (Figure 1D, E). Using digital image analysis, anthracotic pigment was quantified in SQCC tissues of 313 patients (Figure 1F, G, H) for subsequent statistical analyses (e.g., correlations with clinical parameters).

Concentration of carbon pigment in lung SQCC does not correlate with smoking, DNA damage, presence of lymphocytes, programmed death-ligand 1 (PD-L1) expression, or patient survival.

There is no significant association between the amount of carbon pigment and smoking behavior (pack-years, $P=0.91$), DNA damage (γ H2AX expression, $P=0.61$), lymphocyte number (cluster of differentiation 3 (CD3), $P=0.42$; and cluster of differentiation 8 (CD8), $P=0.67$), PD-L1 expression ($P=0.07$), and patient overall survival (OS, $P=0.23$, cutoff=0.005%) (Figure 1 I, J, K). There are no

significant correlations for pigment quantity. Next, the molecular composition was further investigated.

Exogenous compounds such as polycyclic aromatic hydrocarbons (PAHs), tobacco-specific nitrosamines, and aromatic amines are highly abundant in and nearby carbon pigment.

Using high mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry imaging (MSI), eleven exogenous compounds were detected at different abundances throughout the tissues. Importantly, the abundance is highest in and nearby the anthracotic pigment (Figure 2). The exogenous compounds can be grouped into four classes: PAHs, tobacco-specific nitrosamines, aromatic amines, and organohalogens (Figure 2A).

Five PAHs, benzo[a]pyrene (m/z : 287.0639), dibenz(a,h)anthracene (m/z : 313.0800), dibenzo[a,l]pyrene (m/z : 337.0775), benzo[b]pyridine (m/z : 128.0504), and 7-OH-12-methylbenz[a]anthracene sulfate (m/z : 351.0692) are particularly rich in carbon pigment (Figure 2). Within pigment interspersed tissue, three tobacco-specific nitrosamines were detected: 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (nicotine-derived nitrosamine ketone (NNK), m/z : 242.0702), 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL, m/z : 208.1091), and NNAL-N-glucuronide (m/z : 367.1373) (Figure 2B, C). Finally, N-hydroxy-4-aminobiphenyl (m/z : 220.0526) and N-hydroxy-MeIQx (m/z : 264.0650) are aromatic amines and dichloroethane (m/z : 96.9617) is an organohalogen. To determine the quantity of benzo[a]pyrene as an example, we performed a spiking experiment. The minimum amount for detecting benzo[a]pyrene in lung tissues is 2 nmol (Supplemental Figure S1).

In total, the concentration of exogenous compounds correlates with the amount of anthracotic pigment for benzo[a]pyrene ($P=0.0009$), dibenz(a,h)anthracene ($P=0.0056$), dibenzo[a,l]pyrene ($P=0.0405$), NNK ($P=0.0316$), NNAL ($P=0.0338$), and NNAL-N-glucuronide ($P=0.0257$). The correlations

between carbon pigment and exogenous compounds are all positive. In the next, we analyzed the patterns of the exogenous compounds in the carbon pigment within and between the individual patients.

The chemical composition of carbon pigment is highly variable and unique for each patient.

The chemical composition of carbon pigment is heterogeneous in terms of the pattern and abundance of the compounds within the different areas of lung tissue (Figure 3, Figure 3A, Supplemental Figure S2 and S3). The variability of the chemical composition is also visible at the microscopic scale: Figure 3B shows SQCC regions from two patients with unique compositions of carbon-bound compounds. Although the carbon pigments of both regions exhibit a high amount of dibenz(a,h)anthracene, other compounds are present at very different abundances. A multicolor visualization of NNK, NNAL, and NNAL-N-glucuronide also shows an entirely different chemical signature (Figure 3C, Supplemental Figure S4). The chemical composition was shown to be unique and heterogeneous within and between patients. Next, we investigated the differences within the tissue compartments stratified to tumor cells and tumor stroma.

Different quantities and qualities of carbon-bound compounds were found within tumor and stroma.

We used our recently published SPACiAL method for immunophenotype-guided separation of tumor and stromal tissue compartments (21) to investigate the associated carbon pigment (Supplemental Figures S5 and S6). The amount and prevalence of exogenous compounds differ between tumor and stromal regions (Figure 4A, Supplemental Figure S7). Most exogenous molecules were more frequently measured in tumor pigments. However, PAHs and tobacco-specific nitrosamines - if present in stroma - are more abundant there.

The abundance of carbon-bound compounds was next correlated to tumor features and patient characteristics. Figure 4A is a comparative representation of the abundance of compounds correlating with smoking behavior, DNA damage and immunological features.

NNK and dichloroethane are associated with smoking behavior.

Several carbon-bound compounds correlate with smoking behavior, DNA damage, lymphocyte number, PD-L1 expression, and tumor progression (Figure 4). In tumor, NNK and dichloroethane are significantly associated with smoking behavior (Figure 4B, $P=0.0114$ and $P=0.0215$, respectively).

DNA damage in cancer cells and high amounts of PAH are interrelated.

A total of 89.7% of the patients have γ H2AX positive tumor cells, which is indicative of DNA damage. The PAHs benzo[a]pyrene ($P=0.0020$), dibenz(a,h)anthracene ($P=0.0262$), and dibenzo[a,l]pyrene ($P=0.0432$) correlate positively with γ H2AX (Figure 4C).

In tumor tissue, PAHs and NNAL are linked to T cell infiltration and PD-L1 expression.

In tumor regions, dibenz(a,h)anthracene abundance correlates negatively with both the overall number of tumor infiltrating T cells (CD3, $P=0.0429$) and CD8 ($P=0.0450$) (Figure 4D). In stromal areas, dibenzo[a,l]pyrene correlates positively with CD3 and CD8 ($P=0.0292$ and $P=0.0156$, respectively). The intensity of tumoral PD-L1 expression correlates negatively with NNAL abundance ($P=0.0489$).

Tumor stage is associated with benzo[a]pyrene and dibenz(a,h)anthracene.

Benzo[a]pyrene in stroma regions ($P=0.0364$) and dibenz(a,h)anthracene in tumor and stroma regions ($P=0.0400$, $P=0.0439$, respectively) are associated with Union for International Cancer Control (UICC) tumor stages (Figure 4E). In contrast to carbon-bound PAHs, the amount of carbon pigment shows no association with tumor stage ($P=0.1729$).

Patient outcome correlates with amount and spatial location of PAHs, NNKs, and aromatic amines.

Two PAHs, benzo[b]pyridine (tumor, $P=0.0103$) and dibenz(a,h)anthracene (stroma, $P=0.0270$) correlate significantly with OS (Figure 5A, see also Supplemental Figures S8 and S9). Additionally, two tobacco-specific nitrosamines - NNK ($P=0.0071$) and NNAL-N-glucuronide ($P=0.0298$) - are significantly correlated with survival (Figure 5B). However, a higher concentration of NNK is associated with poor OS, while its detoxified form, NNAL-N-glucuronide, is significantly linked to better OS. In contrast, higher amounts of the two aromatic amines - N-hydroxy-MeIQx and N-hydroxy-4-aminobiphenyl - are significantly correlated with poor patient outcome, both in tumor and in stroma (Figure 5C).

We also tested the correlation of CD3, CD8, PD-L1, and pack-years with patient survival. High expression of CD3 ($P=0.0139$) and CD8 ($P=0.0275$) in tumor stroma regions, low expression of PD-L1 in tumor cells ($P=0.0021$), and low pack-years ($P=0.0003$) are associated with good survival (Supplemental Figure S10).

PAH, NNK, and aromatic amines are independent factors for overall survival.

We investigated the association between the survival time of patients and several predictor variables and used the Cox proportional-hazards model. Benzo[b]pyridine ($P=0.0019$), dibenz(a,h)anthracene ($P=0.0150$), NNK ($P=0.0043$), and N-hydroxy-MeIQx ($P=0.0008$) are independent factors for OS (Figure 5D). The highest HR was determined for NNK in stroma (HR=5.0263) and N-hydroxy-MeIQx in tumor (HR=3.0943), indicating that a higher amount of these compounds is deleterious.

After finding significant correlations of individual carbon-bound compounds with tumor features and patient characteristics, we investigated spatial correlations of exogenous and endogenous compounds with a focus on pathway and network analyses.

Spatial correlation networks of metabolites and exogenous compounds reveal substantially different metabolism in tumor and stroma regions.

To investigate the metabolic changes of tumor cells associated with exogenous compound quantities, we evaluated the spatial correlation networks of metabolites in 330 patient tissues. Dense clusters within the networks indicate stronger spatial correlation, and therefore, dependencies between quantities of exogenous and endogenous compounds. Pixel-wise spatial correlations within and between metabolites and eleven exogenous compounds were calculated and filtered ($P < 0.05$). In the two resulting networks, the spatial correlation of 133 metabolites within tumor cells and 159 metabolites in the stroma are visualized (Figure 6).

In tumor, N-hydroxy-MeIQx is associated with altered lipid and glutathione metabolism.

The spatial correlation network in tumor reveals no distinct cluster of exogenous compounds (Figure 6A). N-Hydroxy-MeIQx, which has the most striking effect on patient survival, was detected in a dense cluster of metabolites (max. $r_s = 0.79$), featuring a higher spatial positive correlation with glutathione (GSH, $r_s = 0.408$).

Most endogenous metabolites within the cluster of interest can be associated with lipid metabolism (41.2%), whereby the strongest, albeit not very pronounced positive correlations to N-hydroxy-MeIQx were found for 9(10)-EpOME (9,10-EOA, $r_s = 0.166$), sn-glycero-3-phosphoethanolamine (NGPE, $r_s = 0.173$), and sn-glycerol 3-phosphate (G3P, $r_s = 0.185$). The second most represented pathway is nucleotide metabolism (17.6%). Higher quantities of the purine metabolite deoxyinosine-phosphate (dIMP, $r_s = 0.159$) and the pyrimidine metabolites cytidine (Cyd, $r_s = 0.127$) and deoxycytidine diphosphate (dCDP, $r_s = 0.151$) are associated with an increased N-hydroxy-MeIQx concentration (Figure 6A).

In stroma, PAH and tobacco-specific nitrosamines have a strong impact on amino acid and nucleotide metabolism.

The spatial correlation network for the stroma region differs substantially from the network for the tumor region (Figure 6B). Six exogenous molecules are part of a dense cluster together with endogenous metabolites. The highest spatial correlation was found between the two exogenous compounds dibenzo[a,l]pyrene and dibenz(a,h)anthracene ($r_s=0.679$). Unlike the network for the tumor region, most of the correlating metabolites in the main cluster take part in amino acid or nucleotide metabolism.

Four metabolites with a role in amino acid metabolism are related to tryptophan metabolism. A high abundance of tryptophan metabolites is associated with high PAH and tobacco-specific nitrosamine concentrations. Deoxyadenosine (dA), deoxyinosine (D-Ino), deoxycytidine (dC), and deoxyuridine (dU) are two purine and two pyrimidine metabolites from the nucleotide metabolism pathway which correlate positively with the exogenous compounds (Figure 6B).

Carbon-bound exogenous compounds are also present in idiopathic pulmonary fibrosis (IPF).

In addition to lung cancer, other respiratory pathophysiological conditions, such as interstitial lung diseases, have been linked to environmental pollutants, e.g. due to epigenetic modification (28). With the analysis of IPF tissue, we aim to highlight - albeit not in depth - the presence and possible significance of anthracosis on other respiratory pathophysiological conditions. Similarly to SQCC we found inter and intra-patient heterogeneity of carbon-bound exogenous compounds in both, normal lung parenchyma and IPF. In contrast to the tumor and tumor stroma, spatial correlation networks for IPF tissues reveal two clusters of exogenous compounds and endogenous metabolites. One cluster comprises several PAHs including dibenzo[a,l]pyrene, dibenz(a,h)anthracene, and benzo[a]pyrene, while the other is a mixture of two PAHs - benzo[b]pyrene and 7-hydroxymethyl-12-methylbenz[a]anthracene sulfate - one tobacco-specific nitrosamine and one aromatic amine. Both the

spatially correlating endogenous and exogenous compounds within the clusters and the biological pathways they are related to show only minor similarity to the tumor metabolic networks and pathway analysis. See supplementary information for details (Supplemental Figure S11 - S14).

DISCUSSION

We have discovered a biologically active pool of carbon-bound exogenous compounds in lung cancer tissue. High amounts of these exogenous compounds in various and patient-unique chemical combinations were found in and near anthracotic pigment. Although the detected exogenous compounds are known carcinogens, we show here for the first time that these exogenous compounds also have a strong impact on tumor pathophysiology and survival outcome of lung cancer patients.

Carbon particles accumulate in human lungs and exhibit a large surface area as well as specific surface characteristics for the adsorption of inorganic and organic exogenous compounds (29–31). In mouse lungs, carbon particles were shown to persist indefinitely (32). The long-term persistence and bioavailability of carbon-bound exogenous compounds is supported by further animal studies showing that the detection of benzo[a]pyrene was possible up to 5.6 months after incubation (33).

We used *in situ* high mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry imaging to show that the chemical composition of carbon particles in human lung cancer tissue is much more complex than expected. Each patient showed a unique chemical signature of carbon particles (Figure 3A). Even within one patient, carbon particles exhibit high chemical heterogeneity (Figure 3B). Both extrinsic factors, such as environmental conditions, and intrinsic factors, such as metabolism of exogenous compounds, may play a role for the diverse chemical patterns of carbon pigments. However, our cohort mainly comprises patients with a history of smoking and squamous-cell carcinoma (SQCC) is a clear smoker associated type of cancer (Supplemental Table S1). Hence, the particles analyzed in this study are

likely smoking related. Certain polycyclic aromatic hydrocarbons (PAHs) accumulate in smokers' lungs (34,35), however, their localization and pathophysiological impact remained unclear. Nevertheless, Tomingas *et al.* demonstrated a large increase of benzo[a]pyrene in human bronchial carcinoma in contrast to adjacent tissue (36). In contrast, our study localized exogenous compounds in human lungs and illustrates the significance of these compounds on SQCC depending on their spatial localization. In our patient cohort, NNK, as the most abundant systemic lung carcinogen in cigarette smoke (37), showed the strongest correlation with pack-years (Figure 4B). Indeed, NNK is derived mainly from tobacco smoke. In contrast, PAHs can be derived from numerous environmental sources (38), explaining the lack of a significant correlation.

Unique differences in the metabolism of exogenous compounds are an intrinsic factor for the chemical signature of carbon particles. This is known in carcinogenesis research, and crucial pathways in PAH metabolism arguably differ between patients (39,40). We found that carbon-bound exogenous compounds are present and bioactive in tumor tissue beyond carcinogenesis (Figure 2, 4 and 5), and their abundance may strongly depend on the metabolic activity in individual patients. Our observations revealed a high variation in NNK, its reduced form NNAL, and its detoxification product NNAL-N-glucuronide (41), suggesting unique metabolic activities in patients (Figure 3A).

In our study, an increased concentration of PAHs was associated with increased DNA damage in tumor cells (Figure 4C). Another study focused on carbon particles in mouse lungs and confirmed our finding that anthracosis is associated with DNA damage (32). Alexandrov *et al.* showed that tobacco smoking and PAH exposure cause specific mutations representing the leading mutation signature of lung cancer (42). High mutational burden is associated with an improved objective response to anti-Programmed cell death protein 1 (PD-1) therapy, patient survival, and durable clinical benefit in non-small cell lung cancers (43). Tumor mutational burden and programmed death-ligand 1 (PD-L1) are used as predictive markers for immunotherapies. Thus, our data suggests that these exogenous molecules may influence SQCC therapies.

While we found exogenous compounds as toxic molecules of tobacco smoke and environmental factors, a recent study found the presence of an exogenous molecule as a drug in anthracosis. Cisplatin can be accurately detected in tissues using laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) imaging. Greenhalgh *et al.* (44) applied LA-ICP-MS imaging for 3D *ex vivo* human explant model and demonstrated for the first time a spatial correlation between platinum and anthracosis in lung tissue. The carbon deposits found in lung tissue may affect the movement and thus the efficacy of cisplatin treatment. The authors conclude that cisplatin penetration can be predicted and monitored by LA-ICP-MS imaging as a screening tool (44). This underlines the potential therapeutic implications of anthracosis in non small cell lung carcinoma. Since LA-ICP-MS imaging is capable of measuring metals, it would be interesting to apply this MSI technology to measure metals in the anthracotic pigment, which are also present in cigarette smoke (45).

Lymphocytes are known to be associated with harmful compounds in lung cancer tissues. In our study, the infiltration of CD3+ and CD8+ lymphocytes was correlated with the concentrations of dibenz(a,h)anthracene and dibenzo[a,l]pyrene (Figure 4D). Many PAHs influence patient immunity, as they are important ligands of the aryl hydrocarbon receptor (AhR) on several immune cells (46,47). Systemic AhR activation by an exogenous compound triggers the suppression of the CD8+ T cell response in infected lungs (48). In our study, a suppression of T cells was observed with a higher concentration of dibenz(a,h)anthracene in tumor regions. Another study showed that tobacco smoke and benzo[a]pyrene lead to CD8+ lymphocyte enrichment in mouse lungs, possibly induced by AhR (49). In this study, a similar effect of lymphocyte enrichment was seen in the stroma regions for high concentrations of dibenzo[a,l]pyrene.

We have shown for the first time that exogenous compounds have a strong impact on patient survival (Figure 5). The presence of NNK, NNAL-N-glucuronide, and N-hydroxy-4-aminobiphenyl in SQCC tissues was associated with patient outcomes. A study on 770 resected lung cancers revealed poorer prognosis for smokers (50). We observed the same effect of smoking on survival (Supplemental Figure S10). All substances with a significant effect on survival can be associated with tobacco

smoke. Interestingly, the non-metabolized primary substance from tobacco smoke, NNK, was associated with poor OS. In contrast, the detoxified variant of the metabolite, NNAL-N-glucuronide, was an indicator of a favorable prognosis (Figure 5B). Hence, glucuronidation may be worth investigating in more detail for detoxification of exogenous compounds. This is underlined by another study, which revealed that smokers with an increased urinary level of glucuronidated nitrosamine N'-nitrosornicotine experienced a significantly reduced risk of esophageal cancer (51). Whether and to what extent glucuronidation in lungs of human individuals can be specifically influenced still remains open.

Correlations between endogenous metabolites and exogenous molecules are fundamentally different between tumor and stroma regions. Notably, in tumor regions, N-hydroxy-MeIQx alters lipid and glutathione metabolism. A higher abundance of the metabolite N-hydroxy-MeIQx is associated with increased lipid species and a higher concentration of glutathione (Figure 6A). The effect of MeIQx on lipid profile, potentially caused by dysregulated maturation of autolysosomes, has been described in hepatocytes (52). Otherwise, glutathione has been shown to increase the perceptivity to oxidative stress (53). The formation of 8-hydroxy-2'-deoxyguanosine, an oxidative DNA damage marker, increases in the liver with a specific dose of MeIQx (54). Oxidative stress caused by N-hydroxy-MeIQx may be contained by an increase in glutathione in tumor cells.

In tumor stroma, PAHs and tobacco-specific nitrosamines revealed a strong impact on amino acid and nucleotide metabolism (Figure 6B). Notably, the majority of the changed amino acid metabolites can be associated with tryptophan metabolism, which is of crucial importance to the immune system as its metabolites orchestrate local and systemic responses to control inflammation (55). The indole ring of the critical regulatory molecule tryptophan is cleaved by indoleamine 2,3-dioxygenase (IDO) (56). The activation of AhR increases IDO expression (57). Therefore, PAHs may increase tryptophan metabolism through the activation of AhR. Our second finding that nucleotide metabolism is enhanced is most likely associated with DNA damage and repair caused by these substances.

The effects of carbon-bound exogenous compounds might also be of pathophysiological significance for other lung diseases. Anthracosis has been described to be associated with non-neoplastic diseases such as emphysema (32). We have selected idiopathic pulmonary fibrosis (IPF) as an example for an exploratory analysis of anthracotic pigment in the context of non-neoplastic diseases. IPF is a lung disease of unknown etiology and is characterized by progressive scarring. The underlying pathomechanisms of IPF, with its complex immunological and inflammatory processes and external impacts, have been the focus of recent research. Lifestyle and environmental influences are held responsible for much of its natural history. Because smoking, pneumotoxic medications, and inhalation of dust are known risk factors of IPF (58), we analyzed the presence and constitution of exogenous compounds within anthracotic tissue of IPF patients to uncover differences to the smoking-related SQCC. Indeed, we also found carbon-bound exogenous molecules in IPF anthracotic pigment. Similarly to SQCC we found inter and intra-patient heterogeneity of carbon-bound exogenous compounds in both, normal lung parenchyma and IPF. Furthermore, the network analysis revealed differences in the affected metabolic pathways compared to SQCC tissues. We conclude that exogenous compounds could be an unrecognized factor in the development and progression of IPF. These preliminary findings warrant further investigation.

When the amount of anthracotic pigment is considered over the total volume of both lungs, there is a large and persistent pool of carbon-bound exogenous compounds with possible systemic effects beyond the lungs. Tobacco smoking is also the leading risk factor for bladder cancer (41). As a representative of aromatic amines, 4-aminobiphenyl has been extensively studied to understand the mechanism of bladder carcinogenesis (59). We discovered that N-hydroxy-4-aminobiphenyl, a carbon-bound exogenous compound, was highly abundant in human lung tissue. Given the potentially large amount of anthracotic pigment in both lungs, it is possible that 4-aminobiphenyl is stored in the pigment and released continuously over the long term and thus may contribute to the development of bladder carcinoma. Similarly, other carcinogens could be stored and continuously released via the persistent carbon pool and thus also play a role in the development of tumors outside the lung.

Since carbon particles and carbon-bound exogenous compounds are known to be persistent, and the subsequent removal of the carbon particles from the lung is as of yet not feasible, the most reasonable and implementable courses of action right now are risk assessment and prevention. Apart from smoking cessation, a change in smoking behavior may influence concentrations of biomarkers of exposure. For example, in a large scale study including 5105 participants, e-cigarette users showed a 10% to 98% lower concentration of PAHs compared to exclusive cigarette smokers - albeit it was still significantly increased compared to the levels in never smokers (60). Based on another study, the most significant determinants of PAH exposure beyond smoking are diet and indoor exposures like coal- or wood-heaters, cooking, diverse leisure activities and passive tobacco smoke - and most of these exposures can be deemed preventable (61). In the case of PAHs, there are several physical and chemical remediation strategies to remove PAHs from polluted environments including membrane filtration, soil washing, adsorption, electrokinetic, thermal, oxidation, and photocatalytic treatments (62). Given that we here show that several exogenous compounds are present directly within anthracotic pigment and that they are an unrecognized factor with strong impact on tumor pathophysiology underlines the importance of risk assessment and prevention.

In conclusion, the bioactive pool of exogenous compounds in and nearby the anthracotic pigment is associated with several changes in tumor pathophysiology and has adverse effects on patient outcome. Genome integrity, immune factors, and tumor progression are associated with specific chemical signatures in the anthracotic pigment. The exact localization of exogenous compounds influences patient outcome by altering amino acid, nucleotide, and lipid metabolism. In lung IPF, exogenous substances can also be found in and nearby anthracotic pigment, however, these molecules affect other pathways (Supplemental Figure S11). The exogenous compounds may contribute to the formation and influence the progression of diseases of the lung and other organs. Because all healthy lung tissues contained exogenous compounds (Supplemental Figure S14), a deeper understanding of the unique composition and pathophysiological relevance of anthracosis is needed.

ACKNOWLEDGMENTS

We thank Ulrike Buchholz, Claudia-Mareike Pflüger, Andreas Voss, Cristina Hübner Freitas, and Elenore Samson for excellent technical assistance.

The study was funded by the Ministry of Education and Research of the Federal Republic of Germany (BMBF; Grant Nos. 01ZX1610B and 01KT1615), the Deutsche Forschungsgemeinschaft (Grant Nos. SFB 824TP C04, CRC/TRR 205 S01) and the Deutsche Krebshilfe (No. 70112617) to A. Walch. Funding was provided through the Impulse and Networking Fund of the Helmholtz Association and the Helmholtz Zentrum München (Helmholtz Enterprise-2018-6) to A. Buck.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Takano APC, Justo LT, Dos Santos NV, Marquezini MV, de André PA, da Rocha FMM, et al. Pleural anthracosis as an indicator of lifetime exposure to urban air pollution: An autopsy-based study in Sao Paulo. *Environ Res.* 2019;173:23–32.
2. Saxena RK, McClure ME, Hays MD, Green FHY, McPhee LJ, Vallyathan V, et al. Quantitative assessment of elemental carbon in the lungs of never smokers, cigarette smokers, and coal miners. *J Toxicol Environ Health A.* 2011;74:706–15.
3. Wang D, Minami Y, Shu Y, Konno S, Iijima T, Morishita Y, et al. The implication of background anthracosis in the development and progression of pulmonary adenocarcinoma. *Cancer Sci.* 2003;94:707–11.
4. Konno S, Morishita Y, Fukasawa M, Shu Y, Wang D, Tanaka R, et al. Anthracotic index and DNA methylation status of sputum contents can be used for identifying the population at risk of lung carcinoma. *Cancer.* 2004;102:348–54.
5. Mirsadraee M. Anthracosis of the lungs: etiology, clinical manifestations and diagnosis: a review. *Tanaffos.* 2014;13:1–13.
6. Churg A. The uptake of mineral particles by pulmonary epithelial cells. *Am J Respir Crit Care Med.* 1996;154:1124–40.

7. Mitchev K, Dumortier P, De Vuyst P. “Black Spots” and hyaline pleural plaques on the parietal pleura of 150 urban necropsy cases. *Am J Surg Pathol*. 2002;26:1198–206.
8. Sun JD, Wolff RK, Kanapilly GM, McClellan RO. Lung retention and metabolic fate of inhaled benzo(a)pyrene associated with diesel exhaust particles. *Toxicol Appl Pharmacol*. 1984;73:48–59.
9. Nicole AH Janssen, Miriam E Gerlofs-Nijland, Timo Lanki, Raimo O Salonen, Flemming Cassee, Gerard Hoek, Paul Fischer, Bert Brunekreef, Michal Krzyzanowski. Health effects of black carbon. Bohr R, editor. World Health Organization; 2012.
10. Oh S-Y, Chiu PC. Graphite- and soot-mediated reduction of 2,4-dinitrotoluene and hexahydro-1,3,5-trinitro-1,3,5-triazine. *Environ Sci Technol*. 2009;43:6983–8.
11. Norris JL, Caprioli RM. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem Rev*. 2013;113:2309–42.
12. Nie Z, Jiang Y, Sun J, Xiong C, Liu H, Li Y, et al. Mass spectrometry imaging reveals in situ behaviors of multiple components in aerosol particles. *Angew Chem Int Ed Engl* [Internet]. 2021; Available from: <http://dx.doi.org/10.1002/anie.202103874>
13. Keller MD, Nepl C, Irmak Y, Hall SR, Schmid RA, Langer R, et al. Adverse prognostic value of PD-L1 expression in primary resected pulmonary squamous cell carcinomas and paired mediastinal lymph node metastases. *Mod Pathol*. 2018;31:101–10.
14. W. H. O. Classification WHO Classification of Tumours Editorial Board. Thoracic Tumours: WHO Classification of Tumours. International Agency for Research on Cancer (I A R C) (UN); 2021.
15. Nepl C, Keller MD, Scherz A, Dorn P, Schmid RA, Zlobec I, et al. Comparison of the 7th and 8th Edition of the UICC/AJCC TNM Staging System in Primary Resected Squamous Cell Carcinomas of the Lung-A Single Center Analysis of 354 Cases. *Front Med*. 2019;6:196.
16. Aichler M, Kunzke T, Buck A, Sun N, Ackermann M, Jonigk D, et al. Molecular similarities and differences from human pulmonary fibrosis and corresponding mouse model: MALDI imaging mass spectrometry in comparative medicine. *Lab Invest*. 2018;98:141–9.
17. Sun N, Fernandez IE, Wei M, Witting M, Aichler M, Feuchtinger A, et al. Pharmacometabolic response to pirfenidone in pulmonary fibrosis detected by MALDI-FTICR-MSI. *Eur Respir J* [Internet]. 2018;52. Available from: <http://dx.doi.org/10.1183/13993003.02314-2017>
18. Feuchtinger A, Stiehler T, Jütting U, Marjanovic G, Lubner B, Langer R, et al. Image analysis of immunohistochemistry is superior to visual scoring as shown for patient outcome of esophageal adenocarcinoma. *Histochem Cell Biol*. 2015;143:1–9.
19. Ly A, Buck A, Balluff B, Sun N, Gorzolka K, Feuchtinger A, et al. High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. *Nat Protoc*. 2016;11:1428–43.
20. Aichler M, Borgmann D, Krumsiek J, Buck A, MacDonald PE, Fox JEM, et al. N-acyl Taurines and Acylcarnitines Cause an Imbalance in Insulin Synthesis and Secretion Provoking β Cell Dysfunction in Type 2 Diabetes. *Cell Metab*. 2017;25:1334–47.e4.
21. Prade VM, Kunzke T, Feuchtinger A, Rohm M, Lubner B, Lordick F, et al. De novo discovery of

- metabolic heterogeneity with immunophenotype-guided imaging mass spectrometry. *Mol Metab.* 2020;36:100953.
22. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
 23. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 2018;46:D608–17.
 24. Hoffmann D, Hoffmann I. Letters to the Editor - Tobacco smoke components. *Beiträge zur Tabakforschung International/Contributions to Tobacco Research.* 1998;18:49–52.
 25. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13:2498–504.
 26. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* 2009;19:1639–45.
 27. Travis WD, Nicholson AG. WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. International Agency for Research on Cancer; 2015.
 28. Goobie GC, Nouraie M, Zhang Y, Kass DJ, Ryerson CJ, Carlsten C, et al. Air Pollution and Interstitial Lung Diseases: Defining Epigenomic Effects. *Am J Respir Crit Care Med.* 2020;202:1217–24.
 29. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Carbon black, titanium dioxide, and talc. *IARC Monogr Eval Carcinog Risks Hum.* 2010;93:1–413.
 30. World Health Organization. Regional Office for Europe. Air Quality Guidelines for Europe. WHO Regional Office Europe; 2000.
 31. Möller W, Felten K, Sommerer K, Scheuch G, Meyer G, Meyer P, et al. Deposition, retention, and translocation of ultrafine particles from the central airways and lung periphery. *Am J Respir Crit Care Med.* 2008;177:426–32.
 32. You R, Lu W, Shan M, Berlin JM, Samuel EL, Marcano DC, et al. Nanoparticulate carbon black in cigarette smoke induces DNA cleavage and Th17-mediated emphysema. *Elife.* 2015;4:e09623.
 33. Gerde P, Muggenburg BA, Lundborg M, Tesfaigzi Y, Dahl AR. Respiratory epithelial penetration and clearance of particle-borne benzo[a]pyrene. *Res Rep Health Eff Inst.* 2001;5–25; discussion 27–32.
 34. Goldman R, Enewold L, Pellizzari E, Beach JB, Bowman ED, Krishnan SS, et al. Smoking increases carcinogenic polycyclic aromatic hydrocarbons in human lung tissue. *Cancer Res.* 2001;61:6367–71.
 35. Lodovici M, Akpan V, Giovannini L, Migliani F, Dolaro P. Benzo[a]pyrene diol-epoxide DNA adducts and levels of polycyclic aromatic hydrocarbons in autoptic samples from human lungs. *Chem Biol Interact.* 1998;116:199–212.
 36. Tomingas R, Pott F, Dehnen W. Polycyclic aromatic hydrocarbons in human bronchial carcinoma. *Cancer Lett.* 1976;1:189–95.
 37. Chapman S. Tobacco: Science, Policy and Public Health. *BMJ.* 2005;330:970.2.

38. Nielsen T. Traffic contribution of polycyclic aromatic hydrocarbons in the center of a large city. *Atmos Environ*. 1996;30:3481–90.
39. Hecht SS, Hochalter JB, Carmella SG, Zhang Y, Rauch DM, Fujioka N, et al. Longitudinal study of [D10]phenanthrene metabolism by the diol epoxide pathway in smokers. *Biomarkers*. 2013;18:144–50.
40. Carmella SG, Chen M, Yagi H, Jerina DM, Hecht SS. Analysis of phenanthrols in human urine by gas chromatography-mass spectrometry: potential use in carcinogen metabolite phenotyping. *Cancer Epidemiol Biomarkers Prev*. 2004;13:2167–74.
41. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer*. 2003;3:733–44.
42. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415–21.
43. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348:124–8.
44. Greenhalgh CJ, Karekla E, Miles GJ, Powley IR, Costa C, de Jesus J, et al. Exploration of Matrix Effects in Laser Ablation Inductively Coupled Plasma Mass Spectrometry Imaging of Cisplatin-Treated Tumors. *Anal Chem*. 2020;92:9847–55.
45. Bernhard D, Rossmann A, Wick G. Metals in cigarette smoke. *IUBMB Life*. 2005;57:805–9.
46. Shimizu Y, Nakatsuru Y, Ichinose M, Takahashi Y, Kume H, Mimura J, et al. Benzo[a]pyrene carcinogenicity is lost in mice lacking the aryl hydrocarbon receptor. *Proc Natl Acad Sci U S A*. 2000;97:779–82.
47. Gutiérrez-Vázquez C, Quintana FJ. Regulation of the Immune Response by the Aryl Hydrocarbon Receptor. *Immunity*. 2018;48:19–33.
48. Lawrence BP, Roberts AD, Neumiller JJ, Cundiff JA, Woodland DL. Aryl hydrocarbon receptor activation impairs the priming but not the recall of influenza virus-specific CD8⁺ T cells in the lung. *J Immunol*. 2006;177:5819–28.
49. Wang G-Z, Zhang L, Zhao X-C, Gao S-H, Qu L-W, Yu H, et al. The Aryl hydrocarbon receptor mediates tobacco-induced PD-L1 expression and is associated with response to immunotherapy. *Nat Commun*. 2019;10:1125.
50. Hanagiri T, Sugio K, Mizukami M, Ichiki Y, Sugaya M, Yasuda M, et al. Significance of Smoking as a Postoperative Prognostic Factor in Patients with Non-small Cell Lung Cancer. *J Thorac Oncol*. 2008;3:1127–32.
51. Yuan J-M, Knezevich AD, Wang R, Gao Y-T, Hecht SS, Stepanov I. Urinary levels of the tobacco-specific carcinogen N⁷-nitrosonornicotine and its glucuronide are strongly associated with esophageal cancer risk in smokers. *Carcinogenesis*. 2011;32:1366–71.
52. Song D, Guo R, Huang H, Zheng P, Huang H, Oyang Q, et al. 2-Amino-3,8-dimethylimidazo[4,5-*g*]quinoxaline Alters Autophagosome Maturation, Cellular Lipidomic Profiles, and Expression of Core Pluripotent Factors. *J Agric Food Chem*. 2019;67:7977–85.
53. Traverso N, Ricciarelli R, Nitti M, Marengo B, Furfaro AL, Pronzato MA, et al. Role of

- glutathione in cancer progression and chemoresistance. *Oxid Med Cell Longev*. 2013;2013:972913.
54. Kakehashi A, Wei M, Fukushima S, Wanibuchi H. Oxidative stress in the carcinogenicity of chemical carcinogens. *Cancers* . 2013;5:1332–54.
 55. Routy J-P, Routy B, Graziani GM, Mehraj V. The Kynurenine Pathway Is a Double-Edged Sword in Immune-Privileged Sites and in Cancer: Implications for Immunotherapy. *Int J Tryptophan Res*. 2016;9:67–77.
 56. Grohmann U, Fallarino F, Puccetti P. Tolerance, DCs and tryptophan: much ado about IDO. *Trends Immunol*. 2003;24:242–8.
 57. Julliard W, Fechner JH, Mezrich JD. The aryl hydrocarbon receptor meets immunology: friend or foe? A little of both. *Front Immunol*. 2014;5:458.
 58. Schäfer SC, Funke-Chambour M, Berezowska S. [Idiopathic pulmonary fibrosis-epidemiology, causes, and clinical course]. *Pathologe*. 2020;41:46–51.
 59. Beland FA, Kadlubar FF. Metabolic Activation and DNA Adducts of Aromatic Amines and Nitroaromatic Hydrocarbons. *Chemical Carcinogenesis and Mutagenesis I*. Springer, Berlin, Heidelberg; 1990. page 267–325.
 60. Goniewicz ML, Smith DM, Edwards KC, Blount BC, Caldwell KL, Feng J, et al. Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. *JAMA Netw Open*. 2018;1:e185937.
 61. Pavanello S, Campisi M, Mastrangelo G, Hoxha M, Bollati V. The effects of everyday-life exposure to polycyclic aromatic hydrocarbons on biological age indicators. *Environ Health*. 2020;19:128.
 62. Patel AB, Shaikh S, Jain KR, Desai C, Madamwar D. Polycyclic Aromatic Hydrocarbons: Sources, Toxicity, and Remediation Approaches. *Front Microbiol*. 2020;11:562813.

FIGURE LEGENDS

Figure 1: Carbon pigment is abundant in both normal lung and pulmonary squamous cell carcinoma (SQCC) tissue. **A:** Gross appearance of lung SQCC tissue. The tumor tissue has grayish, focal areas with carbon pigment deposits at the center (a). Additionally, anthracosis with carbon deposits are seen in lung parenchyma (b) and pleura (c). **B and C:** Histology of normal lung tissue exhibiting carbon pigment deposits (hematoxylin and eosin staining). The pigment accumulates in the cytoplasm of macrophages in the bronchial wall. **D and E:** Histopathology of lung SQCC tissue with dispersed intratumoral carbon deposits. High magnification shows the close spatial relationship of carbon particles and cancer cells. **F and G:** Carbon deposits in SQCC (nuclear red stain) and segmentation (blue) by image analysis for the quantification of carbon particles. Subsequent analyses are based on pigment quantification (H-K). **H:** Distribution of pigment amount within tumor tissue. Patients can be divided into those with no, low, or high pigment content. **I:** Distribution of feature characteristics relating to pigment content, smoking behavior, DNA damage, and CD3, CD8, and PD-L1 expression. **J:** Spearman's rank correlation of the pigment area with feature characteristics, showing no significant correlation. **K:** Survival analysis showing that pigment abundance does not correlate with patient survival (n=234, cutoff=0.005%).

Figure 2: Carbon-bound exogenous compounds detected with high mass resolution matrix-assisted laser desorption/ionization (MALDI) fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry imaging. **A:** Skyline spectrum showing maximum peak intensities between 90 and 375 Dalton. Exogenous compounds are highlighted and colored according to their respective class: polycyclic aromatic hydrocarbons (PAHs) (blue), tobacco-specific nitrosamines (red), aromatic amines (green), and an organohalogen (grey). **B:** Tissue region featuring high carbon pigment content (top left, nuclear red stain) and ion distribution of dibenzo[a,l]pyrene, dibenz(a,h)anthracene, NNK, NNAL, and NNAL-N-glucuronide. Note that all five show a close spatial relationship to the pigment, but also differing distribution patterns. Although NNK has focal high intensity in dense carbon deposits, NNAL-N-glucuronide is conversely distributed within the pigment. **C:** Tumor tissue region featuring extensive intratumoral carbon pigment deposits (top left, nuclear red stain). Spatial organization of dibenzo[a,l]pyrene, dibenz(a,h)anthracene, NNK, benzo[a]pyrene, and 7-OH-12-methylbenz[a]anthracene sulfate and the intratumoral carbon pigment. There are obvious differences in the abundance and distribution pattern of PAHs and NNK and that of carbon pigment, indicating intratumoral heterogeneity.

Figure 3: Inter and intra-tumoral heterogeneity in the chemical composition of carbon pigment in squamous cell carcinoma (SQCC). **A:** Signal intensities of carbon-bound compounds in the tissues of ten patients, illustrating the unique and heterogeneous chemical composition of carbon pigment. See Supplemental Figure S3 for all patients. **B:** SQCC tissues from two patients with comparable intratumoral carbon depositions (nuclear red stains) and distribution of carbon-bound compounds: NNK, NNAL, NNAL-N-glucuronide, dibenz(a,h)anthracene, benzo[a]pyrene, and 7-OH-12-methylbenz[a]anthracene sulfate. **C:** Visualization of NNK, NNAL, and NNAL-N-glucuronide demonstrates heterogeneity within one patient tissue. See Supplemental Figure S4 for increased visibility.

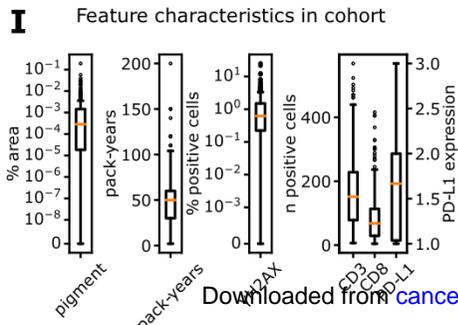
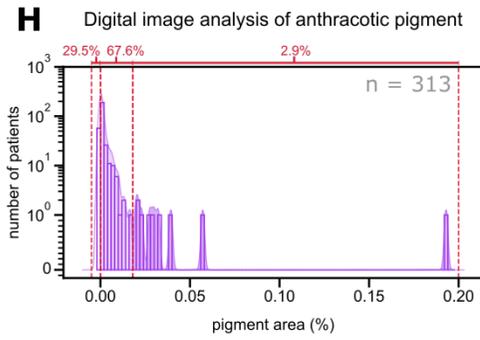
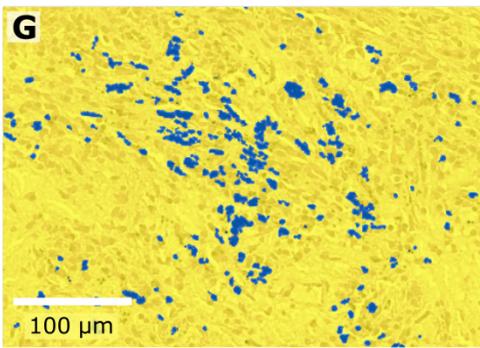
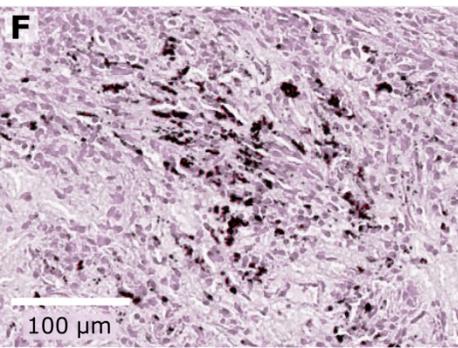
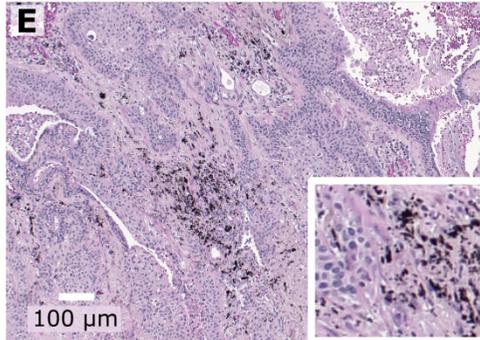
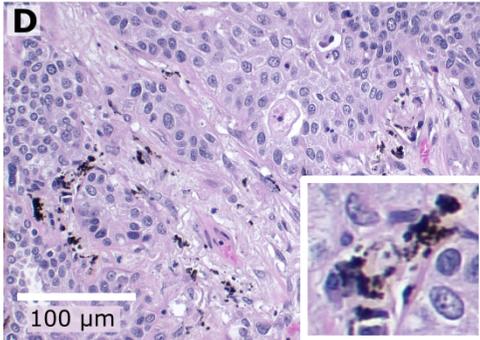
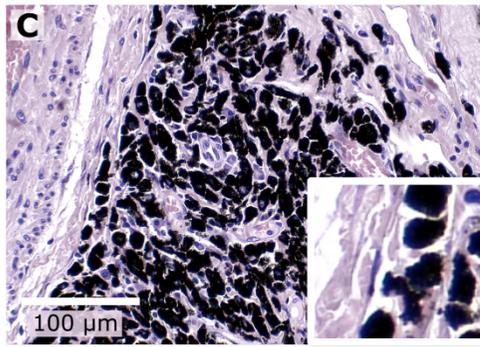
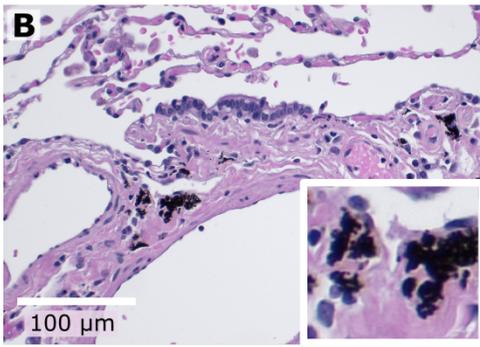
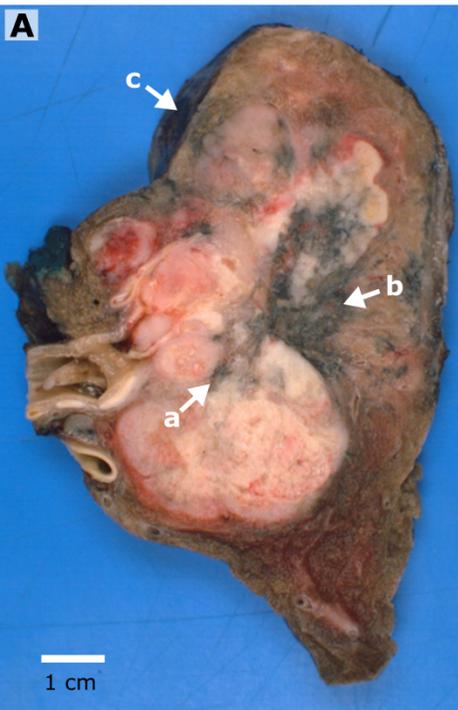
Figure 4: Concentration and prevalence of carbon-bound exogenous compounds and their correlation with smoking behavior, DNA damage, lymphocyte number, PD-L1 expression, and tumor progression. **A:** Intensity distribution of carbon-bound compounds in the tumor (T) and stromal (S) regions. The non-cumulative histograms are visualized as heatmaps to facilitate visual comparison. Since the logarithmic intensities are shown, the counts of intensity=0 are separately shown on the left of each row. To the right of each row the maximum number of patients in a bin is shown. **B:** Significant correlation between NNK and dichloroethane and pack-years (left) and distribution of pack-years (non-cumulative histogram, right). **C:** Significant correlation between polycyclic aromatic hydrocarbons (PAHs) and γ H2AX (left) and distribution of γ H2AX percent positive cells (non-cumulative histogram, right). **D:** Significant correlation between compounds with immunological features (top left) and distribution of the number of positive cells or expression per feature characteristic (histograms). **E:** Benzo[a]pyrene (S), dibenz(a,h)anthracene (T), and dibenz(a,h)anthracene (S), but not carbon pigment are associated with the Union for International Cancer Control (UICC) tumor stage.

Figure 5: The abundance of carbon-bound exogenous compounds is an independent factor for patient outcome. Kaplan–Meier survival analyses (left) and distribution of compound abundance (non-cumulative histograms, right), including the intensity threshold used to split the collective (yellow) for polycyclic aromatic hydrocarbons (PAHs) (**A**), nitrosamines (**B**), and aromatic amines (**C**). On the right, the y-axes are annotated with the maximum frequency per distribution. Only compounds with a significant separation are shown here ($P < 0.05$). All Kaplan–Meier curves, as well as the histograms for all exogenous compounds can be found in the supplementary data (Supplemental Figures S8 and S9). **D:** Cox proportional hazard model for the shown compounds with significant separation in the Kaplan–Meier analyses (log-rank test) and which passed the non-proportional test, as well as the Union for International Cancer Control (UICC) stage. All but two

compounds remained significant in multivariate analysis, indicating that they are independent factors for patient survival.

Figure 6: Spatial correlation between carbon-bound compounds and endogenous metabolites in (A, n=313) tumor and (B, n=268) stroma tissue. Nodes in the spatial correlation networks (left) represent endogenous metabolites (white) and carbon-bound compounds (red). Edges represent positive (blue) and negative (red) spatial correlations, with edge opacity increasing with the correlation coefficient. Circular plots (right) focus on the highlighted compounds from the networks and exclusively on correlations with carbon-bound compounds (center). Tracks: (1) pathway information; (2) compound abbreviation; (3) histogram of the min/max scaled compound signal intensities. Note that within the stroma, multiple carbon-bound compounds form a dense cluster with endogenous metabolites mainly involved in amino acid and nucleotide metabolism, whereas in the tumor, only N-hydroxy-MeIQx is part of a cluster of metabolites involved in lipid metabolism. **Abbreviations (A):** L-cysteate (Cysteate), D-glucose 6-phosphate (G6P), N-hydroxy-MeIQx (OH-MeIQx), glutathione (GSH), (9Z)-stearic acid (9Z-SA), stearic acid (SA), palmitic acid (PLM), sn-glycerol 3-phosphate (G3P), sn-glycero-3-phosphoethanolamine (NGPE), 9(10)-EpOME (9,10-EOA), cholesterol sulfate (CholS), cytidine (Cyd), D-4'-phosphopantothenate (PanP), inositol 1,3,4,5-tetraphosphate (InsP4), CMP-2-aminoethylphosphonate (CMPciliate). **Abbreviations (B):** N-acetylmethionine (AOR), L-homocysteine (Hcy), 5-hydroxy-L-tryptophan (5-HTP), L-kynurenine (L-KYN), L-formylkynurenine (NFK), formyl-N-acetyl-5-methoxykynurenamine (AFMK), pyridoxamine (PM), sedoheptulose 7-phosphate (Sed-7P), benzo[a]pyrene (BP), dibenz(a,h)anthracene (DBahA), dibenzo[a,l]pyrene (DBP), NNAL N-glucuronide (NNAL-NG), 4a-hydroxytetrahydrobiopterin (4a-HTHB), 2,5-diaminopyrimidine nucleoside triphosphate (DAPNTP), gamma-L-glutamyl-L-cysteine (g-Glu-Cys), deoxyadenosine (dA), deoxyinosine (D-Ino), deoxycytidine (dC), deoxyuridine (dU), pantothenate (Vit B5).

Figure 1



Spearman's rank correlation of pigment

coefficient	0.01	0.05	0.07	0.04	-0.16
P-value	0.91	0.61	0.42	0.67	0.07
	pack-years	YH2AX	CD3	CD8	PD-L1

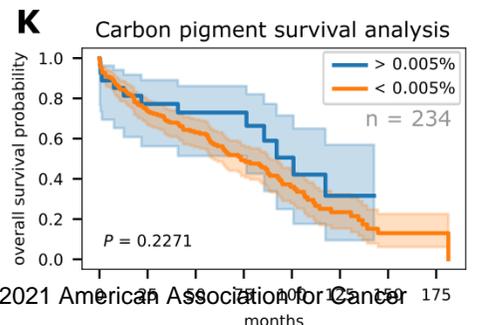


Figure 2

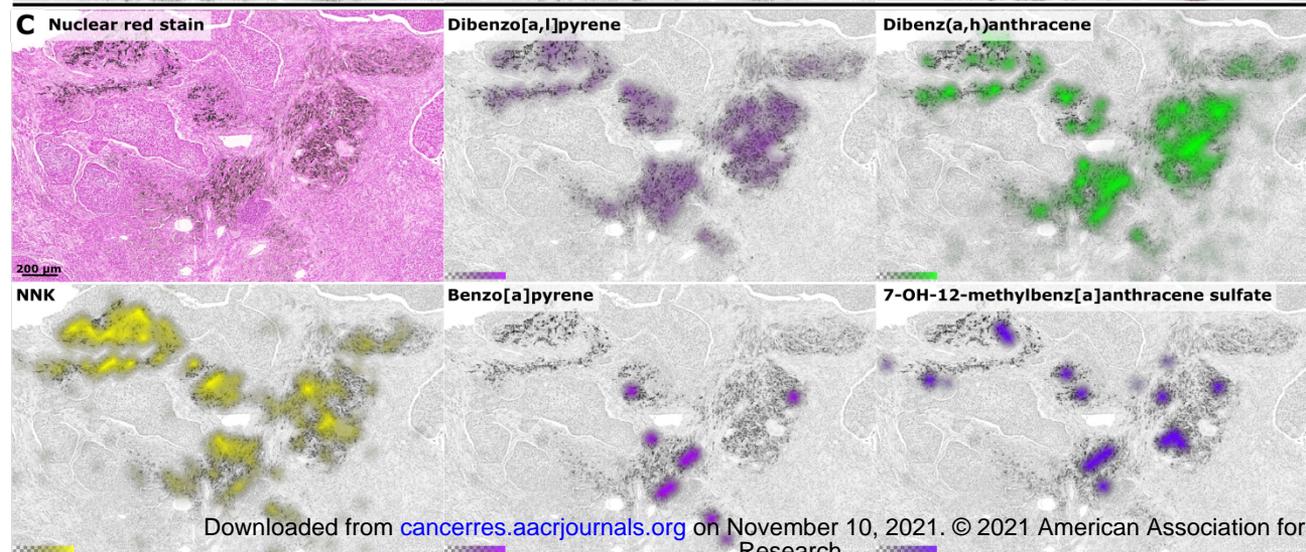
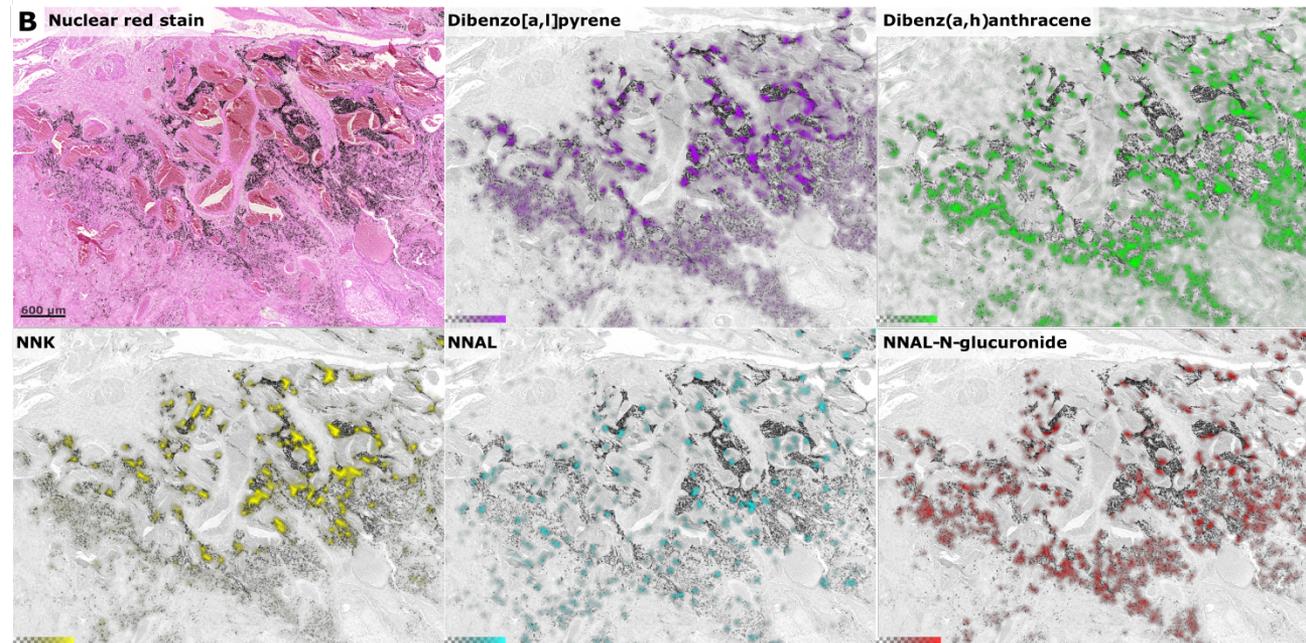
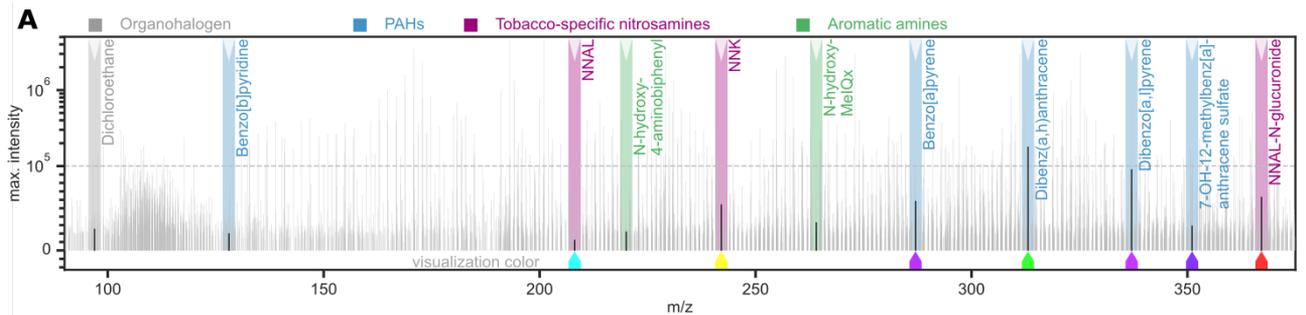


Figure 3

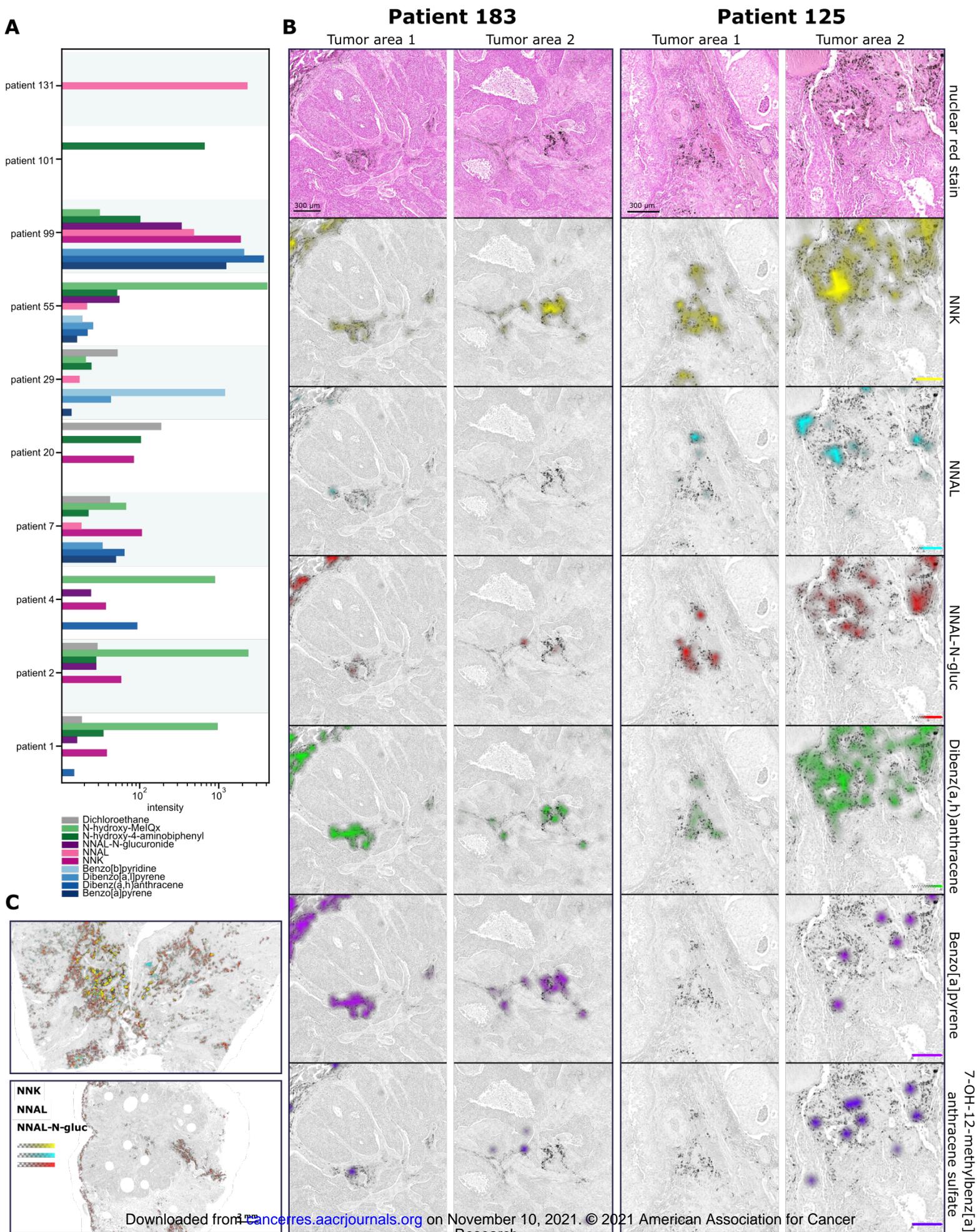
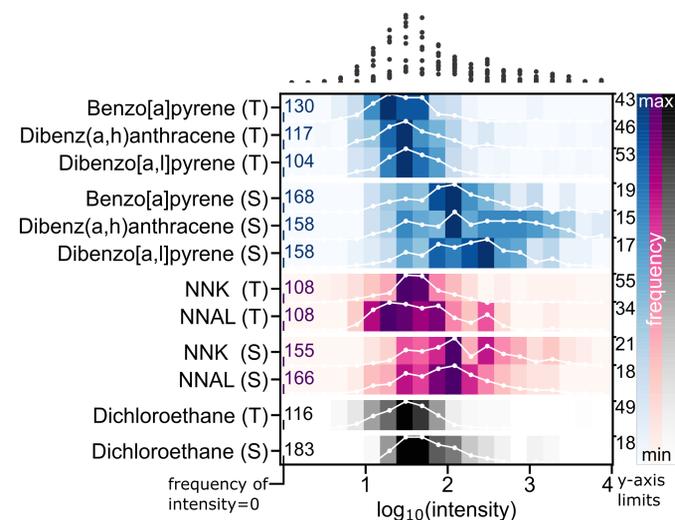
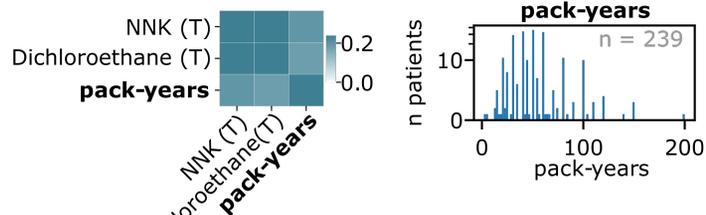


Figure 4

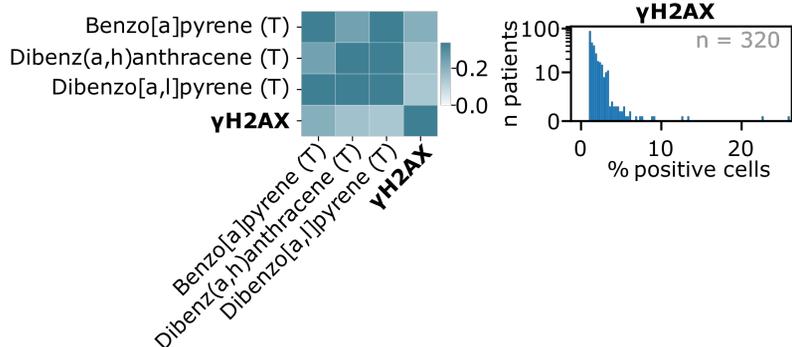
A Compound abundance



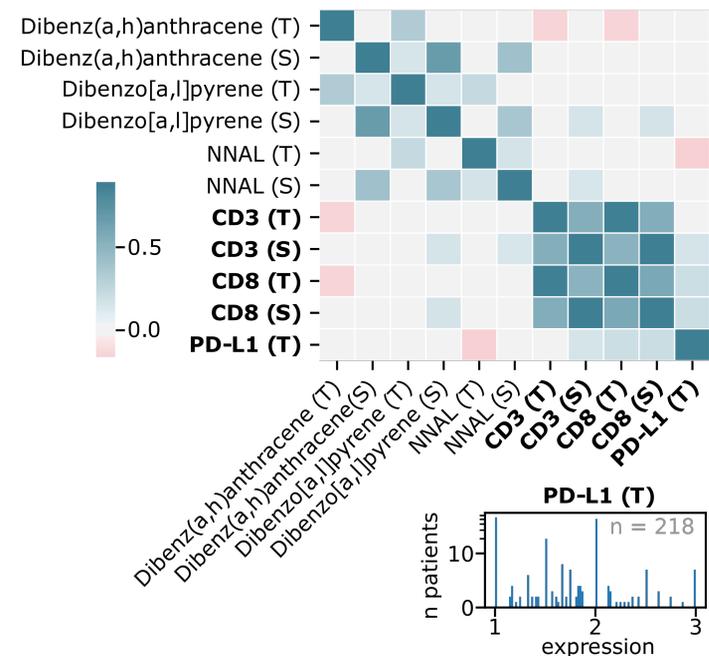
B Smoking behavior



C DNA damage



D Lymphocytes



E Tumor progression

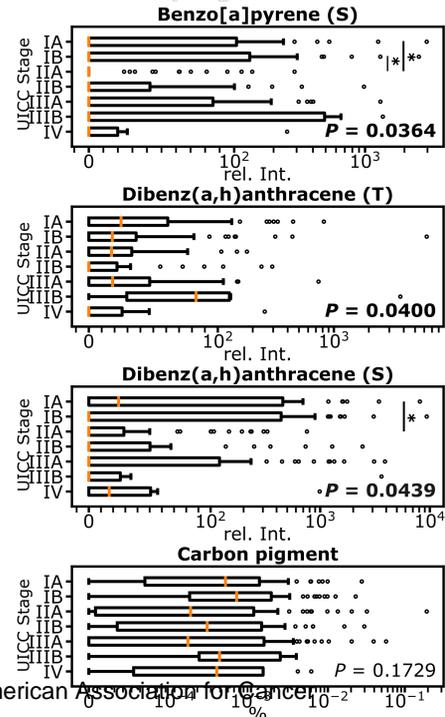


Figure 5

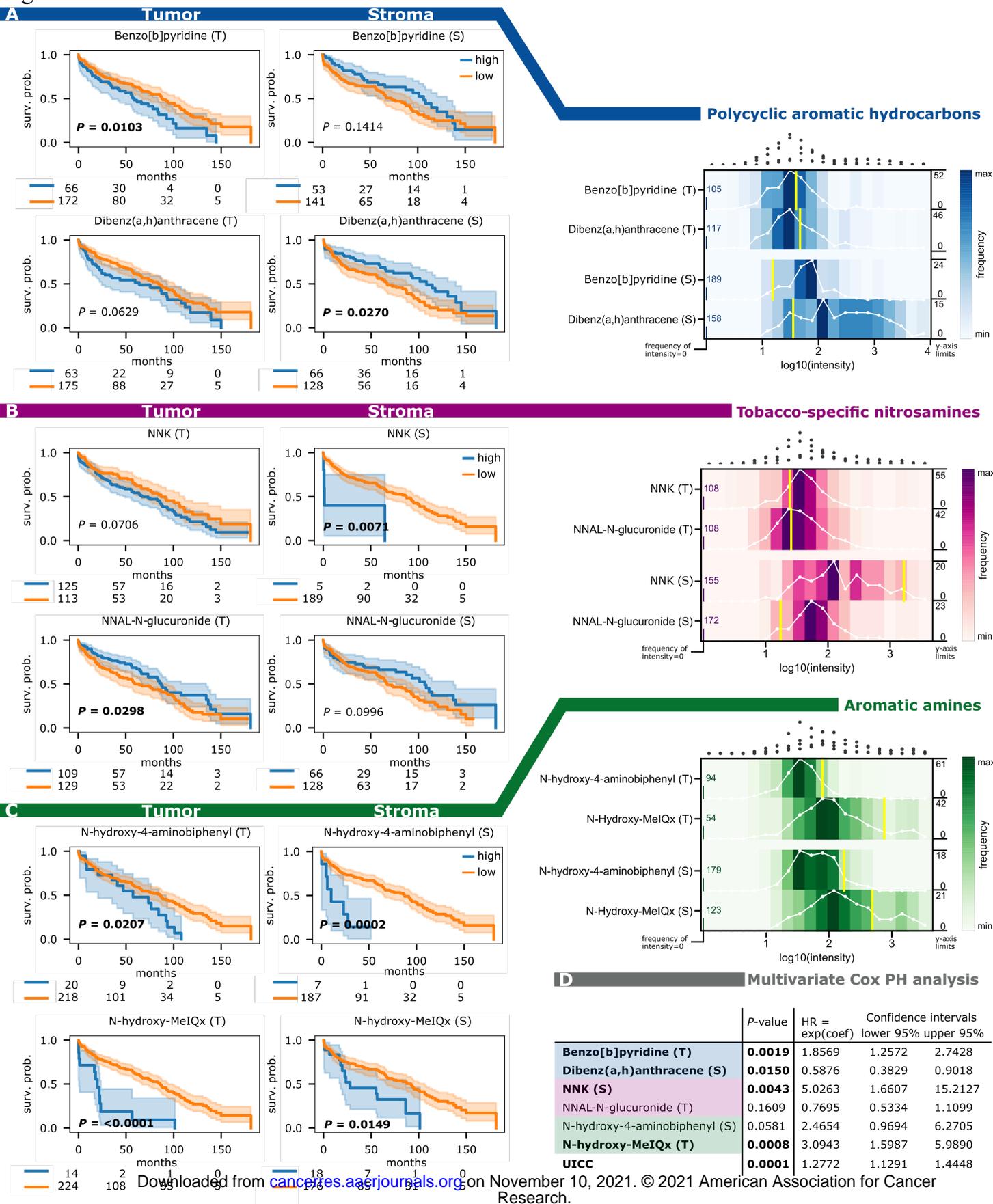
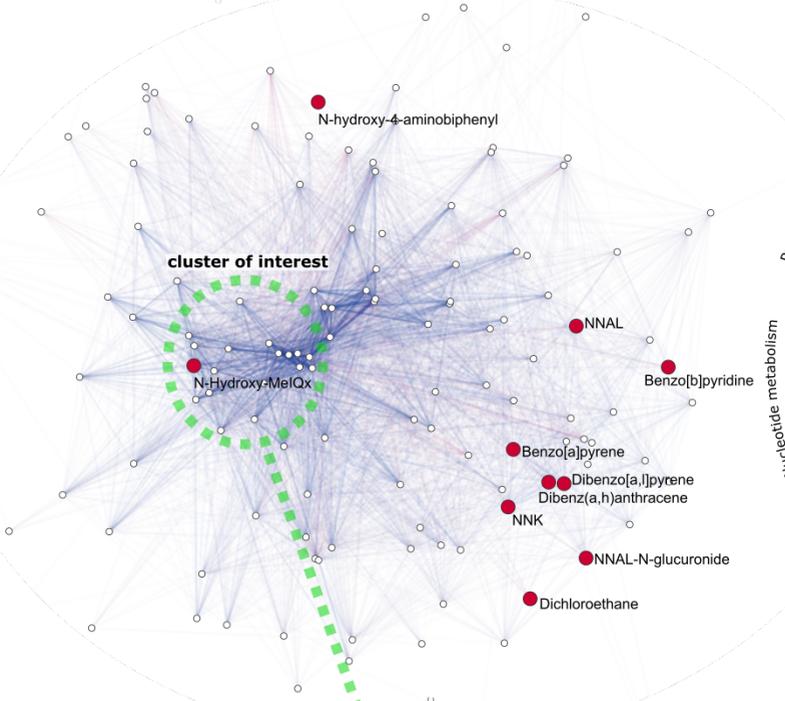
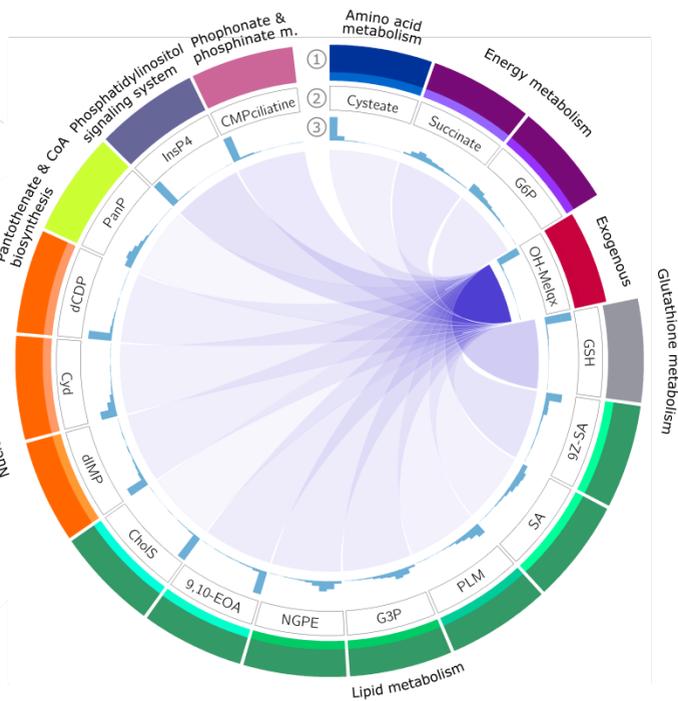


Figure 6

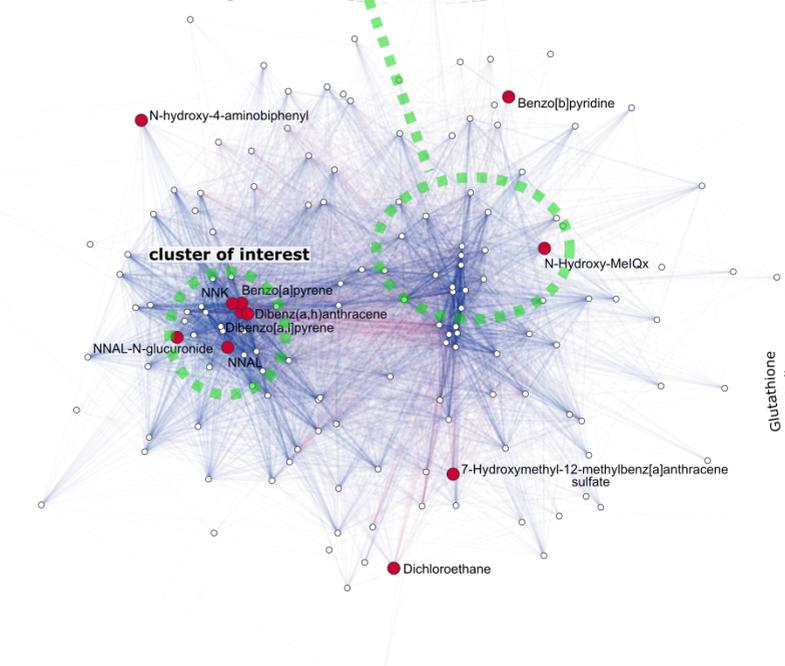
A Tumor - Spatial correlation network



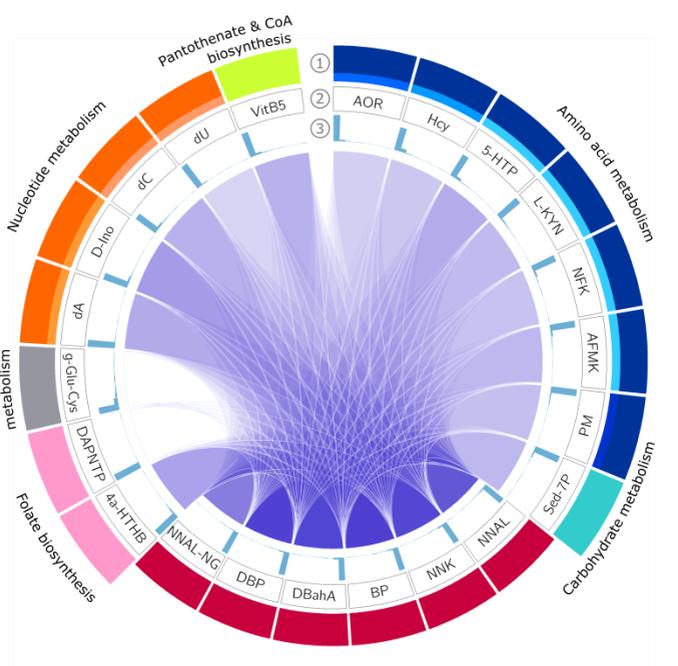
Cluster of interest



B Stroma - Spatial correlation network



Cluster of interest



Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Patterns of carbon-bound exogenous compounds in lung cancer patients and association with disease pathophysiology

Thomas Kunzke, Verena M. Prade, Achim Buck, et al.

Cancer Res Published OnlineFirst October 19, 2021.

Updated version	Access the most recent version of this article at: doi: 10.1158/0008-5472.CAN-21-1175
Supplementary Material	Access the most recent supplementary material at: http://cancerres.aacrjournals.org/content/suppl/2021/10/20/0008-5472.CAN-21-1175.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/early/2021/10/19/0008-5472.CAN-21-1175 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.