Title

Untargeted adductomics of Cys34 modifications to human serum albumin in newborn dried blood spots

Journal

Analytical and Bioanalytical Chemistry, 411(11)

ISSN

1618-2642

Authors

Yano, Yukiko
Grigoryan, Hasmik
Schiffman, Courtney
et al.

Publication Date

2019-04-01

DOI

10.1007/s00216-019-01675-8

Peer reviewed

HHS Public Access

Author manuscript

Anal Bioanal Chem. Author manuscript; available in PMC 2020 April 01.

Published in final edited form as:

Anal Bioanal Chem. 2019 April ; 411(11): 2351–2362. doi:10.1007/s00216-019-01675-8.

Untargeted adductomics of Cys34 modifications to human serum albumin in newborn dried blood spots

Yukiko Yano[1], **Hasmik Grigoryan**[1], **Courtney Schiffman**[2], **William Edmands**[1], **Lauren Petrick**[3], **Katie Hall**[1], **Todd Whitehead**[4,5], **Catherine Metayer**[4,5], **Sandrine Dudoit**[2], and **Stephen Rappaport**[1,5,*]

[1]Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94720, USA

[2]Division of Biostatistics, School of Public Health, University of California, Berkeley, CA 94720, USA

[3]The Senator Frank R. Lautenberg Environmental Health Sciences Laboratory, Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

[4]Division of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720, USA

[5]Center for Integrative Research on Childhood Leukemia and the Environment, University of California, Berkeley, CA 94720, USA

Abstract

Metabolism of chemicals from the diet, exposures to xenobiotics, the microbiome, and lifestyle factors (e.g., smoking, alcohol intake) produce electrophiles that react with nucleophilic sites in circulating proteins, notably Cys34 of human serum albumin (HSA). To discover potential risk factors resulting from *in utero* exposures, we are investigating HSA-Cys34 adducts in archived newborn dried blood spots (DBS) that reflect systemic exposures during the last month of gestation. The workflow includes extraction of proteins from DBS, measurement of hemoglobin (Hb) to normalize for blood volume, addition of methanol to enrich HSA by precipitation of Hb and other interfering proteins, digestion with trypsin, and detection of HSA-Cys34 adducts via nanoflow liquid chromatography-high resolution mass spectrometry. As proof-of-principle, we applied the method to 49 archived DBS collected from newborns whose mothers either actively smoked during pregnancy or were nonsmokers. Twenty-six HSA-Cys34 adducts were detected, including Cys34 oxidation products, mixed disulfides with low-molecular-weight thiols (*e.g.*,

*Corresponding author: srappaport@berkeley.edu.

Conflict of interest

The authors declare they have no conflict of interest.

Ethics approval

The CCLS and Center for Integrative Research on Childhood Leukemia and the Environment were approved by the University of California Committee for the Protection of Human Subjects, the California Health and Human Services Agency Committee for the Protection of Human Subjects, and the institutional review boards of all participating hospitals, as appropriate.

Informed consent

Written informed consent was obtained from all adult volunteer subjects and parents of all participating subjects in the CCLS.

cysteine, homocysteine, glutathione, cysteinylglycine, etc.), and other modifications. Data were normalized with a novel method ('scone') to remove unwanted technical variation arising from: HSA digestion, blood volume, DBS age, mass spectrometry analysis, and batch effects. Using an ensemble of linear and nonlinear models, the Cys34 adduct of cyanide was found to consistently discriminate between newborns of smoking and nonsmoking mothers with a mean fold change (smoking/nonsmoking) of 1.31. These results indicate that DBS-adductomics is suitable for investigating *in utero* exposures to reactive chemicals and metabolites that may influence disease risks later in life.

Keywords

adductomics; human serum albumin; dried blood spots; *in utero* exposures

Introduction

Many carcinogens are reactive electrophiles that are generated through metabolism of chemicals from the diet, environmental exposures, the microbiome, and lifestyle factors such as smoking and alcohol consumption. Although these reactive intermediates are short-lived *in vivo*, they can be quantified by measuring their reaction products (adducts) with circulating proteins, such as hemoglobin (Hb) and human serum albumin (HSA).[1, 2] We have focused on HSA adducts bound to the highly nucleophilic sulfhydryl group at Cys34, which is a powerful antioxidant and scavenger of reactive electrophiles in the interstitial space.[3] Whereas targeted assays are limited to measurement of particular HSA-Cys34 adducts known *a priori*, our adductomics methodology motivates discovery and quantitation of unknown HSA modifications of potential health significance.[4]

In our Cys34 adductomics pipeline, HSA is isolated from plasma/serum, digested with trypsin, and analyzed via nanoflow liquid chromatography-high resolution mass spectrometry (nLC-HRMS) to find and quantitate modifications at the third largest tryptic peptide (T3) (^{21}ALVLIAFAQYLQQC^{34}PFEDHVK^{41}).[5] In four previous studies, we applied this adductomics pipeline to plasma/serum from healthy smokers and nonsmokers in the U.S.,[5] nonsmoking women in China exposed to indoor combustion products and local controls,[6] nonsmoking British patients with lung & heart disease and local controls,[7] and nonsmoking Chinese workers exposed to benzene and local controls.[8] Here, we extended the adductomics assay to measure Cys34 adducts in newborn dried blood spots (DBS).

Because newborn DBS have been routinely collected at birth to screen for inborn errors of metabolism in the U.S. and worldwide,[9] analysis of archived newborn DBS provides an avenue for investigating the etiologies of diseases initiated *in utero*. Retrospective investigations of chromosomal translocations in DNA from newborn DBS provide direct evidence of the prenatal origin of childhood leukemia, the most common childhood cancer.[10–12] Chronic diseases in adult life, such as type 2 diabetes mellitus, cardiovascular disease, and the metabolic syndrome, can also have fetal origins.[13] Since HSA has a residence time of 28 days,[4] measuring Cys34 adducts in newborn DBS would allow us to

investigate exposures to reactive and potentially carcinogenic electrophiles during the last month of gestation.

Here, we describe untargeted measurements of HSA-Cys34 adducts in newborn DBS. A major challenge involved interfering species from the sample matrix, particularly lysed red blood cells and Hb that are not abundant in serum or plasma.[14, 15] Indeed, Hb is present at a 7-fold higher concentration than HSA in whole blood[16] and interferes with tryptic digestion that releases the T3 peptide and its modifications for analysis.[17] Using controlled addition of methanol, we effectively removed Hb from DBS extracts prior to digestion and detected HSA-Cys34 adducts via nLC-HRMS in archived DBS from 49 newborns with mothers who either actively smoked during pregnancy or were nonsmokers. After using a novel method to normalize the data, we applied an ensemble of statistical methods to select adducts that discriminated between newborns of smoking and nonsmoking mothers.

Materials and methods

Chemicals and reagents

Acetonitrile (Ultra Chromasolv, LCMS grade), triethylammonium bicarbonate (TEAB) buffer (1 M), ethylenediamine-tetraacetic acid (EDTA, anhydrous), HSA (lyophilized powder, 97–99%), and porcine trypsin were from Sigma-Aldrich (St. Louis, MO). Methanol (Optima, LCMS grade), formic acid (Optima, LCMS grade), and iodoacetamide (IAA) were from Fisher Scientific (Pittsburgh, PA). Purified human Hb was purchased from MP Biomedicals, LLC (Santa Ana, CA). Isotopically labeled T3 (iT3) with sequence AL-[^{15}N, ^{13}C-Val]-LIAFAQYLQQCPFEDH-[^{15}N,^{13}C-Val]-K was custom-made (>95%, BioMer Technology, Pleasanton, CA), and the carbamidomethylated iT3 (IAA-iT3)[18] was used as an internal standard to monitor retention time and mass drifts. Water was prepared with a PureLab purification system (18.2 mΩ cm resistivity at 25 °C; Elga LabWater, Woodridge, IL).

Preparation of capillary DBS for method development

For method development, capillary DBS were collected with informed consent from adult volunteers by finger prick with a sterile safety lancet (Fisher HealthCare, Houston, TX). The first drop of blood was discarded and subsequent drops were collected on Whatman 903 Protein Saver cards (GE Healthcare, Cardiff, UK). Blood spots were air dried for a minimum of 4 days and stored at −20 °C in glassine envelopes (GE Healthcare, Cardiff, UK) prior to use. Punches of 5 and 6-mm diameter were obtained from these DBS with a Biopunch (Ted Pella Inc., Redding, CA).

Archived newborn DBS

Newborn DBS were obtained with informed consent for 49 healthy control children from the California Childhood Leukemia Study (CCLS).[19] These newborn DBS had been archived by the California Department of Public Health[20] at −20 °C for 14 to 32 y prior to analysis in the current investigation. Twenty-three of these participants had mothers who actively smoked during pregnancy and the remaining 26 had nonsmoking mothers. Interviews with

the biological mother were conducted to collect data on the child's sex, race, and mother's smoking status during pregnancy. A total of 23 smoking/nonsmoking pairs were matched on sex and child's birth year. Smoking/nonsmoking pairs of newborn samples were analyzed together to minimize technical variation. Our methodology was developed for 4.7-mm punches from DBS. Because the archived newborn DBS for the present investigation were remnants from previous analyses, they consisted of areas of filter media equivalent to 4.7-mm punches based on size and weight.

Extraction of proteins and measurements of Hb and total protein

DBS punches were placed in microcentrifuge tubes and extracted with 55 μL of water at room temperature for 15 min with constant agitation at 1400 rpm (Thermomixer, Eppendorf, Hamburg, Germany). Samples were then centrifuged for 10 s and 5-μL aliquots were diluted with 45 μL of water to measure Hb concentrations (for normalization of blood volumes) with a Cytation 5 microplate spectrophotometer (BioTek Instruments, Winooski, VT) at room temperature. The absorbance of duplicate 2.5-μL sample aliquots was measured at 407 nm, which was the experimentally-determined absorbance maximum corresponding to heme in the *ex vivo* oxidation state of Hb.[21, 22] Absorbance readings were converted into Hb concentrations with five-point linear calibration curves using Hb standard solutions ranging from 0.5 to 5.0 mg/mL.

Absorbance measurements at 280 nm were used to calculate total protein concentrations in DBS extracts with correction for nucleic acid interferences at 260 nm and background correction at 320 nm.[23]

Sample preparation for adductomics

Various extraction protocols were tested and the method described below was found to be optimal for isolating HSA from DBS while removing Hb and other proteins from the extract (Results and Discussion provides further details). After Hb measurement, 41 μL of methanol was added to the remaining 50 μL of DBS extract (resulting in 45% methanol), vortexed, and mixed at 37 °C for 30 min with agitation at 1400 rpm (Thermomixer, Eppendorf, Hamburg, Germany). Samples were then stored at 4 °C for 30 min and centrifuged at 14,000×g for 10 min to remove precipitates and cellulose fibers. A 55-μL aliquot of the supernatant was diluted with 95 μL of digestion buffer (50 mM TEAB, 1 mM EDTA, pH 8.0), and the solution was loaded into a Costar Spin-X centrifuge tube filter (0.22 μm cellulose acetate, Corning Incorporated, Corning, NY) and centrifuged at 10,000×g for 10 min. A 130-μL aliquot of the filtered solution (containing ~17% methanol to enhance trypsin activity)[5] was transferred into BaroFlex 8-well strips (Pressure Biosciences Inc., South Easton, MA) to which 1 μL of 10-μg/μL trypsin was added (~1:10 enzyme-to-protein ratio). Pressure-assisted proteolytic digestion was performed with a Barozyme HT48 (Pressure Biosciences Inc., South Easton, MA) instrument, which cycled between ambient pressure (30 s) and 1,380 bar (20 kpsi, 90 s) for 32 min. After digestion, 3 μL of 10% formic acid was added to denature trypsin. Digests were transferred to new tubes and centrifuged for 2 min at 10,000×g. A 100-μL aliquot of the digest was transferred to a 300-μL silanized glass vial (Micosolv Technology Corporation, Leland, NC), and 1 μL of the isotopically labeled internal standard (IAA-iT3, 20 pmol/μL) was added. Samples were stored in liquid nitrogen

prior to nLC-HRMS. The 49 newborn DBS were processed daily in four batches of 12 or 13 samples.

nLC-HRMS analysis

Digests were analyzed by an Orbitrap Elite HRMS coupled to a Dionex Ultimate 3000 nanoflow LC system via a Flex Ion nanoelectrospray ionization source (Thermo Fisher Scientific, Waltham, MA), as described previously.[5] A Dionex PepSwitft monolithic column (100 μm i.d. × 25 cm) (Thermo Scientific, Sunnyvale, CA) was used with mobile phases consisting of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B), with gradient elution (2–35 % B, 26 min) at a flow rate of 750 nL/min. Two 1- μL aliquots were injected for each sample. Full scan mass spectra were acquired in the positive ion mode with a resolution of 120,000 at m/z 400 in the m/z = 350 – 1500 mass range using the Orbitrap. The MS was operated in data-dependent mode to collect tandem MS (MS2) spectra in the linear ion trap. Additional details of nLC-HRMS analysis are available in the Electronic Supplementary Material (ESM).

Identification, quantitation, and annotation of putative adducts

HSA adducts were pinpointed using the adductomics pipeline, as described previously.[5] Adducts were grouped by monoisotopic mass (MIM) within 10 ppm and retention time (RT) within 1.5 min. The means of MIMs and RTs were calculated for each adduct across all samples. Masses added to the thiolate ion of the triply charged T3 peptide (Cys34-S^-) were calculated and annotated as described previously[5] (see ESM for details). Mass accuracies were estimated by the difference between theoretical and observed MIMs.

Peak picking and integration were performed using the Xcalibur Processing Method (version 3.0, Thermo Fisher Scientific, Waltham, MA) based on the average MIMs (5 ppm mass accuracy) and RTs of the putative adducts. Peaks were integrated with the Genesis algorithm after normalizing the RTs using the internal standard (iT3-IAA) and using a RT window of 60 s.

Statistical analysis

All statistical analyses were performed using the R statistical programming environment.[24] Peak areas were log-transformed and the means of duplicate measurements were used after removal of four obvious outliers. Missing values were imputed using k-nearest-neighbor imputation (see ESM for details).[25]

Data normalization was optimized using the Bioconductor R package ‘scone,’[26] which explored different scaling methods and used a regression-based strategy to adjust for unwanted factors of variation.[27] Normalization workflows were ranked by a variety of metrics to reduce unwanted technical variation while preserving the biological variation of interest.[28] The top-ranking normalization scheme according to ‘scone’ used DESeq scaling and removed unwanted variation due to the following factors: digested HSA, blood volume, DBS age, instrument performance, and batch effects. Here, ‘digested HSA’ was quantified by the abundance of the tryptic housekeeping peptide adjacent to T3 with sequence ^{42}LVNEVTEFAK^{51} (doubly charged precursor ion at m/z = 575.3111).[5] ‘Blood

volume' was indicated by measurement of Hb in DBS extracts. 'DBS age' (*i.e.*, 2017 – child birth year) was used to adjust for differences in the extraction efficiency due to the age of the DBS.[29] 'Instrument performance' was indicated by the drift in the abundance of the internal standard over time. 'Batch effect' was used to adjust for differences in the four subsets of samples that were prepared on different days.

To ensure robust associations with mothers' smoking status, a combination of linear and nonlinear models were used to identify discriminating adducts. First, the following multivariable regression model was applied:

$$Y_i = \beta_0 + \beta_1 X_{Smoke} + \beta_2 X_{Sex} + \beta_3 X_{Race} + \beta_4 X_{Batch} + \beta_5 X_{HK} + \beta_6 X_{IS} + \beta_3 X_{Hb} + \beta_3 X_{DBS\ Age} + \varepsilon_i, \quad (1)$$

where Y_i is a vector of logged, DESeq scaled abundances of the ith adduct, X_{Sex}, and X_{Race} (0 = other, 1 = white) are binary vectors, X_{Batch} is a four-level categorical variable indicating batch, X_{HK} is the vector of housekeeping peptide abundances, X_{IS} is the vector of internal standard abundances, X_{Hb} is the vector of Hb measurements, and $X_{DBS\ Age}$ is a vector of DBS sample ages. The coefficient β_1 and estimated p-values were used to rank adducts by their associations with mothers' smoking status. Significance levels were adjusted for multiple testing with the Benjamini-Hochberg method to control the false discovery rate (FDR) at $\alpha = 0.05$.[30] The mean fold change (smoking/nonsmoking) in adduct intensities between newborns of smoking and nonsmoking mothers was calculated as $\exp(\beta_1)$.

Next, a logistic least absolute shrinkage and selection operator (lasso)[31] model was fitted to the logged, scaled, and normalized adduct abundances from 'scone' along with the matching variables (*i.e.*, sex, birth year), to select a subset of adducts that best predicted the mothers' smoking status. To increase stability, the logistic lasso regression was performed on 500 bootstrapped data sets.[32] The percentage of times each adduct was selected by the lasso model out of the 500 iterations was used as a measure of variable importance. This process was performed for a range of values of the lasso penalty parameter (lambda range: 0.12 ~ 0.20) to ensure that the final variable selections were robust.

Finally, adducts were also ranked by their associations with the mothers' smoking status using random forest variable importance.[33] A random forest with 500 trees was used to predict mothers' smoking status with the logged, scaled, and normalized adduct abundances in addition to matching variables. Adducts were ranked by the mean decrease in Gini index, which indicates the total decrease in node impurity (as measured by the Gini index when splitting on the adducts within the decision tree averaged over all trees).[34]

To evaluate the importance of selected adducts that discriminated between the smoking and nonsmoking mothers based on all three statistical approaches, we obtained 10-fold cross-validated area under the estimated receiver operating characteristic (ROC) curve (AUC) by fitting a logistic regression model on smoking status with discriminating adducts as predictors. Cross-validated AUC estimates were obtained with the 'cvAUC' R package.[35]

Results and discussion

Measurement of Hb in archived DBS

Our analysis was performed exclusively with newborn DBS that had been maintained in freezers at −20 °C for 14 to 32 y prior to analysis. Quantitative analysis of Hb in DBS stored at room temperature can be problematic because of oxidation of Hb.[36] Indeed, when DBS were stored at room temperature for several months compared to storage at −20 °C, we observed that the color changed from deep red to dark brown indicating oxidation of Hb,[37] and decreased water solubility of Hb (data not shown). The absorption spectra in the 250 – 750 nm range for the Hb calibration curve (ESM Fig. S1A) and Hb measured from extracts of 10 randomly selected 4.7-mm punches from archived newborn DBS (ESM Fig. S1B) both showed maximum absorbance at 407 nm. The Hb calibration curves measured for each of the four batches of newborn DBS showed a strong linear relationship between Hb concentrations and absorbance measurements at 407 nm ($r^2 > 0.99$, ESM Fig. S1C). Details on oxidation of Hb is provided in Supporting Information.

Adductomics analysis of DBS

In preparing DBS for adductomics, HSA must first be extracted from the filter paper and isolated from whole blood. Previous analyses of proteins extracted from DBS have used mixed aqueous-organic solutions to selectively precipitate proteins in solution.[38, 39] Since Hb is one of the most prominent interfering proteins in whole blood, we tested various mixtures of organic solvents (ethanol, methanol, acetonitrile, and 1-propanol) to precipitate Hb while retaining HSA in solution (data not shown), and found ethanol and methanol to be most effective. When the concentration of ethanol and methanol were gradually increased from 30 to 60% (v/v), HSA remained in solution at concentrations less than 40% ethanol or 45% methanol and increasingly precipitated at concentrations up to 60% for both ethanol and methanol (ESM Fig. S2). Methanol was more effective at precipitating Hb with a ~95% decrease in concentration at 45% methanol when compared to DBS extracted with water (ESM Fig. S2).

The recovery of HSA was also influenced by the total protein concentration of the DBS extract, where higher total protein concentrations led to a lower recovery of HSA after precipitation (ESM Fig. S3). Based on preliminary analysis of ten 4.7-mm punches from archived newborn DBS, we found that the average total protein concentration for newborns (4.98 mg/mL) was approximately equivalent to a 6-mm punch from an adult DBS. The observed higher total protein concentration of newborn blood reflects the higher Hb concentrations in newborns compared to adults.[40] Therefore, we used 6-mm punches from adult DBS to find the optimal concentration of methanol in the extraction mixture to isolate HSA. In addition, the isoelectric points (pI) of fetal and adult Hb differ (fetal Hb: pI 6.98, adult Hb: pI 6.87),[41] and fetal Hb precipitates more readily at neutral pH. When comparing 40, 45, 48, and 50% methanol, we observed that Hb did not precipitate with 40% methanol and that there was a ~40% loss of HSA when the methanol concentration was increased to 50% (ESM Fig. S4). Based on this result, we chose 45% methanol to isolate HSA in the DBS extract. We also found that incubating the samples at 37 °C (as opposed to room temperature) after addition of methanol to the aqueous DBS extract was essential for

denaturing and precipitating Hb. When we tested extraction with 45% methanol on four 4.7-mm punches from archived newborn DBS, there was no loss of HSA and the residual Hb concentration was ~0.02 mg/mL (1.2% of the initial concentration) (ESM Fig. S5).

Digestion of HSA was optimized by testing various digestion programs using the pressure-cycling technology and by adjusting the proteolytic enzyme-to-protein ratio (E:P). While conventional proteomics approaches perform reduction and alkylation of proteins prior to digestion,[42] we did not apply these techniques in order to preserve Cys34 disulfides and to prevent the formation of artifacts. The digestion time was tested at 16, 32 and 64 min to determine the optimal time needed for a high yield of digestion. Both the total ion chromatogram and base peak chromatogram were examined for the presence of undigested proteins and yield of peptides.[17] While chromatograms from 32 and 64 min digests were comparable, there were fewer peptides and a more prominent peak for undigested proteins with the 16 min digestion, suggesting that 32 min was sufficient (ESM Fig. S6). Undigested protein was observed despite longer digestions, and probably reflects the lack of denaturation and reduction of disulfide bonds. It may also be due to the presence of residual non-HSA proteins, including Hb, which increase competition for trypsin cleavage sites and thereby interfere with the digestion of HSA.[42] We tested various 30 min digestion programs consisting of shorter and longer cycles at high pressure, but there was little difference in the resulting chromatograms (data not shown). The E:P was optimized to ensure an amount of trypsin that was sufficient for digestion while preventing autolysis.[43] When the E:P was increased from 1:18 to 1:3, trypsin activity showed a plateau at about 1:10, after which a further increase in trypsin did not improve the digestion (ESM Fig. S7). Increasing trypsin to a ratio of 1:5 resulted in incomplete digestion and more trypsin autolysis products, which could lead to ion suppression during MS detection.[44]

Analysis of archived newborn DBS

Normalization of blood volumes based on Hb measurements proved to be effective in removing unwanted variation in adduct abundances as indicated by ‘scone.’ In addition, there was a relatively strong negative correlation (Pearson’s $r = -0.56$) between Hb concentrations and the second estimated factor of unwanted variation (ESM Fig. S8) using the RUVg method of Risso et al., a factor analysis method that estimates factors of unwanted variation via decomposition of a subset of the data consisting of negative control adducts.[27] This shows that Hb is indeed a factor of unwanted variation in DBS analysis. Furthermore, the weights of the DBS and Hb concentrations were highly correlated (Pearson’s $r = 0.93$), suggesting that Hb is a good predictor of blood volume in newborn DBS.

Relative log abundance (RLA) plots,[45] which were obtained by standardizing each adduct by the median abundance across samples and logging the resulting ratio, were used to visually inspect the reproducibility of replicate measurements and to assess the normalization scheme. Two subjects (of nonsmoking mothers) were removed from the analysis and only one measurement was used for one subject due to high variation in adduct abundances based on the RLA plot of duplicate measurements (ESM Fig. S9). Three adducts

were removed from the analysis since they were missing in over half of the samples (ESM Fig. S10). This left 47 subjects and 26 putative Cys34 adducts for analysis.

Newborns of smoking and nonsmoking mothers, matched by sex and birth year, were similar with respect to race, DBS weight, and Hb concentrations (ESM Table S1). The distribution of adduct peak areas in each sample before and after normalization for unwanted factors of variation (*i.e.*, Hb concentration, DBS age, housekeeping peptide, internal standard, and batch effects) is shown in Fig. 1. By comparing the RLA plots from before (Fig. 1A) and after (Fig. 1B) normalization, it can be seen that this scheme effectively removed unwanted variation.

Table 1 summarizes the 26 adducts that were detected in the DBS of newborns from smoking and nonsmoking mothers. Representative MS2 spectra of the adducts are shown in ESM Fig. S11. Nineteen of the 26 adducts have been previously reported, including truncations, unmodified T3, methylated T3, Cys34 sulfoxidation products (*e.g.*, sulfinic and sulfonic acids), a cyanide adduct, and Cys34 disulfides of low-molecular-weight thiols.[5, 6, 8] Only three of the remaining 7 adducts had putative annotations, *i.e.*, the Cys34 sulfenamide (811.09), a CH_2 crosslink (815.76), and the sodium adduct (819.09). Aside from the unmodified T3 peptide (811.76), the Cys34 sulfinic acid (822.42) and the *S*-glutathione (GSH) disulfide (913.45) were the most abundant adducts across all samples (ESM Fig. S12).

In studying reaction pathways leading to Cys34 sulfoxidation products, Grigoryan et al. reported an intramolecular cyclic sulfinamide adduct (816.42) with the added mass (+O, $-H_2$), which results from the formation of a cross-link between Cys34 and the amide group of the adjacent Gln33.[18] Two different pathways were proposed for the formation of the sulfinamide adduct: (1) from dehydration of Cys34 sulfenic acid (Cys34-SOH) resulting in the sulfenamide adduct (mass difference [$-H_2$]) with the Cys34-Gln33 cross-link, which is then oxidized to the sulfinamide adduct; (2) from oxidation of the sulfenic acid to the sulfinic acid (822.42, Cys34-SO_2H), from which loss of water results in the sulfinamide adduct (Fig. 2). The second reaction pathway appeared to be more likely because the intermediate sulfenamide adduct had not been detected in plasma/serum samples.[5, 6, 18] However, in newborn DBS we detected both the sulfenamide (811.09) and sulfinamide (816.42) adducts, suggesting that formation of the sulfinamide adduct is possible via both pathways. In addition, we detected the sulfonamide adduct (821.75, added mass [$+O_2$, $-H_2$]), which results from oxidation of the sulfinamide adduct (Fig. 2).[46] The formation of these intramolecular cyclic adducts may have been promoted by the drying of DBS with subsequent dehydration of sulfenic, sulfinic, and sulfonic acids to produce the sulfenamide, sulfinamide, and sulfonamide adducts, respectively (Fig. 2). It is also possible that these intramolecular cyclic adducts (particularly sulfenamide) were detected in the present analysis due to an increased stability of these analytes in DBS compared to plasma and serum. Analytes in DBS are typically less reactive than in liquid blood because they are stabilized through adsorption onto a solid cellulose matrix (*i.e.*, the filter paper).[47] Proteins commonly degrade in aqueous solutions due to aggregation, oxidation, and hydrolysis[48] that appear to be minimized during long-term storage of DBS in a freezer. In fact, we did not observe the T3 dimer in our analysis of newborn DBS although this dimer is routinely

detected in plasma/serum samples.[5–8] Furthermore, adducts in DBS may be less susceptible to formation of artifacts because HSA is immobilized by the filter paper and less likely to interact with other molecules during storage.

Adducts that discriminated newborns of smoking and nonsmoking mothers

We had anticipated that Cys34 adducts of two toxic contaminants of cigarette smoke, *i.e.*, ethylene oxide and acrylonitrile, might be more abundant in newborns of smoking mothers given our previous detection of these adducts in plasma from adult smokers and their absence in plasma from nonsmokers.[5] However, these adducts were not detected in the newborn DBS, possibly due to low concentrations of the precursor molecules in newborn blood. Mothers may have stopped smoking during the third trimester or before the last month of pregnancy, and this may explain why we did not see all of the expected adducts in the present analysis. It would be interesting to determine whether Hb adducts of ethylene oxide and acrylonitrile, which have been reported in smokers' blood,[49, 50] could be detected in newborn DBS from smoking mothers.

We used three different statistical approaches to select adducts that discriminated newborns of smoking and nonsmoking mothers. The volcano plot (Fig. 3A) shows the relationship between the smoker/nonsmoker fold change of a given adduct and the statistical significance of the difference in adduct abundance obtained from Equation 1. Discriminating adducts were also identified using lasso on bootstrapped data, which ranked adducts by the percentage of times each adduct was introduced into the model out of 500 iterations (Fig. 3B) and by random forest (Fig. 3C).

Of the 26 adducts that were tested, the Cys34 adduct of cyanide (820.09) was ranked highest by all three statistical methods. As seen from the volcano plot (Fig. 3A), the cyanide adduct showed a marked difference between newborns of smoking and nonsmoking mothers, with a smoker/nonsmoker mean fold change of 1.31 (nominal p-value = 0.0017; FDR-adjusted p-value = 0.044). The cyanide adduct was also top-ranked by both the lasso model (Fig. 3B) and random forest (Fig. 3C). The elevated levels of the cyanide adduct among newborns of smoking mothers are consistent with inhalation of hydrogen cyanide from tobacco smoke.[51]

Adduct 830.43 was ranked second, based on its fold change of 0.82 (nominal p-value = 0.049, Fig. 3A) and by lasso (Fig. 3B), and was ranked fourth by random forest (Fig. 3C). We suspect that adduct 830.43 represents *S*-addition of acrolein to Cys34 given an added mass that corresponds to $+C_3H_5O$[52] within 0.7 ppm (Table 1). This accurate mass and the corresponding MS2 spectrum (ESM Fig. S11) of 830.43 match those attributed by Colombo et al. to a T3 adduct of acrolein that had been derived from HSA after incubation with a cigarette smoke extract.[53] Although acrolein is a prominent constituent of cigarette smoke,[54] in our samples abundances of 840.43 were *lower* in newborns of smoking mothers. Since acrolein is also a normal product of lipid peroxidation,[52] our results cannot elucidate the potential utility of adduct 830.43 as a potential biomarker of acrolein in neonatal DBS.

Because the cyanide adduct was unambiguously more abundant in newborns of smoking mothers, we performed a ROC analysis using logistic regression with the cyanide adduct as

the predictor and estimated a cross-validated AUC of 0.79 (95% CI: 0.65, 0.93). Although this AUC estimate is likely to be optimistic, because we did not have an independent test set for the ROC analysis, this indicates good discrimination between newborns of smoking and nonsmoking mothers.

The half-life of cyanide in blood is less than one hour, which makes it difficult to obtain accurate measurements of cyanide exposure from the direct analysis of blood from smokers.[55] While more stable metabolites of cyanide, such as thiocyanate, are often used as surrogate measures of cyanide exposure, pairwise correlations are small between such metabolites and cyanide exposures.[55, 56] Since the residence time of HSA is ~1 month,[4] the cyanide adduct of Cys34 arguably represents a more accurate measure of chronic low-level exposure to cyanide.[57]

Using adductomics to discover biomarkers of *in utero* exposures

Adducts of HSA represent biomarkers of *in utero* exposures during the last month of gestation. A good example of such exposures is maternal smoking during pregnancy, which has been consistently associated with increased risks of adverse birth outcomes (*e.g.*, low birth weight, preterm birth)[58] and has also been suggested to increase the risk of diseases later in life, including various types of cancer.[59–62] However, the long term effects of *in utero* tobacco-smoke exposures on the risk of childhood cancer have been inconsistent, with many studies reporting null associations.[63] One limitation of these epidemiological investigations has been reliance on maternal self-reports to retrospectively characterize fetal exposures to tobacco smoke.[63, 64] Exposure misclassification due to recall and reporting bias is a particular concern among pregnant women, who may feel uncomfortable discussing their smoking histories during pregnancy, and can result in underestimation of fetal health effects from smoking mothers.[65]

Biomarkers complement interview-based exposure assessment by providing objective measures of exposure that are not susceptible to recall bias. Nicotine and its metabolite cotinine are commonly measured in biological fluids (*e.g.*, urine, blood, saliva) to assess tobacco smoke exposures.[65] For retrospective analyses of fetal exposures, archived newborn DBS are particularly attractive because they are readily available in California's repository that is maintained at −20 °C. In addition, newborn DBS enable direct measurement of fetal exposures that can accumulate in the placenta and exceed those of the mother.[66] Metabolites of nicotine, mainly cotinine, have been measured in newborn DBS to improve smoking surveillance among pregnant women.[67, 68] However, the half-life of cotinine is only about 28 h in infants, and cotinine may only be detected in newborns of heavy smokers who smoke throughout pregnancy.[69] Likewise, the half-lives of cyanide and its metabolites, such as thiocyanate, vary greatly in a given subject even within a single day, and are, therefore, imprecise measures of cyanide exposure in blood from smokers.[55, 56]

Since the residence time of HSA is 28 days,[4] stable Cys34 adducts detected in newborn DBS theoretically represent exposures received during the last month of gestation and should be only marginally affected by the day-to-day variability in exposure to the adducting species.[70] Using time-course experiments, Fasco et al. showed that the cyanide adduct of

Cys34 was stable after formation in human blood and encouraged its measurement for evaluating cyanide exposures in populations at risk.[57] In our study of newborn DBS, the Cys34 cyanide adduct effectively discriminated between mothers who reported smoking during pregnancy vs. those who did not, adding further validity to use of this adduct as a biomarker of prenatal exposures to toxicants from smoking.

It is possible that some HSA modifications may be artefacts of sample processing or storage. However, the purpose of our study was to identify adducts that discriminated between newborns of smoking and nonsmoking mothers. All blood spots from the two groups were sampled, stored, and treated identically, were matched by birth year and sex, and were randomized for analysis. This ensures that the differences we observed in the HSA modifications between the two groups were not artefacts.

Conclusion

By substantially modifying an untargeted HSA-Cys34 adductomics pipeline,[5] we were able to analyze Cys34 adducts extracted from newborn DBS. We then validated the methodology with archived DBS from newborns of smoking and nonsmoking mothers. Despite the small sample size and limited blood volume derived from 4.7-mm DBS punches, we were able to detect significantly higher levels of the Cys34 cyanide adduct among newborns whose mothers smoked during pregnancy. Since archived newborn DBS represent snapshots of *in utero* exposures, our adductomics method can be applied to investigate early-life exposures associated with childhood diseases. We are currently applying this methodology to archived newborn DBS from the CCLS to discover HSACys34 adducts associated with childhood leukemia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

The biospecimens used in this study were obtained from the California Biobank Program, (SIS request number 26), in accordance with Section 6555(b), 17 CCR. The NIEHS, USEPA, and California Department of Public Health are not responsible for the results or conclusions drawn by the authors of this publication.

Funding Sources

This work was supported by U.S. National Institutes of Health through grants P01ES018172, P50ES018172, R01ES009137 and P42ES0470518 from the National Institute for Environmental Health Sciences and grant R33CA191159 from the National Cancer Institute, by the U.S. Environmental Protection Agency through grants RD83451101 and RD83615901, and by a grant for a pilot project from Children with Cancer, a registered Charity in the U.K.

References

1. Rubino FM, Pitton M, Di Fabio D, Colombi A (2009) Toward an “omic” physiopathology of reactive chemicals: Thirty years of mass spectrometric study of the protein adducts with endogenous and xenobiotic compounds. Mass Spectrom Rev 28:725–784. doi: 10.1002/mas.20207 [PubMed: 19127566]

2. Törnqvist M, Fred C, Haglund J, Helleberg H, Paulsson B, Rydberg P (2002) Protein adducts: quantitative and qualitative aspects of their formation, analysis and applications. J Chromatogr B 778:279–308. doi: 10.1016/S1570-0232(02)00172-1

3. Aldini G, Vistoli G, Regazzoni L, Gamberoni L, Facino RM, Yamaguchi S, Uchida K, Carini M (2008) Albumin Is the Main Nucleophilic Target of Human Plasma: A Protective Role Against Pro-atherogenic Electrophilic Reactive Carbonyl Species? Chem Res Toxicol 21:824–835. doi: 10.1021/tx700349r [PubMed: 18324789]

4. Rappaport SM, Li H, Grigoryan H, Funk WE, Williams ER (2012) Adductomics: Characterizing exposures to reactive electrophiles. Toxicol Lett 213:83–90. doi: 10.1016/j.toxlet.2011.04.002 [PubMed: 21501670]

5. Grigoryan H, Edmands W, Lu SS, Yano Y, Regazzoni L, Iavarone AT, Williams ER, Rappaport SM (2016) Adductomics Pipeline for Untargeted Analysis of Modifications to Cys34 of Human Serum Albumin. Anal Chem 88:10504–10512. doi: 10.1021/acs.analchem.6b02553 [PubMed: 27684351]

6. Lu SS, Grigoryan H, Edmands WMB, Hu W, Iavarone AT, Hubbard A, Rothman N, Vermeulen R, Lan Q, Rappaport SM (2017) Profiling the Serum Albumin Cys34 Adductome of Solid Fuel Users in Xuanwei and Fuyuan, China. Environ Sci Technol 51:46–57. doi: 10.1021/acs.est.6b03955 [PubMed: 27936627]

7. Liu S, Grigoryan H, Edmands WMB, Dagnino S, Sinharay R, Cullinan P, Collins P, Chung KF, Barratt B, Kelly FJ, Vineis P, Rappaport SM (2018) Cys34 Adductomes Differ between Patients with Chronic Lung or Heart Disease and Healthy Controls in Central London. Environ Sci Technol 52:2307–2313. doi: 10.1021/acs.est.7b05554 [PubMed: 29350914]

8. Grigoryan H, Edmands WMB, Lan Q, Carlsson H, Vermeulen R, Zhang L, Yin S-N, Li G-L, Smith MT, Rothman N, Rappaport SM (2018) Adductomic signatures of benzene exposure provide insights into cancer induction. Carcinogenesis 39:661–668. doi: 10.1093/carcin/bgy042 [PubMed: 29538615]

9. Therrell BL, Padilla CD, Loeber JG, Kneisser I, Saadallah A, Borrajo GJC, Adams J (2015) Current status of newborn screening worldwide: 2015. Semin Perinatol 39:171–187. doi: 10.1053/j.semperi.2015.03.002 [PubMed: 25979780]

10. Gale KB, Ford AM, Repp R, Borkhardt A, Keller C, Eden OB, Greaves MF (1997) Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. Proc Natl Acad Sci 94:13950–13954. doi: 10.1073/pnas.94.25.13950 [PubMed: 9391133]

11. Greaves M (2005) In utero origins of childhood leukaemia. Early Hum Dev 81:123–129. doi: 10.1016/j.earlhumdev.2004.10.004 [PubMed: 15707724]

12. Wiemels JL, Walsh KM, de Smith AJ, Metayer C, Gonseth S, Hansen HM, Francis SS, Ojha J, Smirnov I, Barcellos L, Xiao X, Morimoto L, McKean-Cowdin R, Wang R, Yu H, Hoh J, DeWan AT, Ma X (2018) GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. Nat Commun 9:286. doi: 10.1038/s41467-017-02596-9 [PubMed: 29348612]

13. Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of In Utero and Early-Life Conditions on Adult Health and Disease. N Engl J Med 359:61–73. doi: 10.1056/NEJMra0708473 [PubMed: 18596274]

14. Chambers AG, Percy AJ, Hardie DB, Borchers CH (2013) Comparison of Proteins in Whole Blood and Dried Blood Spot Samples by LC/MS/MS. J Am Soc Mass Spectrom 24:1338–1345. doi: 10.1007/s13361-013-0678-x [PubMed: 23821375]

15. Martin NJ, Cooper HJ (2014) Challenges and opportunities in mass spectrometric analysis of proteins from dried blood spots. Expert Rev Proteomics 11:685–695. doi: 10.1586/14789450.2014.965158 [PubMed: 25308552]

16. Chambers AG, Percy AJ, Yang J, Borchers CH (2015) Multiple Reaction Monitoring Enables Precise Quantification of 97 Proteins in Dried Blood Spots. Mol Cell Proteomics 14:3094–3104. doi: 10.1074/mcp.O115.049957 [PubMed: 26342038]

17. Hustoft HK, Reubsaet L, Greibrokk T, Lundanes E, Malerod H (2011) Critical assessment of accelerating trypsination methods. J Pharm Biomed Anal 56:1069–1078. doi: 10.1016/j.jpba.2011.08.013 [PubMed: 21873015]

18. Grigoryan H, Li H, Iavarone AT, Williams ER, Rappaport SM (2012) Cys34 Adducts of Reactive Oxygen Species in Human Serum Albumin. Chem Res Toxicol 25:1633–1642. doi: 10.1021/tx300096a [PubMed: 22591159]
19. Metayer C, Zhang L, Wiemels JL, Bartley K, Schiffman J, Ma X, Aldrich MC, Chang JS, Selvin S, Fu CH, Ducore J, Smith MT, Buffler PA (2013) Tobacco Smoke Exposure and the Risk of Childhood Acute Lymphoblastic and Myeloid Leukemias by Cytogenetic Subtype. Cancer Epidemiol Biomarkers Prev 22:1600–1611. doi: 10.1158/1055-9965.EPI-13-0350 [PubMed: 23853208]
20. California Department of Public Health (CDPH) Newborn Screening Specimens Use and Storage. https://www.cdph.ca.gov/Programs/CFH/DGDS/Pages/nbs/NBSDBS-Storage.aspx. Accessed 11 May 2018
21. Waterman MR (1978) [48] Spectral characterization of human hemoglobin and its derivatives In: Methods Enzymol. pp 456–463
22. Meng F, Alayash AI (2017) Determination of extinction coefficients of human hemoglobin in various redox states. Anal Biochem 521:11–19. doi: 10.1016/j.ab.2017.01.002 [PubMed: 28069451]
23. BioTek Instruments (2012) Multi-Volume Based Protein Quantification.
24. R Development Core Team (2016) R: A Language and Environment for Statistical Computing In: R Found. Stat. Comput Vienna, Austria https://www.r-project.org/. Accessed 8 Mar 2017
25. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, Botstein D, Altman RB (2001) Missing value estimation methods for DNA microarrays. Bioinformatics 17:520–525. doi: 10.1093/bioinformatics/17.6.520 [PubMed: 11395428]
26. Cole M, Risso D (2018) scone: Single Cell Overview of Normalized Expression data. R Packag. version 1.2.0
27. Risso D, Ngai J, Speed TP, Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol 32:896–902. doi: 10.1038/nbt.2931 [PubMed: 25150836]
28. Cole MB, Risso D, Wagner A, DeTomaso D, Ngai J, Purdom E, Dudoit S, Yosef N (2017) Performance Assessment and Selection of Normalization Procedures for Single-Cell RNA-Seq. bioRxiv 1–30. doi: 10.1101/235382
29. Petrick L, Edmands W, Schiffman C, Grigoryan H, Perttula K, Yano Y, Dudoit S, Whitehead T, Metayer C, Rappaport S (2017) An untargeted metabolomics method for archived newborn dried blood spots in epidemiologic studies. Metabolomics 13:27. doi: 10.1007/s11306-016-1153-z [PubMed: 29706849]
30. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289–300.
31. Tibshirani R (1996) Regression Shrinkage and Selection via the Lasso. J R Stat Soc Ser B 58:267–288.
32. Bach FR (2008) Bolasso: Model Consistent Lasso Estimation through the Bootstrap In: Proc. 25th Int. Conf. Mach. Learn ACM Press, New York, New York, USA, pp 33–40
33. Liaw A, Wiener M (2002) Classification and Regression by randomForest. R News 2:18–22.
34. Calle ML, Urrea V (2011) Letter to the Editor: Stability of Random Forest importance measures. Brief Bioinform 12:86–89. doi: 10.1093/bib/bbq011 [PubMed: 20360022]
35. LeDell E, Petersen M, Van Der Laan M (2016) Cross-Validated Area Under the ROC Curve Confidence Intervals. R Packag. version 1.1.0
36. Bremmer RH, de Bruin DM, de Joode M, Buma WJ, van Leeuwen TG, Aalders MCG (2011) Biphasic Oxidation of Oxy-Hemoglobin in Bloodstains. PLoS One 6:e21845. doi: 10.1371/journal.pone.0021845 [PubMed: 21789186]
37. Bremmer RH, Nadort A, van Leeuwen TG, van Gemert MJC, Aalders MCG (2011) Age estimation of blood stains by hemoglobin derivative determination using reflectance spectroscopy. Forensic Sci Int 206:166–171. doi: 10.1016/j.forsciint.2010.07.034 [PubMed: 20729018]
38. Funk WE, Waidyanatha S, Chaing SH, Rappaport SM (2008) Hemoglobin Adducts of Benzene Oxide in Neonatal and Adult Dried Blood Spots. Cancer Epidemiol Biomarkers Prev 17:1896–1901. doi: 10.1158/1055-9965.EPI-08-0356 [PubMed: 18708378]

39. DeWilde A, Sadilkova K, Sadilek M, Vasta V, Hahn SH (2008) Tryptic Peptide Analysis of Ceruloplasmin in Dried Blood Spots Using Liquid Chromatography-Tandem Mass Spectrometry: Application to Newborn Screening. Clin Chem 54:1961–1968. doi: 10.1373/clinchem.2008.111989 [PubMed: 18845768]
40. Andropoulos DB (2011) Appendix B: Pediatric Normal Laboratory Values In: Greg. Pediatr. Anesth Wiley-Blackwell, Oxford, UK, pp 1300–1314
41. (1982) 3. Fetal hemoglobin: Structure and function. Scand J Clin Lab Invest 42:32–37. doi: 10.1080/00365518209168384
42. Hustoft HK, Malerod H, Wilson SR, Reubsaet L, Lundanes E, Greibrokk T (2012) A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. Integr Proteomics. doi: 10.5772/2473
43. Klammer AA, MacCoss MJ (2006) Effects of Modified Digestion Schemes on the Identification of Proteins from Complex Mixtures. J Proteome Res 5:695–700. doi: 10.1021/pr050315j [PubMed: 16512685]
44. van den Broek I, Niessen WMA, van Dongen WD (2013) Bioanalytical LC–MS/MS of protein-based biopharmaceuticals. J Chromatogr B 929:161–179. doi: 10.1016/j.jchromb.2013.04.030
45. De Livera AM, Dias DA, De Souza D, Rupasinghe T, Pyke J, Tull D, Roessner U, McConville M, Speed TP (2012) Normalizing and Integrating Metabolomics Data. Anal Chem 84:10768–10776. doi: 10.1021/ac302748b [PubMed: 23150939]
46. Go Y-M, Chandler JD, Jones DP (2015) The cysteine proteome. Free Radic Biol Med 84:227–245. doi: 10.1016/j.freeradbiomed.2015.03.022 [PubMed: 25843657]
47. Wagner M, Tonoli D, Varesio E, Hopfgartner G (2016) The use of mass spectrometry to analyze dried blood spots. Mass Spectrom Rev 35:361–438. doi: 10.1002/mas.21441 [PubMed: 25252132]
48. Björkesten J, Enroth S, Shen Q, Wik L, Hougaard DM, Cohen AS, Sörensen L, Giedraitis V, Ingelsson M, Larsson A, Kamali-Moghaddam M, Landegren U (2017) Stability of Proteins in Dried Blood Spot Biobanks. Mol Cell Proteomics 16:1286–1296. doi: 10.1074/mcp.RA117.000015 [PubMed: 28501802]
49. Fennell TR, MacNeela JP, Morris RW, Watson M, Thompson CL, Bell DA (2000) Hemoglobin adducts from acrylonitrile and ethylene oxide in cigarette smokers: Effects of glutathione s-transferase T1-null and M1-null genotypes. Cancer Epidemiol Biomarkers Prev 9:705–712. [PubMed: 10919741]
50. Carlsson H, von Stedingk H, Nilsson U, Törnqvist M (2014) LC–MS/MS Screening Strategy for Unknown Adducts to N-Terminal Valine in Hemoglobin Applied to Smokers and Nonsmokers. Chem Res Toxicol 27:2062–2070. doi: 10.1021/tx5002749 [PubMed: 25350717]
51. Talhout R, Schulz T, Florek E, Van Benthem J, Wester P, Opperhuizen A (2011) Hazardous Compounds in Tobacco Smoke. Int J Environ Res Public Health 8:613–628. doi: 10.3390/ijerph8020613 [PubMed: 21556207]
52. Aldini G, Dalle-Donne I, Facino RM, Milzani A, Carini M (2007) Intervention strategies to inhibit protein carbonylation by lipoxidation-derived reactive carbonyls. Med Res Rev 27:817–868. doi: 10.1002/med.20073 [PubMed: 17044003]
53. Colombo G, Aldini G, Orioli M, Giustarini D, Gornati R, Rossi R, Colombo R, Carini M, Milzani A, Dalle-Donne I (2010) Water-Soluble α,β-Unsaturated Aldehydes of Cigarette Smoke Induce Carbonylation of Human Serum Albumin. Antioxid Redox Signal 12:349–364. doi: 10.1089/ars.2009.2806 [PubMed: 19686037]
54. Weiner D, Levy Y, Khankin EV, Reznick AZ (2008) Inhibition of salivary amylase activity by cigarette smoke aldehydes. J Physiol Pharmacol 59 Suppl 6:727–37. [PubMed: 19218700]
55. Vinnakota CV, Peetha NS, Perrizo MG, Ferris DG, Oda RP, Rockwood GA, Logue BA (2012) Comparison of cyanide exposure markers in the biofluids of smokers and non-smokers. Biomarkers 17:625–633. doi: 10.3109/1354750X.2012.709880 [PubMed: 22889346]
56. Fasco MJ, Stack RF, Lu S, Hauer CR, Schneider E, Dailey M, Aldous KM (2011) Unique Cyanide Adduct in Human Serum Albumin: Potential as a Surrogate Exposure Marker. Chem Res Toxicol 24:505–514. doi: 10.1021/tx100344e [PubMed: 21366342]
57. Fasco MJ, Hauer CR, Stack RF, O'Hehir C, Barr JR, Eadon GA (2007) Cyanide Adducts with Human Plasma Proteins: Albumin as a Potential Exposure Surrogate. Chem Res Toxicol 20:677–684. doi: 10.1021/tx6003425 [PubMed: 17373827]

58. DiFranza JR, Aligne CA, Weitzman M (2004) Prenatal and postnatal environmental tobacco smoke exposure and children's health. Pediatrics 113:1007–1015. [PubMed: 15060193]
59. Knopik VS, Maccani MA, Francazio S, McGeary JE (2012) The epigenetics of maternal cigarette smoking during pregnancy and effects on child development. Dev Psychopathol 24:1377–1390. doi: 10.1017/S0954579412000776 [PubMed: 23062304]
60. de Smith AJ, Kaur M, Gonseth S, Endicott A, Selvin S, Zhang L, Roy R, Shao X, Hansen HM, Kang AY, Walsh KM, Dahl GV, McKean-Cowdin R, Metayer C, Wiemels JL (2017) Correlates of Prenatal and Early-Life Tobacco Smoke Exposure and Frequency of Common Gene Deletions in Childhood Acute Lymphoblastic Leukemia. Cancer Res 77:1674–1683. doi: 10.1158/0008-5472.CAN-16-2571 [PubMed: 28202519]
61. Metayer C, Petridou E, Aranguré JMM, Roman E, Schüz J, Magnani C, Mora AM, Mueller BA, de Oliveira MSP, Dockerty JD, McCauley K, Lightfoot T, Hatzipantelis E, Rudant J, Flores-Lujano J, Kaatsch P, Miligi L, Wesseling C, Doody DR, Moschovi M, Orsi L, Mattioli S, Selvin S, Kang AY, Clavel J (2016) Parental Tobacco Smoking and Acute Myeloid Leukemia. Am J Epidemiol 184:261–273. doi: 10.1093/aje/kww018 [PubMed: 27492895]
62. Gonseth S, de Smith AJ, Roy R, Zhou M, Lee S-T, Shao X, Ohja J, Wrensch MR, Walsh KM, Metayer C, Wiemels JL (2016) Genetic contribution to variation in DNA methylation at maternal smoking-sensitive loci in exposed neonates. Epigenetics 11:664–673. doi: 10.1080/15592294.2016.1209614 [PubMed: 27403598]
63. Momen NC, Olsen J, Gissler M, Li J (2016) Exposure to maternal smoking during pregnancy and risk of childhood cancer: a study using the Danish national registers. Cancer Causes Control 27:341–349. doi: 10.1007/s10552-015-0707-0 [PubMed: 26689564]
64. Whitehead TP, Metayer C, Wiemels JL, Singer AW, Miller MD (2016) Childhood Leukemia and Primary Prevention. Curr Probl Pediatr Adolesc Health Care 46:317–352. doi: 10.1016/j.cppeds.2016.08.004 [PubMed: 27968954]
65. Florescu A, Ferrence R, Einarson T, Selby P, Soldin O, Koren G (2009) Methods for Quantification of Exposure to Cigarette Smoking and Environmental Tobacco Smoke: Focus on Developmental Toxicology. Ther Drug Monit 31:14–30. doi: 10.1097/FTD.0b013e3181957a3b [PubMed: 19125149]
66. Holbrook BD (2016) The effects of nicotine on human fetal development. Birth Defects Res Part C 108:181–192. doi: 10.1002/bdrc.21128
67. Spector LG, Murphy SE, Wickham KM, Lindgren B, Joseph AM (2014) Prenatal Tobacco Exposure and Cotinine in Newborn Dried Blood Spots. Pediatrics 133:e1632–e1638. doi: 10.1542/peds.2013-3118 [PubMed: 24819573]
68. Yang J, Pearl M, Jacob P, DeLorenze GN, Benowitz NL, Yu L, Havel C, Kharrazi M (2013) Levels of Cotinine in Dried Blood Specimens from Newborns as a Biomarker of Maternal Smoking Close to the Time of Delivery. Am J Epidemiol 178:1648–1654. doi: 10.1093/aje/kwt182 [PubMed: 24068198]
69. Spector LG, Hecht SS, Ognjanovic S, Carmella SG, Ross JA (2007) Detection of Cotinine in Newborn Dried Blood Spots. Cancer Epidemiol Biomarkers Prev 16:1902–1905. doi: 10.1158/1055-9965.EPI-07-0230 [PubMed: 17855712]
70. Rappaport SM, Waidyanatha S, Qu Q, Shore R, Jin X, Cohen B, Chen LC, Melikian A a., Li G, Yin S, Yan H, Xu B, Mu R, Li Y, Zhang X, Li K (2002) Albumin adducts of benzene oxide and 1,4-benzoquinone as measures of human benzene metabolism. Cancer Res 62:1330–1337. [PubMed: 11888901]

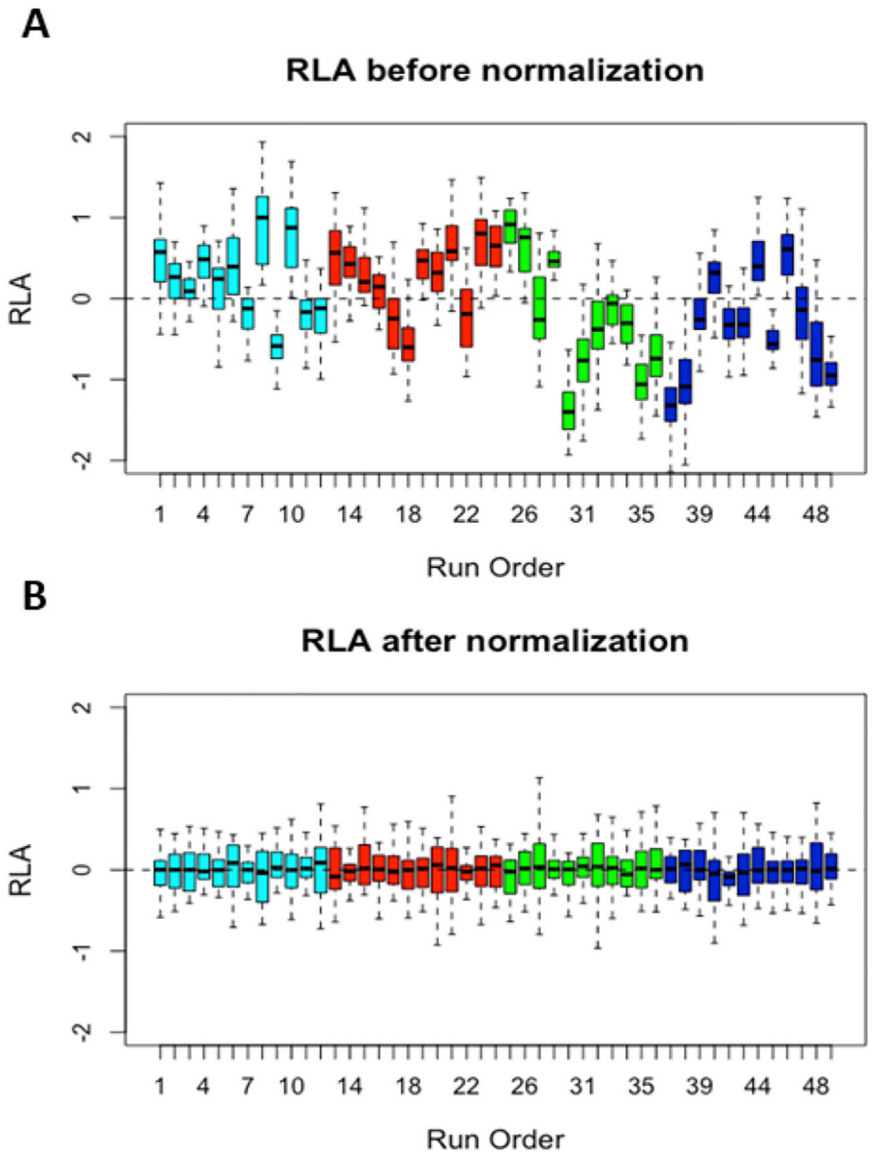

Fig. 1.

Relative log abundance (RLA) plots from (A) before and (B) after normalization with 'scone.' Colors represent sample batches

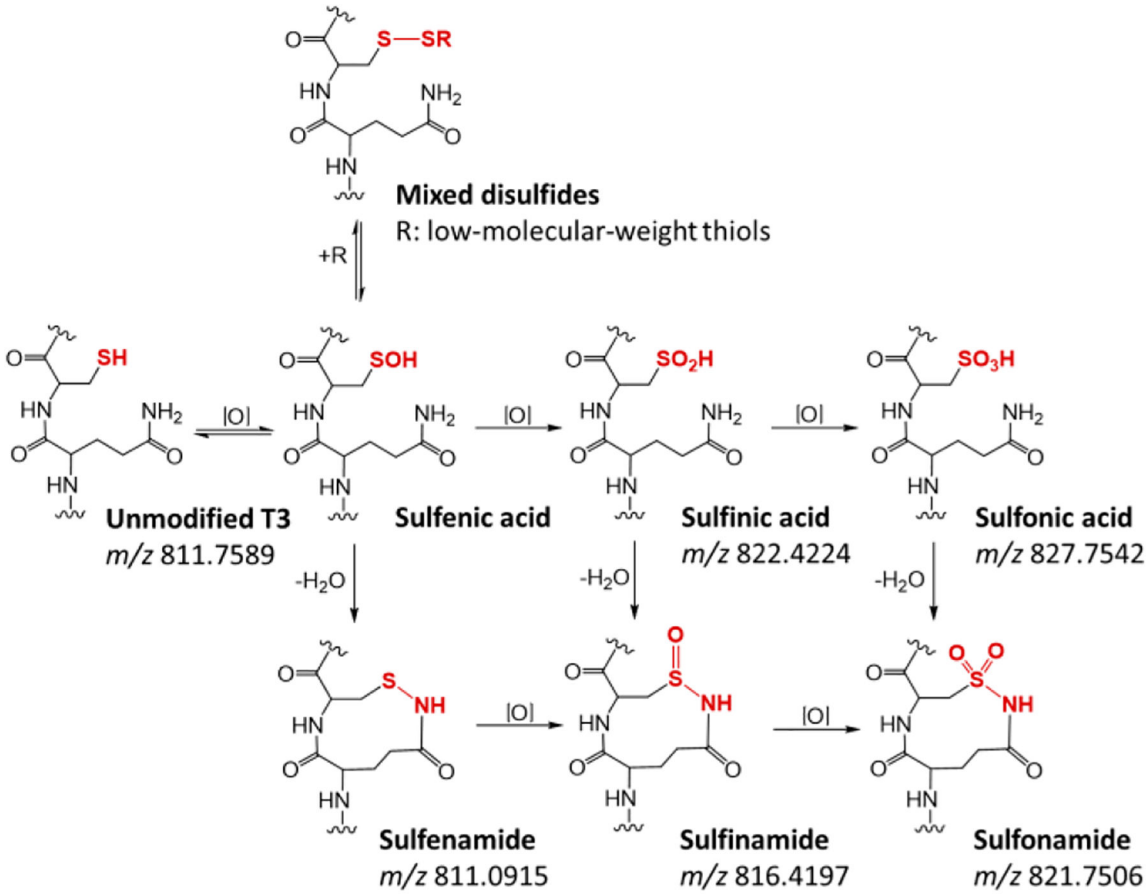

Fig. 2.

Reaction pathways proposed for the formation of Cys34 oxidation products (Adapted from Grigoryan et al. 2012).[18] Observed monoisotopic masses are shown

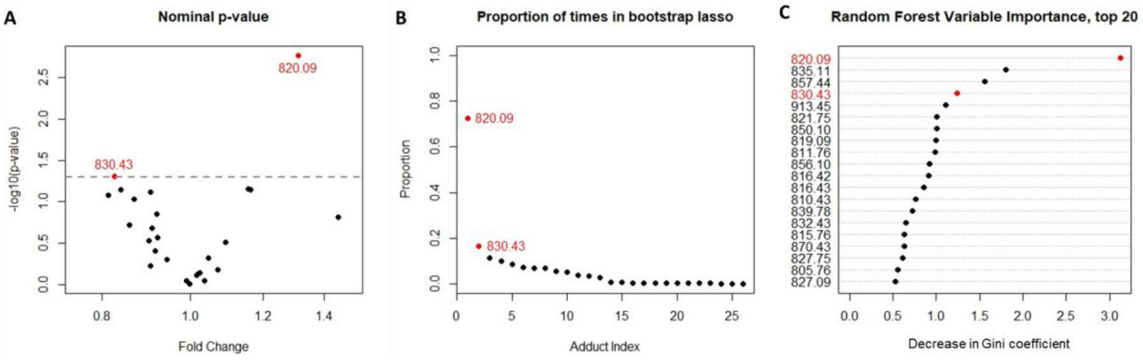

Fig. 3.

Selection of adducts that discriminate newborns by maternal smoking status. Points represent adducts and the two highest-ranked adducts are shown in red (820.09 is the cyanide adduct and 830.43 is an unknown adduct, possibly S-addition of acrolein). (A) Volcano plot showing the mean fold change for each adduct between newborns with smoking/nonsmoking mothers, and the corresponding statistical significance). The dashed line represents a nominal p-value of 0.05. (B) Proportion of times each adduct was selected into the lasso model out of 500 iterations. (C) The top 20 adducts ranked by random forest variable importance

Table 1

Adducts detected in 47 archived DBS from newborns whose mothers smoked during pregnancy and those who did not

Adduct	RT (min)	MIM observed (*m/z*, 3+)	MIM theoretical (*m/z*, 3+)	ΔMass (ppm)	Mass added to Cys34-S$^-$ (Da)	Elemental composition of added mass	Annotation
796.43[b,c,d]	27.08	796.4304	796.4301	0.39	−45.9854	$-CH_2S$	Cys34→Gly
805.76[b,c,d]	26.73	805.7616	805.7618	−0.22	−17.9919	$-SH_2$, +O	Cys34→Oxoalanine
810.43	26.85	810.4335			−3.9761		Unknown
811.09	26.60	811.0915	811.0875	4.98	−2.0020	$-H_2$	Cys34 Sulfenamide
811.76[a,b,c,d]	27.67	811.7589	811.7594	−0.61	1.0078	+H	Unmodified T3[e]
815.76	27.26	815.7593	815.7594	−0.07	12.0013	$+CH_2O$, $-H_2O$	CH_2 crosslink
816.42[a,b,c,d]	25.43	816.4197	816.4191	0.72	13.9825	$-H_2$, +O	Cys34 Sulfinamide[e]
816.43[a,b,c,d]	27.53	816.4312	816.4312	−0.05	15.0248	$+CH_3$	Methylation (not at Cys34)
819.09	27.59	819.0865	819.0867	−0.23	22.9907	+Na	S-Sodiation
820.09[a,b,d]	27.72	820.0914	820.0911	0.31	26.0053	+CN	S-Cyanylation
821.75[a]	27.33	821.7506	821.7507	−0.08	30.9831	$-H_2$, O_2	Cys34 Sulfonamide
822.42[a,b,c,d]	26.70	822.4224	822.4226	−0.28	32.9984	$+HO_2$	Cys34 Sulfinic acid[e]
827.09[a,b]	27.25	827.0947	827.0945	0.24	47.0153	$+CH_3O_2$	S-(O)-O-CH_3
827.75[a,b,c,d]	27.00	827.7542	827.7543	−0.08	48.9938	$+HO_3$	Cys34 Sulfonic acid[e]
830.43[b]	27.20	830.4342	830.4348	−0.74	57.0337	$+C_3H_5O$	Unknown (possible S-addition of acrolein)
832.43	27.43	832.4260	832.4262	−0.24	63.0091	$+CH_3O_3$	Unknown
835.11[b,d]	27.29	835.1064	835.1066	−0.35	71.0503	$+C_4H_7O$	S-Addition of crotonaldehyde[e]
839.78[b]	27.58	839.7777	839.7785	−0.99	85.0643	$+C_5H_9O$	S-Addition of tiglic aldehyde[e]
845.11	27.16	845.1110	845.1102	0.98	101.0642	$+C_5H_9O_2$	Unknown
850.10[b,c,d]	27.51	850.0950	850.0958	−0.88	116.0162	$+C_4H_6NOS$	S-Addition of hCys (−H2O)
851.42[a,b,c,d]	26.12	851.4249	851.4274	−2.90	120.0060	$+C_3H_6NO_2S$	S-Addition of Cys[e]
856.10[a,b,c,d]	26.46	856.0990	856.0993	−0.32	134.0282	$+C_4H8NO_2S$	S-Addition of hCys[e]
857.44	27.67	857.4418	857.4419	−0.10	138.0567	$+C_7H_8NO_2$	Unknown
870.43[a,b,c,d]	25.75	870.4346	870.4345	0.11	177.0351	$+C_5H_9N_2O_3S$	S-Addition of CysGly[e]
913.45[a,b,c,d]	26.06	913.4487	913.4487	−0.06	306.0773	$+C_{10}H_{16}N_3O_6S$	S-Addition of GSH[e]
918.12[a]	26.39	918.1205	918.1206	−0.16	320.0926	$+C_{11}H_{18}N_3O_6S$	Unknown (likely S-addition of GSH and CH_2 crosslink)

Cys, cysteine; hCys, homocysteine; CysGly, cysteinylglycine; GSH, glutathione; MIM, monoisotopic mass; RT, retention time.

[a] Also detected by Liu et al. 2018.[7]

[b] Also detected by Grigoryan et al. 2018.[8]

[c] Also detected by Lu et al. 2017.[6]

[d] Also detected by Grigoryan et al. 2016.[5]

[e] Annotation confirmed with a synthetic standard